ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ
И МАТЕМАТИЧЕСКОЙ ГЕОФИЗИКИ
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК

МАРЧУКОВСКИЕ НАУЧНЫЕ ЧТЕНИЯ – 2019

Труды

Международной конференции
"Актуальные проблемы
вычислительной и прикладной математики"

1–5 июля 2019 г.
Академгородок, Новосибирск, Россия

УДК 519.6
ББК 22.19
М30

М30 Марчуковские научные чтения – 2019 : Труды конференции "Актуальные проблемы вычислительной и прикладной математики" / Ин-т вычислительной математики и матем. геофизики СО РАН. Новосибирск, 1–5 июля 2019 г. Новосибирск : ИВМиМГ СОРАН, 2019. 587 с.

ISBN 978-5-901548-42-4

Целью Международной конференции "Актуальные проблемы вычислительной и прикладной математики" является привлечение специалистов по численному анализу, прикладной математике и вычислительным технологиям к обсуждению актуальных вопросов математики и математического моделирования, а также вопросов практического применения современных численных методов. Основные темы конференции: вычислительная алгебра и методы аппроксимации, численное решение дифференциальных уравнений, методы Монте-Карло и численное статистическое моделирование, математическое моделирование в задачах физики атмосферы, океана, климата и охраны окружающей среды, обратные задачи, математическое моделирование в задачах геофизики и электрофизики, математические модели и методы в науках о Земле, математическое моделирование в информационных технологиях, компьютерная биология.

Конференция проводится при поддержке

Новосибирского государственного университета
Министерства науки и высшего образования Российской Федерации
Сибирского отделения Российской академии наук
Института вычислительной математики им. Г. И. Марчука РАН
ФИЦ Институт цитологии и генетики СО РАН

Спонсор
ЗАО РСК Технологии

Информационная поддержка
Пресс-служба СО РАН

Сайт конференции: http://conf.nsc.ru/amca2019/ru

ISBN 978-5-901548-42-4

© Институт вычислительной математики
и математической геофизики СО РАН, 2019

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СИСТЕМЫ ТЕЛЕУПРАВЛЕНИЯ ЗУР ЗРК МАЛОЙ ДАЛЬНОСТИ

Т. А. Аверина[1,2], И. М. Косачев[3], К. Н. Чугай[4]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

[3] *Военная академия Республики Беларусь, 220057, Минск*

[4] *НИИ Вооруженных Сил Республики Беларусь, 220103, Минск*

УДК 519.676
DOI: 10.24411/9999-016A-2019-10001

В статье построена модель системы со случайной структурой, описывающей математическую модель системы телеуправления ЗУР ЗРК малой дальности (МД) с комплексированной (радиолокационно-лазерной) информационной системой. С помощью разработанного статистического алгоритма проведена тестовая проверка модели
Ключевые слова: структурная схема, система случайной структуры, стохастическое дифференциальное уравнение, статистический алгоритм.

Введение

В настоящее время широкое распространение получили динамические системы со случайными изменениями условий функционирования и возмущений, приводящими к внезапному изменению структуры в целом — к структурной неопределенности. Систематическое изложение таких задач и методов их анализа дано в работах И.Е. Казакова, В.М. Артемьева, В.А. Бухалева [1].

Эти модели появились в ряде отраслей науки и в научных исследованиях, связанных с моделированием сложных явлений и процессов управления. Это, например, задачи автоматизации управления системой, имеющей на неперекрывающихся временных интервалах различные режимы работы и разные структуры. Такие системы получили название динамических систем со случайной сменой структуры (ДССС) или более короткое название — системы со случайной структурой.

Примерами систем со случайной структурой могут служить системы управления сближением летательных аппаратов, системы поиска и захвата информационного сигнала в задачах навигации и управления полетом летательных аппаратов, системы комбинированного наведения на цель, а также системы с возможными нарушениями и отказами

В данной статье рассматривается имитационная математическая модель типовой системы телеуправления ЗУР перспективного ЗРК МД с комплексированной (радиолокационно-лазерной) информационной системой.

1 Построение стохастической модели системы управления ЗУР ЗРК малой дальности

Стохастическую модель системы управления ЗУР ЗРК малой дальности будем рассматривать как математическую модель системы управления (СУ) беспилотным летательным аппаратом (ЛА) в классе ДССС.

Работа выполнена при частичной финансовой поддержке Госзадания 0315-2016-0002.

ISBN 978-5-901548-42-4

Моделируемая система имеет два состояния: основное (первое) — поиск цели и сопровождение ЛА по угловым координатам и дальности с помощью радиолокационной информационной системы; второе — захват цели и ее автосопровождение по угловым координатам и дальности лазерным локатором. При нарушении работы лазерной информационной системы происходит быстрый обратный переход на радиолокационную информационную систему, т. е. система снова возвращается в первое состояние структуры. Переходы из r-й структуры в l-ю задаются либо с помощью интенсивностей перехода $\nu^{(r,l)}$, либо с помощью условных вероятностей $P^{(r,l)}(t, A)$, которые рассчитываются как свертка плотности распределения вероятности (ПРВ) времен «жизни», предшествующих l-й структуры.

ПРВ времен «жизни» ДССС в каждой из возможных структур $f^{(1)}$ и $f^{(2)}$ определяется по результатам натурных (полунатурных) испытаний.

На рис. 1 представлена типовая структурная схема контура наведения телеуправляемого ЛА в вертикальной плоскости с комплексированной (радиолокационно-лазерной) информационной системой. На данном рисунке обозначено:

FF — передаточные функции формирующих фильтров в виде колебательных звеньев, обеспечивающие формирование на своем выходе ошибки измерения разности угловых координат цели и ЛА в вертикальной плоскости в l-й структуре (фазовая координата $Y_{12}^{(l)}(t)$);

$R(t)$ — детерминированная функция, аппроксимирующая текущую наклонную дальность до ЛА;

CGD (УВК) — передаточная функция устройства выработки команд наведения ЛА в вертикальной плоскости;

PM — передаточная функция рулевой машинки ЛА в виде апериодического звена;

$3UP$ — передаточная характеристика ЛА в виде колебательного звена с нестационарными параметрами;

DG (ДГ) — передаточная функция демпфирующего гироскопа;

$K3$ — передаточная функция нестационарного кинематического звена контура телеуправления ЛА.

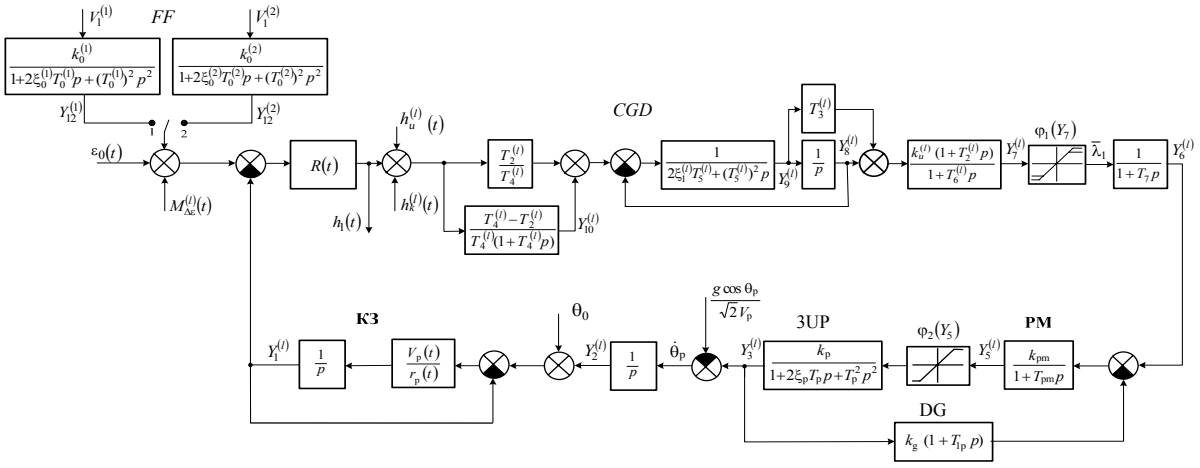


Рис. 1: Структурная схема контура наведения телеуправляемой ЛА в вертикальной плоскости

При разработке структурной схемы контура телеуправления ЛА, представленной на рис. 1, были приняты следующие допущения и ограничения:

1. рассматривается плоскостная модель контура телеуправления ЛА лишь в одной (вертикальной) плоскости наведения;

2. рассматривается случай сильного радиолокационного и лазерного сигналов, отраженных от цели, при котором отношение мощности сигнала к мощности помехи гораздо больше единицы;

3. импульсный характер передаваемых на борт ЛА команд управления не учитывается, так как контур телеуправления является узкополосной следящей системой;

4. ошибки измерения разности угловых координат одиночной цели и в вертикальной и горизонтальной плоскостях являются гауссовыми случайными процессами с экспоненциально-колебательными корреляционными функциями вида $K_y^{(l)}(\tau) = D_y^{(l)} e^{\alpha^{(l)}|\tau|} \cos \beta^{(l)}|\tau|$. При исследовании только контура телеуправления ЛА это позволяет заменить радиолокационную и лазерную системы измерения разности

угловых координат цели и ракеты двумя эквивалентными формирующими фильтрами в виде колебательных звеньев. Параметры этих формирующих фильтров рассчитываются исходя из значений параметров $D_y^{(l)}$, $\alpha^{(l)}$, $\beta^{(l)}$ входящих в экспоненциально-колебательные корреляционные функции для каждой структуры системы управления.

Входным задающим воздействием в рассматриваемом контуре телеуправления ЛА является текущий угол места цели в вертикальной плоскости $\varepsilon_0(t)$.

Детерминированная функция $M_{\Delta\varepsilon}^{(l)}(t)$ учитывает нестационарное математическое ожидание ошибки измерения разности угловых координат цели и ЛА в вертикальной плоскости наведения в l-й структуре.

С помощью детерминированной функции $h_{\text{u}}^{(l)}(t)$ формируется упреждающая кинематическая траектория полета ЛА относительно линии визирования цели.

Функция $h_{\text{k}}^{(l)}(t)$ учитывает компенсацию динамической ошибки контура наведения ЛА в l-й структуре.

Функция $g\cos(\theta_h)/(\sqrt{2}V_p)$ учитывает влияние массы ЛА на скорость изменения угла тангажа (угла наклона вектора скорости ЛА к горизонту).

Фазовые координаты вводятся только на выходах инерционных звеньев: интеграторов, апериодических, колебательных. Физический смысл фазовых координат (см. рис. 1) исследуемой системы управления ЛА пояснен ниже.

По известным правилам была составлена система со случайной структурой для каждого l-го ($l = 1, 2$) состояния структуры:

$$\frac{dy_1^{(l)}}{dt} = \frac{V_p(t)}{r_p(t)}(\theta_0 - y_1^{(l)} + y_2^{(l)}),$$

$$\frac{dy_2^{(l)}}{dt} = -g\frac{\cos(\theta_p)}{\sqrt{2}V_p(t)} + y_3^{(l)},$$

$$\frac{dy_3^{(l)}}{dt} = y_4^{(l)},$$

$$\frac{dy_4^{(l)}}{dt} = -\frac{1}{T_p^2(t)}y_3^{(l)} - \frac{2\xi_p(t)}{T_p(t)}y_4^{(l)} + \frac{k_p(t)}{T_p^2(t)}\varphi_2^{(l)}(y_5),$$

$$\frac{dy_5^{(l)}}{dt} = -\frac{k_g k_{pm}}{T_{pm}}(y_3^{(l)} + T_{1p}y_4^{(l)}) + \frac{1}{T_{pm}}(k_{pm}y_6^{(l)} - y_5^{(l)} + k_{pm}V_n(t)),$$

$$\frac{dy_6^{(l)}}{dt} = \frac{1}{T_7}(-y_6^{(l)} + k_u^{(l)}\varphi_1^{(l)}(y_7) + V_n(t)) + \frac{k_u^{(l)}T_2^{(l)}}{T_6^{(l)}T_7}\varphi_1^{(l)}(y_7)(y_8^{(l)} + T_3^{(l)}y_9^{(l)}),$$

$$\frac{dy_7^{(l)}}{dt} = \frac{k_u^{(l)}(T_2^{(l)})^2T_3^{(l)}}{T_4^{(l)}(T_5^{(l)})^2T_6^{(l)}}\Big(R(t)[\varepsilon_0(t) + M_{\Delta\varepsilon}^{(l)} - y_1^{(l)} + y_{12}^{(l)}] + h_u^{(l)} + h_k^{(l)}\Big) + \frac{k_u^{(l)}(T_5^{(l)})^2 - T_2^{(l)}T_3^{(l)}T_6^{(l)}}{(T_5^{(l)})^2T_6^{(l)}}y_8^{(l)} - \frac{1}{T_6^{(l)}}y_7^{(l)} + \frac{k_u(l)(T_2^{(l)}+T_3^{(l)})T_5^{(l)} - 2\xi_1T_2^{(l)}T_3^{(l)}}{T_5^{(l)}T_6^{(l)}}y_9^{(l)} + \frac{k_u^{(l)}T_2^{(l)}T_3^{(l)}}{(T_5^{(l)})^2T_6^{(l)}}y_{10}^{(l)},$$

$$\frac{dy_8^{(l)}}{dt} = y_9^{(l)},$$

$$\frac{dy_9^{(l)}}{dt} = \frac{T_2^{(l)}}{T_4^{(l)}(T_5^{(l)})^2}\Big(R(t)[\varepsilon_0(t) + M_{\Delta\varepsilon}^{(l)} - y_1^{(l)} + y_{12}^{(l)}] + h_u^{(l)} + h_k^{(l)}\Big) - \frac{2\xi_1^{(l)}}{T_5^{(l)}}y_9^{(l)} + \frac{1}{(T_5^{(l)})^2}(y_{10}^{(l)} - y_8^{(l)}),$$

$$\frac{dy_{10}^{(l)}}{dt} = \frac{T_4^{(l)} - T_2^{(l)}}{(T_4^{(l)})^2}[R(t)(\varepsilon_0(t) + M_{\Delta\varepsilon}^{(l)} - y_1^{(l)} + y_{12}^{(l)}) + h_u^{(l)} + h_{k1}^{(l)}] - \frac{1}{T_4^{(l)}}y_{10}^{(l)},$$

$$\frac{dy_{11}^{(l)}}{dt} = -\frac{1}{(T_0^{(l)})^2}y_{11}^{(l)} - \frac{2\xi_0^{(l)}}{T_0^{(l)}}y_{11}^{(l)} + \frac{k_0^{(l)}}{(T_0^{(l)})^2}V_1(t),$$

$$\frac{dy_{12}^{(l)}}{dt} = y_{11}^{(l)},$$

$y_1^{(l)}$ — угол места ЛА на выходе кинематического звена в l-й структуре СУ;
$y_2^{(l)}$ — угол тангажа в l-й структуре СУ;
$y_3^{(l)}$ — угловая скорость вращения вектора скорости ЛА в вертикальной плоскости в l-й структуре СУ;
$y_4^{(l)}$ — угловое ускорение вращения вектора скорости ЛА в вертикальной плоскости в l-й структуре СУ;

$y_5^{(l)}$ — угол отклонения рулей высоты ЛА в l-й структуре СУ;
$y_6^{(l)}$ — команда управления ЛА на выходе бортового дешифратора команд в l-й структуре СУ;
$y_7^{(l)}$ — команда управления ЛА в вертикальной плоскости на выходе УКВ (до ограничителя), передаваемая на борт ЛА в l-й структуре СУ;
$y_8^{(l)}$ — сигнал на выходе второго формирующего фильтра УКВ в l-й структуре СУ;
$y_9^{(l)}$ — производная сигнала на выходе второго формирующего фильтра УКВ в l-й структуре СУ;
$y_{10}^{(l)}$ — сигнал на выходе первого формирующего фильтра УКВ в l-й структуре СУ;
$y_{11}^{(l)}$ — производная ошибки измерения разности углов места цели и ЛА на выходе ФФ в l-й структуре СУ;
$y_{12}^{(l)}$ — ошибка измерения разности углов места цели и ЛА на выходе ФФ в l-й структуре СУ.

Данные СДУ являются нестационарными, нелинейными с аддитивными белыми шумами. Нелинейность: $\varphi_1^{(l)}(y_7)$ ограничивает величину команды управления, а $\varphi_2^{(l)}(y_5)$ — угол отклонения рулей ЛА.

Входным задающим воздействием для исследуемой системы управления ЛА является текущий угол места цели $\varphi_0(t)$.

2 Статистическое моделирование системы управления

При смене структур рассматриваются естественные условия восстановления. Исключение составляет $y_{12}^{(l)}(t)$, которая имеет восстановление в окрестности функций $M_{\Delta\varepsilon}^{(1)}(t)$ и $M_{\Delta\varepsilon}^{(2)}(t)$). Искомой случайной величиной исследуемой СУ является промах ЛА в вертикальной плоскости в l-й структуре, который рассчитывается по формуле

$$h_1^{(l)} = (\varepsilon_0 + M_{\Delta\varepsilon}^{(l)}(t) - y_1^{(l)} + y_{12}^{(l)})R(t_v),$$

где t_v — точка встречи.

Показателем эффективности функционирования системы управления ЛА является вероятность попадания в круг заданного относительно цели радиуса RPB. При нормальной ПРВ промаха рассчитывается по следующей формуле:

$$P(-R_{pb} \le h_1 \le R_{pb}) = \int_{-R_{pb}}^{R_{pb}} f(h_1)dh_1 = \Phi\Big(\frac{R_{pb} - M_{h1}}{\sigma_{h1}}\Big) + \Phi\Big(\frac{R_{pb} - M_{h1}}{\sigma_{h1}}\Big),$$

где h_1 — итоговый (с учетом двух состояний структуры) промах ЛА в вертикальной плоскости наведения ракеты; M_{h_1} — итоговое математическое ожидание промаха ЛА в вертикальной плоскости наведения; σ_{h_1} — итоговое среднеквадратическое отклонение промаха ЛА в вертикальной плоскости наведения; $\Phi(u)$ — функция Лапласа.

Статистическое моделирование решений системы со случайной структурой позволяет оценить различные вероятностные характеристики решения, в том числе и ПРВ промаха. Исследование разработанной стохастической модели в классе ДССС осуществлялось на ПК с процессором Intel Core i5 3330 (3.00 ГГц). с использованием алгоритма статистического моделирования систем со случайной структурой с распределенными переходами.

Модифицированный алгоритм численного моделирования перехода из l-го состояния для систем с условной марковской структурой при распределенных переходах:

(1) Пусть в момент t_k система находилась в l-м состоянии и вектор состояния равен Y_k.

(2) Моделируем вспомогательную случайную величину α_1 — равномерную на интервале $(0, 1)$ и заводим счетчик, полагая $z = 1$.

(3) Моделируем возможный момент выхода из l-го состояния: $t_{k+1} = t_k + \tau$, где τ — случайная величина с плотностью распределения $p(x) = \nu_l^m * exp(-\nu_l^m x)$, $\nu_l^m = \sum\limits_{i \ne l} \nu_{li}^m$, (по формуле $\tau = -ln\alpha/\nu_l^m$, α — равномерная на интервале $(0, 1)$ случайная величина).

(4) Вычисляем возможный r-й номер новой структуры, распределенный с вероятностью $p_r = \frac{\nu_{lr}^m}{\nu_l^m}$, $r \ne l$, $r = 1, ...N_0$.

(5) На интервале $[t_k, t_{k+1}]$ численным методом для СДУ решаем уравнения ДССС для l-й структуры, находим Y_{k+1}- вектор состояния системы в момент времени t_{k+1} (шаг численного метода должен быть согласован с интенсивностью перехода, например, $h \leq 0.1/\nu_l^m$).

(6) Полагаем $t_k := t_{k+1}$, $Y_k := Y_{k+1}$.

(7) Полагаем $z := z * (1 - \nu_{lr}(\mathbf{Y}_k, t_k)/\nu_{lr}^m)$. Проверяем условие смены структуры: если $1 - \alpha_1 > z$, то идем на 8); иначе идем на 3).

(8) Меняем номер структуры на r-й; вычисляем Y_k согласно заданной условной плотности восстановления q_{lr}.

Следует отметить, что данный алгоритм построен для произвольных систем с условной марковской структурой при распределенных переходах [2]. В более простом, частном случае, когда интенсивности перехода постоянны $\nu_{lr}(Y,t) = \nu_{lr} = const$ (такие системы называются системами с независимой марковской структурой при распределенных переходах), в алгоритме будут отсутствовать пункты 2 и 7.

Для моделирования равномерно распределенных случайных величин на интервале $(0,1)$ использовался «генератор» псевдослучайных чисел RAND128 [3] (с модулем 2^{128} и множителем 5^{100109}). При тестовых расчетах задавались следующие безразмерные параметры модели

$$
\begin{gathered}
k_0 = 0,707,\ \xi_0 = 0,27, \\
T_0^{(1)} = 0.76,\ T_0^{(2)} = 0.37, \\
k_u^{(1)} = 1,\ k_u^{(2)} = 3, \\
\xi_1 = 0.8,\ k_{pM} = 0.5, \\
T_{pM} = 0.035,\ k_g = 0.375, \\
T_2^{(1)} = 1.8,\ T_2^{(2)} = 1.0, \\
T_3^{(1)} = 0.33, T_3^{(2)} = 0.174, \\
T_4^{(1)} = 0.14,\ T_4^{(2)} = 0.0725, \\
T_5^{(1)} = 0.14,\ T_5^{(2)} = 0.063, \\
T_6^{(1)} = 5.0,\ T_6^{(2)} = 2.5, \\
T_7 = 0,04,\ T_{1p}(t) = k_0 + k_1 t, \\
T_p(t) = k_4 + k_5 t,\ k_p(t) = k_2 + k_3 t, \\
\xi_p(t) = k_6 + k_7 t, \\
r_p(t) = k_{11} + k_8 t + \tfrac{k_9}{2} t^2 + \tfrac{k_{10}}{3} t^3, \\
V_p(t) = k_8 + k_9 t + k_{10} t^2.
\end{gathered}
$$

Заключение

В данной статье разработана стохастическая модель системы со случайной структурой, описывающей математическую модель системы телеуправления ЗУР ЗРК малой дальности (МД) с комплексированной (радиолокационно-лазерной) информационной системой. Эта система построена как система управления беспилотным ЛА в классе ДССС.

Тестовые расчеты данной модели статистическим алгоритмом показывают ее адекватность на безразмерных параметрах, согласованных с результатами натурных и полунатурных испытаний [4]. В дальнейшем предполагается

1. разработать аналитическую математическую модель этого же контура телеуправления ЛА согласно методике [5];

2. провести дополнительные исследования модифицированным статистическим алгоритмом [6] стохастической модели СУ, построенной в данной статье и

3. сравнить результаты расчетов аналитического и статистического моделирования.

Список литературы

[1] Казаков, И.Е., Артемьев В. М., Бухалев. В. А. Анализ систем случайной структуры. М. : Физматлит, 1993. — 272 с.

[2] Averina T.A. (2013). A modified algorithm for statistical simulation of random structure systems with distributed transitions. *Numerical Analysis and Applications*. Vol. **6**, pp. 91–97.

[3] Mikhailov G.A., Marchenko M.A. (2002). Parallel realization of statistical simulation and random number generators. *Russ. J. Numer. Anal. Math. Modelling*. Vol. **17**, pp. 113–124.

[4] Новые высокоэффективные методы наведения телеуправляемых ЗУР / И. М. Косачев, О. Г. Бойцов, В. В. Меликаев // Сб. докл. Междунар. воен.-науч. конф., Минск, 10–16 марта 2000 г. — Минск : ВА РБ, 2000. — С. 176–191.

[5] Косачев, И. М., Ерошенков М.Г. Аналитическое моделирование стохастических систем. Минск : Наука и техника, 1993. — 264 с.

[6] Averina T.A. (2016). A randomized maximum cross-section method to simulate random structure systems with distributed transitions. *Numerical Analysis and Applications*. Vol. **9**, pp. 179–190.

Аверина Татьяна Александровна — к.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: ata@osmf.sscc.ru;
Косачев Иван Михайлович — д.т.н., гл. науч. сотр. Военной академии Республики Беларусь
e-mail: kosachev1301@mail.ru;
Чугай Константин Николаевич — к.т.н. НИИ Вооруженных Сил Республики Беларусь;
e-mail: konstantin.ch40@gmail.com.
Дата поступления — 30 апреля 2019 г.

ПРИМЕНЕНИЕ МЕТОДА МАКСИМАЛЬНОГО СЕЧЕНИЯ В ЗАДАЧАХ ОЦЕНИВАНИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

Т. А. Аверина[1,2], К. А. Рыбаков[3]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

[3] *Московский авиационный институт (национальный исследовательский университет), 125993, Москва*

УДК 519.676
DOI: 10.24411/9999-016A-2019-10002

В работе рассматривается задача оценивания случайных процессов в динамических системах, математическая модель которых описывается стохастическими дифференциальными уравнениями с пуассоновской составляющей. Решение задачи оценивания (фильтрации, интерполяции, экстраполяции) предполагает моделирование траекторий решения стохастического дифференциального уравнения. Процедура моделирования траекторий включает моделирование пуассоновского потока, для которого применен метод максимального сечения.

Ключевые слова: стохастические дифференциальные уравнения, пуассоновский поток, оценивание, фильтрация, интерполяция, экстраполяция, метод частиц, метод максимального сечения, статистическое моделирование.

Введение

Во многих практических задачах возникает необходимость оценивания параметров или состояния динамической системы по косвенным измерениям на фоне случайных помех. Такие задачи возникают при управлении подвижными объектами [6,7], при обработке навигационной информации [17,18], в радиотехнике [19,20], анализе финансовых рынков [23]. В этой работе задача оценивания решается для систем с непрерывным временем при описании процессов стохастическими дифференциальными уравнениями (СДУ) с пуассоновской составляющей, т.е. рассматриваются системы диффузионно-скачкообразного типа [22,24].

Оценивание параметров или состояния динамической системы осуществляется по результатам измерений как для текущего момента времени (задача фильтрации), так и с запаздываем или опережением по времени (задачи интерполяции и экстраполяции соответственно). В качестве метода решения перечисленных задач используется метод частиц, предполагающий моделирование ансамбля траекторий динамической системы. Ранее в [15] задача фильтрации для систем диффузионно-скачкообразного типа решалась методом частиц (сформирован алгоритм фильтра частиц), однако для моделирования пуассоновского потока применялся наиболее простой метод. Здесь алгоритмы оценивания предлагается формировать на основе метода максимального сечения и его модификации [2,9,10]. Эти методы обеспечивают меньшую трудоемкость при моделировании пуассоновского потока.

1 Оценивание случайных процессов в стохастических системах диффузионно-скачкобразного типа

Будем рассматривать модель стохастической системы наблюдения с непрерывным временем, которая включает уравнение модели объекта наблюдения и уравнение измерительной системы:

$$dX(t) = f\big(t, X(t)\big)dt + \sigma\big(t, X(t)\big)dW(t) + \int_{\Xi} v\big(t, X(t^-), \xi\big)\nu(dt \times d\xi), \quad X(0) = X_0, \tag{1}$$

Работа выполнена в рамках госзадания ИВМиМГ СО РАН (проект 0315-2016-0002) и при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-08-00530-а).

ISBN 978-5-901548-42-4

$$dY(t) = c\big(t, X(t)\big)dt + \zeta(t)dV(t), \quad Y(0) = Y_0 = 0. \tag{2}$$

Уравнение (1) является СДУ в смысле Ито с пуассоновской составляющей. В нем используются следующие обозначения: $X \in \mathbb{R}^n$ — n-мерный вектор состояния, $t \in \mathrm{T} = [0, T]$, T — заданный отрезок времени функционирования системы; $f(t,x)\colon \mathrm{T} \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma(t,x)\colon \mathrm{T} \times \mathbb{R}^n \to \mathbb{R}^{n\times s}$, $v(t,x,\xi)\colon \mathrm{T} \times \mathbb{R}^n \times \Xi \to \mathbb{R}^n$, $\Xi = \mathbb{R}^q$; $W(t)$ — s-мерный стандартный винеровский случайный процесс, ν — пуассоновская случайная мера на $\mathrm{T} \times \Xi$ с характеристической мерой Π_ν, заданной функцией $\pi(t,x,\xi)\colon \mathrm{T} \times \mathbb{R}^n \times \Xi \to \mathbb{R}_+$. Закон распределения начального вектора состояния X_0 задан. Начальный вектор состояния X_0, винеровский процесс $W(t)$ и мера ν независимы. В частном случае X_0 — детерминированный n-мерный вектор.

В СДУ (2) используются следующие обозначения: $Y \in \mathbb{R}^m$ — m-мерный вектор измерений; $c(t,x)\colon \mathrm{T} \times \mathbb{R}^n \to \mathbb{R}^m$, $\zeta(t)\colon \mathrm{T} \to \mathbb{R}^{m\times d}$; $V(t)$ — d-мерный стандартный винеровский случайный процесс, не зависящий от винеровского процесса $W(t)$ и меры ν. При этом предполагается, что матрица $\eta(t) = \zeta(t)\zeta^{\mathrm{T}}(t)$ невырождена, т.е. $\det\eta(t) \neq 0$ и существует обратная матрица $\eta^{-1}(t)$ при любом $t \in \mathrm{T}$.

В наиболее общем случае для математической модели объекта наблюдения и измерительной системы процессы $X(t)$ и $Y(t)$ входят в правые части уравнений (1), (2), а также (2) является СДУ в смысле Ито с пуассоновской составляющей [12].

Траектории случайного процесса $X(t)$ имеют разрывы, образующие пуассоновский поток $\{\tau_j\}$, $\tau_0 = 0$. Введем обозначения для интенсивности этого потока: $\lambda(t,x)\colon \mathrm{T}\times\mathbb{R}^n \to \mathbb{R}_+$, и плотности вероятности скачков (случайных приращений вектора состояния в моменты разрывов): $\psi(t,\delta)\colon \mathrm{T}\times\mathbb{R}^n \to \mathbb{R}_+$. Эти две характеристики определяют характеристическую меру Π_ν, пуассоновскую случайную меру ν и функцию $v(t,x,\xi)$ [4]. Таким образом, при малых Δt

$$\Pr\big(P(t+\Delta t) - P(t) = 1 \,|\, X(t) = x\big) = \lambda(t,x)\Delta t + o(\Delta t),$$

где $P(t)$ — пуассоновский процесс, считающий число разрывов для процесса $X(t)$ на отрезке времени $[0,t]$, Pr — вероятность, а плотность вероятности $\psi(\tau_j,\delta)$ определяет величину скачка Δ_j в момент τ_j:

$$X(\tau_j) = X(\tau_j^-) + \Delta_j, \quad j = 1, 2, \ldots \tag{3}$$

Перечисленные функции $f(t,x)$, $\sigma(t,x)$, $\lambda(t,x)$, $\psi(t,\delta)$, $c(t,x)$ и $\zeta(t)$ заданы, они удовлетворяют условиям существования и единственности решений СДУ с пуассоновской составляющей [8, 24], кроме того, $\mathrm{E}|X_0|^2 < \infty$, где E означает математическое ожидание.

Задача оптимального оценивания для стохастической системы наблюдения (1), (2) заключается в нахождении оценки $\hat{X}(\theta)$ вектора состояния по доступным к текущему моменту времени t измерениям $Y_0^t = \{Y(\tau),\ \tau \in [0,t]\}$ в виде $\hat{X}(\theta) = \Phi(\theta, Y_0^t)$, т.е. требуется определить функцию $\Phi(\theta,\,\cdot\,)$, преобразующую измерения Y_0^t в оценку вектора состояния объекта наблюдения, $\theta \in \mathrm{T}$.

Функцию $\Phi(\theta,\,\cdot\,)$ можно найти из условия оптимальности оценки, т.е. исходя из заданного критерия качества, обеспечивающего минимум среднего риска [11,16]. Зададим квадратичную функцию потерь $\Pi(\varepsilon) = \varepsilon^{\mathrm{T}} L\varepsilon$, для которой L — положительно определенная матрица весовых коэффициентов размера $n \times n$, тогда решение задачи минимизации $\mathrm{E}\Pi(\mathcal{E}(\theta)) \to \min$, где минимум берется по всем допустимым функциям $\Phi(\theta,\,\cdot\,)$ и $\mathcal{E}(\theta) = X(\theta) - \hat{X}(\theta)$ — вектор ошибки оценивания, записывается в форме

$$\hat{X}(\theta) = \Phi(\theta, Y_0^t) = \mathrm{E}[X(\theta) \,|\, Y_0^t],$$

т.е. оптимальная оценка представляет собой условное математическое ожидание вектора состояния, она обеспечивает минимум среднеквадратического отклонения ошибки $\mathcal{E}(\theta)$.

При $\theta = t$ задача оценивания называется задачей фильтрации, при $\theta < t$ — задачей интерполяции (задачей сглаживания) и при $\theta > t$ — задачей экстраполяции (задачей прогнозирования).

Задачу интерполяции разделяют на интерполяцию в фиксированной точке ($\theta = \mathrm{const}$), интерполяцию на фиксированном промежутке ($\theta \in [\theta_*, \theta^*] \subseteq [0,t]$) и интерполяцию с постоянным запаздыванием ($t - \theta = \mathrm{const}$). Используются термины «прямая интерполяция» и «обратная интерполяция». Аналогичная классификация применяется и для задач экстраполяции (с опережением по времени вместо запаздывания). Более подробно эти задачи описаны в [16, 19, 20].

2 Метод частиц для оценивания случайных процессов

Для решения задачи оптимальной фильтрации в работе [13] был предложен алгоритм, построенный на моделировании специального случайного процесса с обрывами и ветвлениями траекторий. Траектории такого

процесса между моментами времени ветвлений или между ветвлением и обрывом определяются уравнением объекта наблюдения (1), а измерения, описываемые уравнением (2), управляют интенсивностью появления обрывов и ветвлений (вероятность обрыва растет, если траектория далека от оптимальной оценки; вероятность ветвления растет, если траектория близка к оптимальной оценке). Затем этот алгоритм был дополнен таким образом, чтобы можно было решать не только задачу фильтрации, но и задачу экстраполяции [5]: на этапе экстраполяции траектории определяются только уравнением (1), а обрывы и ветвления для них невозможны. Решить задачу интерполяции таким образом сложнее, поэтому воспользуемся методом частиц [21].

В методе частиц предлагается наряду с траекториями объекта наблюдения (1) моделировать траектории весовой функции по определенному правилу. Тогда вклад каждой траектории объекта наблюдения в искомую оценку вектора состояния зависит от соответствующей траектории весовой функции, которая формируется на основе измерений (2) (вес уменьшается, если траектория далека от оптимальной оценки; вес растет, если траектория близка к оптимальной оценке). При решении задачи оптимальной фильтрации метод частиц составляет основу для фильтра частиц [21]. Связь специального случайного процесса с обрывами и ветвлениями траекторий и весовой функцией показана в работе [14]. В этой работе предлагается применить метод частиц не только в задаче фильтрации, но и для решения задач интерполяции и экстраполяции.

В небольшой по объему статье не удастся привести подробные алгоритмы фильтрации, интерполяции и экстраполяции, особенно с учетом того, что последние две задачи могут иметь различные формулировки, упомянутые в предыдущем разделе. Поэтому остановимся только на основных моментах: моделировании траекторий стохастических систем диффузионно-скачкобразного типа с применением метода максимального сечения, моделировании траекторий весовой функции и нахождении оптимальной оценки.

Определим весовую функцию согласно [21]:

$$\omega(t) = \exp\biggl\{\int_0^t c^{\mathrm{T}}\bigl(\tau, X(\tau)\bigr)\eta^{-1}(\tau)dY(\tau) - \frac{1}{2}\int_0^t c^{\mathrm{T}}\bigl(\tau, X(\tau)\bigr)\eta^{-1}(\tau)c\bigl(\tau, X(\tau)\bigr)d\tau\biggr\}. \tag{4}$$

Далее, введем равномерную сетку $\{t_k\}$ с заданным постоянным шагом h, определяющую разбиение отрезка времени T:

$$t_{k+1} = t_k + h, \quad k = 1, 2, \ldots, N; \quad t_0 = 0, \quad t_N = T, \quad N = \frac{T}{h}.$$

Обозначим через X_k численное решение СДУ [2, 23], которое отличается от уравнения (1) отсутствием пуассоновской составляющей ($\lambda(t, x) \equiv 0$), в момент времени t_k. Например, для стохастического метода Эйлера имеем

$$X_{k+1} = X_k + hf(t_k, X_k) + \sqrt{h}\sigma(t_k, X_k)\zeta_k, \tag{5}$$

ζ_k — s-мерный случайный вектор, координаты которого независимы и имеют стандартное нормальное распределение для всех k.

Приближенный аналог соотношения (4) имеет вид

$$\omega_{k+1} = \omega_k \exp\biggl\{c^{\mathrm{T}}(t_k, X_k)\eta^{-1}(t_k)\bigl(Y(t_{k+1}) - Y(t_k)\bigr) - \frac{1}{2}c^{\mathrm{T}}(t_k, X_k)\eta^{-1}(t_k)c(t_k, X_k)h\biggr\}, \quad \omega_0 = 1. \tag{6}$$

Для моделирования моментов времени разрывов траекторий используется метод максимального сечения [1–3, 9]. Если существует λ^*, для которого $\lambda(t) \leqslant \lambda^* = \mathrm{const}$, то моделирование времени τ, через которое произойдет разрыв траектории, осуществляется по правилу

$$\tau = \theta_N, \quad N = \min\biggl\{\vartheta\colon\ \alpha_\vartheta \leqslant \frac{\lambda(t_* + \theta_\vartheta)}{\lambda^*}\biggr\}, \quad \theta_\vartheta = \sum_{i=1}^{\vartheta}\xi^i, \tag{7}$$

где $\xi^1, \xi^2, \ldots, \xi^\vartheta, \ldots$ — последовательность независимых случайных величин, имеющих показательное распределение с параметром λ^*: $\xi^i = -\ln\beta_i/\lambda^*$; $\alpha_1, \alpha_2, \ldots, \alpha_\vartheta, \ldots, \beta_1, \beta_2, \ldots, \beta_\vartheta, \ldots$ — последовательности независимых случайных величин, имеющих равномерное распределение на интервале $(0, 1)$; $\lambda(t)$ — значение функции $\lambda(t, x)$ на траекториях решения уравнения (1), т.е. $\lambda(t) = \lambda(t, X(t))$; $t_* \in \{\tau_j\}$ — начальный момент времени $t = 0$ или момент времени последнего разрыва траектории. В момент разрыва траектории моделируется величина скачка Δ согласно плотности вероятности $\psi(t_* + \tau, \delta)$.

Применяя более экономичный модифицированный метод максимального сечения [1–3, 9], в соотношении (7) полагаем

$$N = \min\biggl\{\vartheta\colon\ 1 - \alpha > \prod_{i=1}^{\vartheta}\biggl(1 - \frac{\lambda(t_* + \theta_i)}{\lambda^*}\biggr)\biggr\},$$

где α — равномерно распределенная на интервале $(0,1)$ случайная величина.

Для приближенного решения задачи оптимального оценивания выберем число M моделируемых траекторий пары $(X(t),\omega(t))$ и обозначим через (X_k^i,ω_k^i) дискретную аппроксимацию выборочной траектории $(X(t),\omega(t))$ с номером $i=1,2,\ldots,M$. Для нахождения реализаций (X_k^i,ω_k^i) используются соотношения (5) и (6) с учетом дополнительных узлов сетки, соответствующих моментам разрывов траекторий случайного процесса $X(t)$, в которых реализация вектора состояния изменяется согласно (3). Это означает, что сетка по времени должна быть суперпозицией равномерной сетки с шагом h, а также моментов разрывов траекторий случайного процесса $X(t)$, удовлетворяющего уравнению (1).

Тогда приближенное решение задачи оптимального оценивания имеет вид нормированного взвешенного среднего:

$$\hat{X}_\kappa = \left(\sum_{i=1}^{M}\omega_k^i\right)^{-1}\sum_{i=1}^{M}\omega_k^i X_\kappa^i,$$

где индекс k соответствует текущему моменту времени $t=t_k$, а индекс κ — моменту времени $\theta=t_\kappa$, для которого вычисляется оценка вектора состояния. Аналогично могут быть найдены моментные характеристики более высокого порядка, а также оценка плотности вероятности или функции распределения вектора состояния (с учетом весовых коэффициентов ω_k^i).

Таким образом, при $\kappa=k$ получается приближенное решение задачи фильтрации, при $\kappa<k$ — задачи интерполяции и при $\kappa>k$ — задачи экстраполяции.

Отметим, что в задаче интерполяции в отличие от задачи фильтрации необходимо хранить траектории случайного процесса $X(t)$ в памяти компьютера. В задаче фильтрации достаточно хранить только реализации вектора состояния для текущего времени, они определяют начальные данные в задаче экстраполяции [5].

Список литературы

[1] Аверина Т. А. Модифицированный алгоритм статистического моделирования систем со случайной структурой с распределенными переходами // Сибирский журнал вычислительной математики. 2013. Т. 16. № 2. С. 97–105.

[2] Аверина Т. А. Построение алгоритмов статистического моделирования систем со случайной структурой. Новосибирск: РИЦ НГУ, 2015.

[3] Аверина Т. А. Использование модификаций метода максимального сечения для моделирования систем со случайной структурой с распределенными переходами // Сибирский журнал вычислительной математики. 2016. Т. 19. № 3. С. 235–247.

[4] Аверина Т. А., Рыбаков К. А. Два метода анализа стохастических систем с пуассоновской составляющей // Дифференциальные уравнения и процессы управления. 2013. № 3. С. 85–116.

[5] Аверина Т. А., Рыбаков К. А. Приближенное решение задачи прогнозирования для стохастических систем диффузионно-скачкообразного типа // Сибирский журнал вычислительной математики. 2017. Т. 20. № 1. С. 1–11.

[6] Бухалев В. А. Оптимальное сглаживание в системах со случайной скачкообразной структурой. М.: Физматлит, 2013.

[7] Бухалев В. А., Скрынников А. А., Болдинов В. А. Алгоритмическая помехозащита беспилотных летательных аппаратов. М.: Физматлит, 2018.

[8] Королюк В. С., Портенко Н. И., Скороход А. В., Турбин А. Ф. Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.

[9] Михайлов Г. А., Аверина Т. А. Алгоритм «максимального сечения» в методе Монте-Карло // Доклады АН. 2009. Т. 428. № 2. С. 163–165.

[10] Михайлов Г. А., Войтишек А. В. Численное статистическое моделирование. Методы Монте-Карло. М.: Издательский центр «Академия», 2006.

[11] Пантелеев А. В., Руденко Е. А., Бортаковский А. С. Нелинейные системы управления: описание, анализ и синтез. М.: Вузовская книга, 2008.

[12] Руденко Е. А. Непрерывная конечномерная локально-оптимальная фильтрация диффузионно-скачкообразных сигналов // Известия РАН. Теория и системы управления. 2018. № 4. С. 14–43.

[13] Рыбаков К. А. Статистические алгоритмы оптимальной фильтрации сигналов в нелинейных диффузионно-скачкообразных стохастических системах // Вестник УГАТУ. 2016. Т. 20. № 4 (74). С. 107–113.

[14] Рыбаков К. А. О вычислении весовых коэффициентов в непрерывном фильтре частиц // Научный вестник МГТУ ГА. 2018. Т. 21. № 2. С. 32–39.

[15] Рыбаков К. А. Непрерывные фильтры частиц для стохастических систем диффузионно-скачкообразного типа // Актуальные проблемы прикладной математики, информатики и механики. Международная научно-техническая конференция, Воронеж, 17–19 декабря 2018 г.: Сб. тр. конф. Воронеж: Изд-во «Научно-исследовательские публикации», 2018. С. 1613–1619.

[16] Сосулин Ю. Г. Теоретические основы радиолокации и радионавигации. М.: Радио и связь, 1992.

[17] Степанов О. А. Основы теории оценивания с приложениями к задачам обработки навигационной информации. Т. 1. Введение в теорию оценивания. СПб.: АО «Концерн «ЦНИИ «Электроприбор», 2017.

[18] Степанов О. А. Основы теории оценивания с приложениями к задачам обработки навигационной информации. Т. 2. Введение в теорию фильтрации. СПб.: АО «Концерн «ЦНИИ «Электроприбор», 2017.

[19] Тихонов В. И., Кульман Н. К. Нелинейная фильтрация и квазикогерентный прием сигналов. М.: Советское радио, 1975.

[20] Тихонов В. И., Харисов В. Н. Статистический анализ и синтез радиотехнических устройств и систем. М.: Радио и связь, 1991.

[21] Bain A., Crisan D. Fundamentals of Stochastic Filtering. Springer, 2009.

[22] Hanson F. B. Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation. SIAM, 2007.

[23] Platen E., Bruti-Liberati N. Numerical Solution of Stochastic Differential Equations with Jumps in Finance. Springer, 2010.

[24] Situ R. Theory of Stochastic Differential Equations with Jumps and Applications. Springer, 2005.

Аверина Татьяна Александровна — к.ф.-м.н., ст. науч. сотр. Института вычислительной математики и математической геофизики СО РАН;
доцент Новосибирского государственного университета;
e-mail: ata@osmf.sscc.ru;
Рыбаков Константин Александрович — к.ф.-м.н., доцент Московского авиационного института (национального исследовательского университета)
e-mail: rkoffice@mail.ru.
Дата поступления — 22 апреля 2019 г.

МОДЕЛИРОВАНИЕ ТУРБУЛЕНТНОГО ТЕЧЕНИЯ НА ОСНОВЕ ГЕНЕРАЦИИ СТОХАСТИЧЕСКОГО ПОЛЯ СКОРОСТЕЙ

А. В. Александров[1], Л. В. Дородницын[2], А. П. Дубень[1]

[1] *Институт прикладной математики им. М.В. Келдыша РАН, 125047, Москва*

[2] *Московский государственный университет имени М.В. Ломоносова, 119991, Москва*

УДК 519.245, 532.517.4
DOI: 10.24411/9999-016A-2019-10003

Рандомизированный спектральный метод, ранее предложенный для генерации однородных изотропных турбулентных полей скорости, практически реализован в трехмерном случае. Аналитически и численно демонстрируется совпадение статистических и спектральных свойств турбулентных полей, искусственно сгенерированных на основе данного метода, со свойствами их физических прототипов. Анализируется влияние различных способов учета зависимости от времени на вид автокорреляционной функции. Валидационный LES расчет, использующий полученные турбулентные поля в качестве начальных условий, демонстрирует хорошее совпадение результатов с экспериментальными данными.

Ключевые слова: синтетическая турбулентность, рандомизированный спектральный метод, однородная изотропная турбулентность.

Введение

При численном моделировании турбулентных течений с высокими числами Рейнольдса применение DNS и LES расчетов существенно ограничивается их высокой вычислительной стоимостью. В настоящее время всё чаще и чаще используются искусственно сгенерированные (синтетические) турбулентные поля, обладающие заданными статистическими и физическими характеристиками. Данное направление начато в 1970 году Крайчнаном [1]. Сегодня для расчета течений возле сложных конфигураций целесообразно комбинировать осредненную модель во всей области с детальным LES описанием в малой подобласти потока, где необходима высокая точность, свойственная вихреразрешающим подходам. Это стимулирует использование синтетических турбулентных полей в качестве граничных условий для области LES.

Простейшим способом задания случайного турбулентного поля в качестве входных граничных условий для области LES является использование «белого шума» [2]. Однако на практике это приводит к быстрому затуханию сгенерированных пульсаций вниз по потоку.

Принято считать [3], что синтетическое турбулентное поле на входной границе LES должно обладать теми же свойствами, которые известны о реальном физическом турбулентном поле. Наряду с бездивергентностью турбулентных пульсаций скорости (как в жидкости, так и в газе при дозвуковом течении), обычно требуется также совпадение статистических характеристик: одноточечных моментов 1-го, 2-го и 3-го порядков, двухточечных пространственных корреляций и временных автокорреляций. Помимо этого, должен воспроизводиться энергетический спектр турбулентности.

Среди способов генерации синтетических турбулентных полей наиболее популярными являются спектральные или Фурье-методы [4]– [6], методы синтетических вихрей [7], а также методы, основанные на использовании специальных численных фильтров [8].

В настоящей работе рассматривается стохастический спектральный метод, построенный в [9]– [11]. Конкретно, результаты [12, 13], где данный подход применялся к двумерным задачам вычислительной аэроакустики, обобщаются авторами на трехмерный случай. Стохастический выбор параметров метода лучше

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 18-01-00726).

ISBN 978-5-901548-42-4

соответствует природе турбулентности, чем полудетерминированный, традиционно используемый в спектральных методах генерации поля скоростей.

1 Рандомизированный спектральный метод

Метод генерации искусственных турбулентных полей, основанный на суперпозиции Фурье-мод впервые был предложен в работе [1] и получил развитие в ряде работ [4]– [6]. В рамках этого метода, амплитуды косинусоидальных и/или синусоидальных мод определяются на базе энергетического спектра турбулентности, а изотропность и несжимаемость обеспечиваются выбором направления волнового вектора.

Будем рассматривать подход, основанный на гармониках со случайными волновыми числами и получивший название «рандомизированный спектральный метод» (РСМ) [9]– [11]. Технология была успешно апробирована на задачах аэроакустики в [12,13] где был реализован метод генерации синтетического двумерного гауссова изотропного турбулентного поля. Достоинством алгоритма являлось хорошее воспроизведение заданных спектральных и корреляционных характеристик турбулентности сгенерированными полями. Авторами настоящей работы было выполнено обобщение данного метода на трехмерный случай и проведено как аналитическое, так и численное исследование свойств полученного генератора.

Поле пульсационных скоростей реализуется в виде суммы гармоник:

$$\mathbf{u}'(\mathbf{x},t) = \frac{\sigma}{\sqrt{N}} \sum_{n=1}^{N} \mathbf{u}^n(\mathbf{x},t). \tag{1}$$

Каждая гармоника представляется формулой

$$\begin{aligned} \mathbf{u}^n(\mathbf{x},t) = \mathbf{Q}(\vec{\omega}^n)\Big[\vec{\xi}^n \sin\big(k_n\,(\vec{\omega}^n\cdot\mathbf{x}) - k_n\,(\vec{\omega}^n\cdot\mathbf{U})\,t - c_n t/\tau_0\big) + \\ + \vec{\eta}^n \cos\big(k_n\,(\vec{\omega}^n\cdot\mathbf{x}) - k_n\,(\vec{\omega}^n\cdot\mathbf{U})\,t - c_n t/\tau_0\big)\Big]. \end{aligned} \tag{2}$$

Здесь k_n, c_n, $\vec{\omega}^n$, $\vec{\xi}^n$, $\vec{\eta}^n$ — случайные числа и векторы. Матрица $\mathbf{Q}(\vec{\omega})$ размерности 3×3 с элементами $q_{jl} = \delta_{jl} - \omega_j\omega_l$, где вектор $\vec{\omega} = (\omega_1,\omega_2,\omega_3)$, $|\vec{\omega}| = 1$, равномерно распределен по сфере, обеспечивает бездивергентность поля скорости $\nabla\cdot\mathbf{u}^n = 0$. Два набора независимых случайных трехмерных векторов $\vec{\xi}^n$, $\vec{\eta}^n$ имеют стандартное гауссовское распределение $N(0,1)$.

Волновые числа $k = k_n \in [k_{\min}, k_{\max}]$ составляют набор независимых случайных величин, распределенных с плотностью вероятности, определяемой энергетическим спектром: $\rho(k) = E(k)/\sigma^2$,

Зависимость от времени в (2) задается в форме переноса гармоники со скоростью фонового внешнего потока $\mathbf{U}$ и хаотического блуждания, где τ_0 — временной масштаб турбулентности, c_n — случайная величина с гауссовым распределением $N(0,1)$. Похожий способ применялся в [14].

В качестве $E(k)$ можно выбрать, например, модифицированный спектр Кармана–Пао из [6], задаваемый формулой

$$E(k) = A\,\frac{(k/k_e)^4}{\Big(1+(k/k_e)^2\Big)^{17/6}}\,f_\eta f_{\rm cut}\,,$$

где $f_\eta(k)$ и $f_{\rm cut}(k)$ — эмпирические функции. Свободные параметры: нормировочная константа A и характерное волновое число k_e энергонесущих гармоник — выбираются в соответствии с данными осредненного турбулентного поля (RANS):

$$\int_{k_{\min}}^{k_{\max}} E(k)\,dk = \sigma^2, \quad 2\nu\int_{k_{\min}}^{k_{\max}} k^2\,E(k)\,dk = \varepsilon.$$

2 Анализ численных реализаций трехмерного турбулентного поля скорости

Свойства описанного РСМ-генератора исследовались на примере искусственно сгенерированного однородного изотропного турбулентного поля в кубе с безразмерной стороной 2π. На рис. 1 приведено изображение сгенерированного поля скорости.

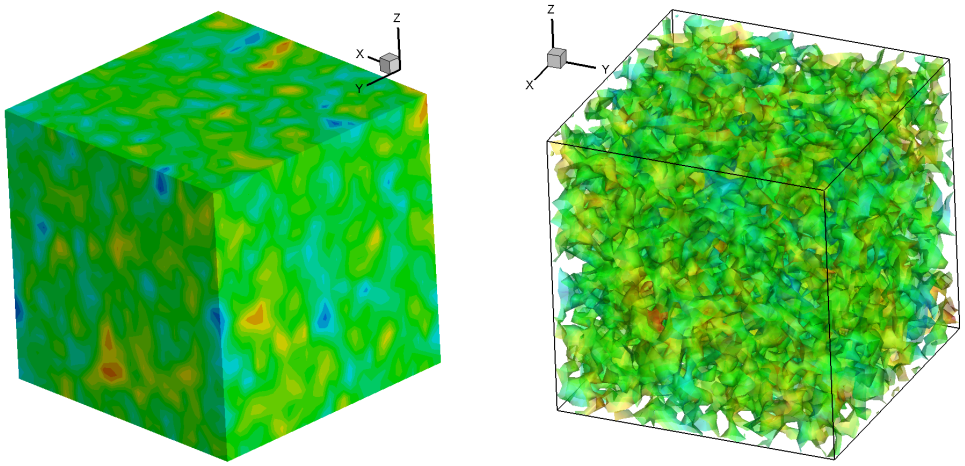


Рис. 1: Сгенерированное поле скорости: x-компонента (слева) и Q-критерий (справа)

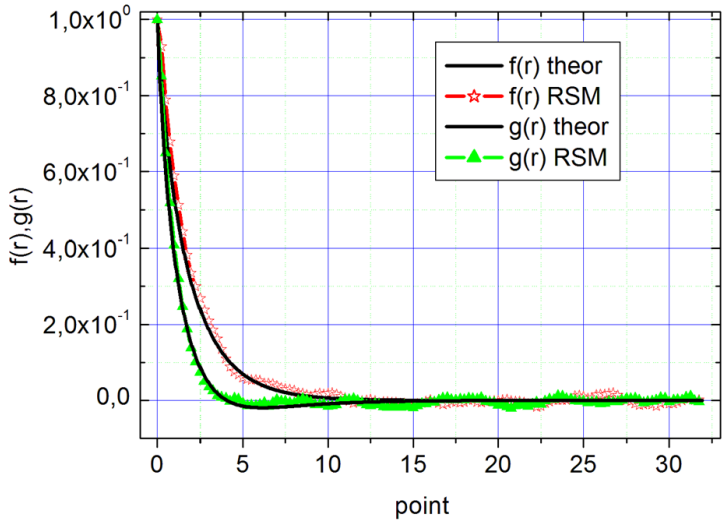


Рис. 2: Продольная и поперечная корреляционные функции

Оценивались одноточечные моменты скоростей 1-го, 2-го и 3-го порядков, двухточечные пространственные моменты второго порядка. Проводилось сравнение энергетических спектров численно сгенерированных полей с исходной функцией $E(k)$. Аналитически и численно исследовалось влияние различных способов учета зависимости от времени на автокорреляционную функцию.

Выделим важнейшие результаты.

2.1 Двухточечные пространственные корреляции

Ключевую роль в описании свойств изотропной турбулентности играют продольная и поперечная корреляционные функции Кармана–Ховарта

$$f(r) = \frac{\langle u_x(r,0,0)\, u_x(0,0,0)\rangle}{\langle u_x(0,0,0)^2\rangle}, \quad g(r) = \frac{\langle u_y(r,0,0)\, u_y(0,0,0)\rangle}{\langle u_y(0,0,0)^2\rangle}.$$

(Угловые скобки обозначают математическое ожидание.) На рис. 2 приведены зависимости двух корреляционных функций от расстояния для 5000 реализаций турбулентного поля и графики теоретически ожидаемых значений этих величин.

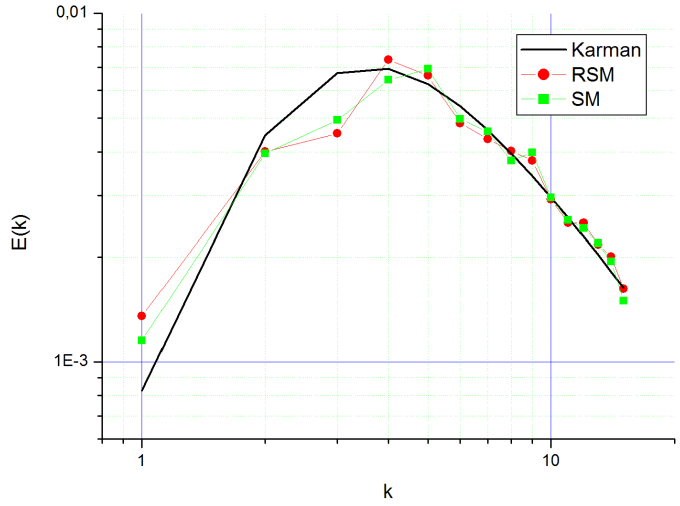


Рис. 3: Энергетический спектр полей, полученных на основе РСМ и метода [5]

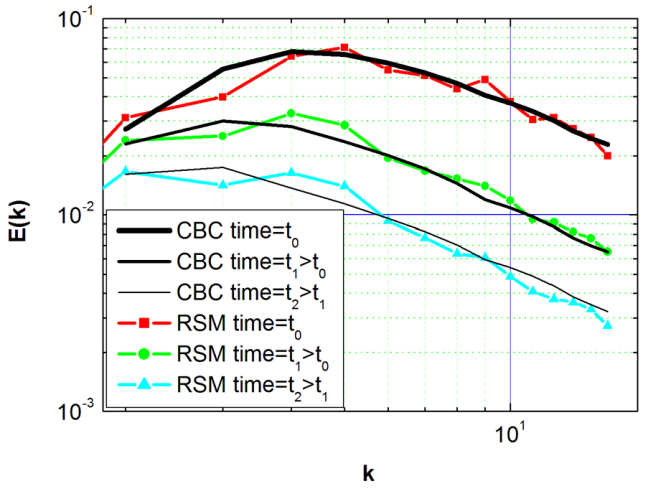


Рис. 4: Расчетные и экспериментальные энергетические спектры в задаче о затухании однородной изотропной турбулентности

2.2 Энергетический спектр

На рис. 3 приведены энергетические спектры турбулентных полей скорости, сгенерированных на сетке 32^3 с использованием $N = 10000$ гармоник на основе спектрального метода из [5] и рандомизированного спектрального метода из настоящей работы. В инерционном интервале оба численных спектра очень близки к исходному спектру Кармана – Пао.

3 Валидационный расчет

Для валидации методики РСМ был выполнен расчет классической задачи о затухании однородной изотропной турбулентности, экспериментально исследованной в работе [15]. С помощью генератора РСМ задавалось начальное турбулентное поле, отвечающее экспериментальному энергетическому спектру на некий момент времени $E(k, t_0)$. Далее проводился расчет детерминированной эволюции данного распределения методом LES с подсеточной моделью Смагоринского [16].

Использовался вычислительный алгоритм, реализованный в программном комплексе NOISEtte [17] и неоднократно валидированный [18, 19] для проведения вихреразрешающего моделирования. Расчет выполнялся в кубе с безразмерной стороной 2π, на сетке 32^3 узлов, с периодическими граничными условиями. Характерное число Рейнольдса задачи составляло $\mathrm{Re} = 1620$.

Энергетические спектры строились на два последующих момента времени, соответствующих эксперименту [15]. Графики, приведенные на рис. 4, показывают хорошее совпадение с экспериментом.

Заключение

Проведенные аналитические и численные исследования свойств трехмерных изотропных турбулентных полей скорости, построенных на основе предложенного генератора, демонстрируют хорошее совпадение статистических характеристик как с теоретическими, так и с экспериментальными данными. Поэтому данный генератор, полученный обобщением на трехмерный случай PCM метода из [12, 13], может стать базой для дальнейшего использования полей, полученных на его основе, при задании нестационарных входных граничных условий для LES расчетов.

Дальнейшее развитие предложенного способа генерации синтетических турбулентных полей скорости может заключаться в его адаптации к анизотропному случаю, что было бы актуальным для большинства инженерных задач. В ряде работ синтетическое анизотропное поле приводится к изотропному виду и обратно путем масштабирования тензора рейнольдсовых напряжений [14] либо его разложения Холецкого [6]. Обе методики полностью совместимы с настоящей технологией PCM.

Список литературы

[1] Kraichnan R. Diffusion by a random velocity field // Phys. of Fluids. 1970. V. 13, iss. 1. P. 22–31.

[2] Batten P., Goldberg U., Chakravarthy S. Interfacing statistical turbulence closures with Large-Eddy Simulation // AIAA J. 2004. V. 42, iss. 3. P. 485–492.

[3] Dhamankar N. S., Blaisdell G. A., Lyrintzis A. S. Overview of turbulent inflow boundary conditions for Large-Eddy Simulations // AIAA J. 2018. V. 56, iss. 4. P. 1317–1334.

[4] Bechara W., Bailly C., Lafon P., Candel S. M. Stochastic approach to noise modeling for free turbulent flows // AIAA J. 1994. V. 32, iss. 3. P. 455–463.

[5] Saad T., Cline D., Stoll R., Sutherland J. C. Scalable tools for generating synthetic isotropic turbulence with arbitrary spectra // AIAA J. 2016. V. 55, iss. 18. P. 327–331.

[6] Shur M. L., Spalart P. R., Strelets M. K., Travin A. K. Synthetic turbulence generators for RANS-LES interfaces in zonal simulations of aerodynamic and aeroacoustic problems // Flow Turbulence Combust. 2014. V. 93, iss. 1. P. 63–92.

[7] Адамьян Д. Ю., Травин А. К. Усовершенствованный метод генерации синтетических вихрей для задания нестационарных входных граничных условий при расчете турбулентных течений // ТВТ. 2011. Т. 49, вып. 5. С. 728–736.

[8] Klein M., Sadiki A., Janicka J. A digital filter based generation of inflow data for spatially developing direct numerical or large eddy simulations // J. Comput. Phys. 2003. V. 186, iss. 2. P. 652–665.

[9] Sabelfeld K. K. Monte Carlo Methods in boundary value problems. Springer, Heidelberg–Berlin–New York, 1991.

[10] Kurbanmuradov O., Sabelfeld K. Stochastic spectral and Fourier-wavelet methods for vector Gaussian random fields // Monte Carlo Methods Appl. 2006. V. 12, iss. 5–6. P. 395–445.

[11] Kurbanmuradov O., Sabelfeld K., Kramer P. R. Randomized spectral and Fourier-wavelet methods for multidimensional Gaussian random vector fields // J. Comput. Phys. 2013. V. 245. P. 218–234.

[12] Боровская И. А. Моделирование однородных случайных полей по заданному спектру в задачах аэроакустики // Матем. моделирование. 2007. Т. 19, вып. 7. С. 67–76.

[13] Боровская И. А., Козубская Т. К., Курбанмурадов О., Сабельфельд К. К. О моделировании однородных случайных полей и сигналов и их использовании в задачах аэроакустики // Матем. моделирование. 2007. Т. 19, вып. 10. С. 76–88.

[14] Smirnov A., Shi S., Celik I. Random flow generation technique for large eddy simulations and particle-dynamics modeling // J. Fluids Eng. 2001. V. 123, iss. 2. P. 359–371.

[15] Comte-Bellot G., Corrsin S. Simple Eulerian time correlation of full- and narrow-band velocity signals in grid-generated 'isotropic' turbulence // J. Fluid Mech. 1971. V. 48, part 2. P. 273–337.

[16] Smagorinsky J. General circulation experiments with the primitive equations. I. The basic experiment // Month. Wea. Rev. 1963. V. 91. P. 99–164.

[17] Абалакин И. В., Бахвалов П. А., Горобец А. В., Дубень А. П., Козубская Т. К. Параллельный программный комплекс NOISETTE для крупномасштабных расчетов задач аэродинамики и аэроакустики // Вычислительные методы и программирование. 2012. Т. 13. С. 110–125.

[18] Дубень А. П. Вычислительные технологии для моделирования сложных пристеночных турбулентных течений на неструктурированных сетках // Матем. моделирование. 2013. Т. 25, вып. 9. С. 4–16.

[19] Дубень А. П., Козубская Т. К., Потапов Д.В. Моделирование нестационарных изотропных турбулентных течений на неструктурированных сетках с использованием реберно-ориентированных алгоритмов // Матем. моделирование. 2017. Т. 29, вып. 5. С. 27–45.

Александров Анатолий Витальевич — к.ф.-м.н., ст. науч.сотр. Института прикладной математики им. М.В. Келдыша РАН;
e-mail: anatoly.v.alexandrov@gmail.com;
Дородницын Людвиг Вацлавович — к.ф.-м.н., ст. науч.сотр. Московского государственного университета имени М.В. Ломоносова;
e-mail: dorodn@cs.msu.su;
Дубень Алексей Петрович — к.ф.-м.н., науч.сотр. Института прикладной математики им. М.В. Келдыша РАН;
e-mail: alexey.duben@gmail.com.
Дата поступления — 30 мая 2019 г.

ДЕКОМПОЗИЦИЯ И РАСПАРАЛЛЕЛИВАНИЕ ВЫЧИСЛИТЕЛЬНОЁМКИХ ФУНКЦИЙ ОБРАБОТКИ ДАННЫХ МИКРОСЕЙСМИЧЕСКОГО МОНИТОРИНГА

К. С. Алсынбаев[1]

[1] *НИИ прикладной информатики и математической геофизики БФУ имени И.Канта, 236001, Калининград*

УДК 519.683:550.8
DOI: 10.24411/9999-016A-2019-10004

Программное обеспечение обработки данных микросейсмического мониторинга характеризуются большим объёмом входной информации, высокой вычислительной ёмкостью алгоритмов. Программы могут включать функции распознавания и геометрического моделирования объектов в нерегулярном облаке точек, использовать алгоритмы с быстро растущим временем счета при увеличении объёма облака. Предлагаемые подходы для ускорения выполнения функций основываются на декомпозиции и распараллеливании. Одни из них базируются на классических методах разбиения векторных данных и декомпозиции алгоритмов, другие используют переход к воксельной модели данных.

Ключевые слова: микросейсмический мониторинг, кластеризация, тетраэдризация, воксель, параллельные вычисления, распознавание.

Введение

Программное обеспечение обработки данных микросейсмического мониторинга можно условно разделить на две части. Первая представляет собой обработку исходных оцифрованных сигналов от поверхностных датчиков. Обычно посредством решения обратной кинематической задачи вычисляются локации сейсмических событий и преобразуются в представление набора (облака) трёх- или четырёх-мерных параметризованных точек, соответствующих событиям (далее будем называть их точками-событиями). Среди параметров — время проявления сигнала сейсмической эмиссии, энергия, скорость распространения сейсмических волн в среде и др. Во второй части решаются задачи выделения полезного сигнала посредством пространственно-временной фильтрации, выявления характерных особенностей каждой точки-события и их совокупностей для интерпретации происходящих в околоскважинной области процессов, создаётся трёхмерная динамическая визуализация полученных результатов. Одна из главных целей — выявление или подчёркивание особенностей процессов развития трещиноватости. В коллективе автора статьи созданы средства визуального подчёркивания тех или иных особенностей точек-событий для экспертной оценки картины данных пользователем [1]. Разрабатываются экспериментальные средства выявления уплотнений и образований в облаке точек-событий с использованием средств геометрического моделирования.

В силу большого объёма обрабатываемых данных и сложности используемых алгоритмов программы, как правило, имеют значительный объём вычислений. Практическое использование программного обеспечения остро ставит задачу сокращения времени обработки. Собственно, возможны два пути: оптимизация алгоритмов и их распараллеливание.

Есть алгоритмы, объём вычислений которых растёт приблизительно линейно от объёма обрабатываемых данных, и вычисления локализованы для анализируемых точек, например, в задаче вычисления плотности точек (чем больше в окрестности точки соседей, тем выше плотность). Очевидно, такие алгоритмы легко распараллеливаются. Задача сильно усложняется, когда в алгоритме рассматриваемая точка может быть

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 16-29-15090).

ISBN 978-5-901548-42-4

связана с другой в любой области пространства, или объём вычислений быстро растёт от количества точек. К таковым относятся алгоритмы кластеризации и тетраэдризации, способы ускорения которых будут рассмотрены в статье.

Распараллеливание, даже если оно возможно, не решает кардинально задачи ускорения алгоритмов с временной сложностью больше $O(N^2 \log N)$. Одним из подходов может быть изменение алгоритма таким образом, чтобы разделить обрабатываемые данные на фрагменты с ограниченным количеством элементов в каждом и организовать их раздельную обработку, чтобы время счета алгоритма для любого фрагмента не выросло до неприемлемых значений. При таком способе разделения обработки данных необходимо решить задачу последующего объединения локальных результатов, что может представлять непростую проблему.

Характеризуемый выше подход изменения алгоритма уместно назвать декомпозицией алгоритма. Декомпозиция не только уменьшает временную сложность алгоритма, но и позволяет распараллелить задачу.

1 Модификация алгоритма кластеризации

1.1 Построение дерева для точек-событий

Для максимальной эффективности разделения данных нужно обеспечить приблизительно равные объёмы обрабатываемых фрагментов, что достигается применением механизма построения деревьев, когда исходная область иерархически разбивается на смежные фрагменты, пока фрагмент не будет содержать достаточно малое количество точек. Согласно общепринятой терминологии, при манипуляциях с деревом образуются узлы и вершины. Вершины указывают на конечную непустую область пространства, узлы — на область, подвергнутую разбиению или на конечную пустую область.

На практике могут использоваться бинарные деревья (например, BSP-деревья [2]), или, в зависимости от размерности пространства, октодеревья и 16-деревья. В наших реализациях выбран второй подход, поскольку теория бинарных деревьев развита для оптимизации последующего поиска, а в нашем случае поиск отсутствует, основные временные затраты при построении дерева и обратной операции, важнее оптимизация для сложных операций, ориентированных по координатам.

Используется рекурсивный способ построения дочерних узлов. Теоретически временная сложность такого алгоритма в вырожденных случаях, когда большинство точек сконцентрированы в небольшой части рассматриваемого пространства, может достигать $O(N^2)$. Но облако точек-событий микросейсмических данных содержит в целом равномерно распределённые точки, «сгущения» представляют умеренно плотные образования. Построение дерева вглубь ограничиваются минимально допустимым размером ребра области, отвечающей вершине, поэтому количество V вершин дерева и глубина дерева H ограничены. Поскольку известна оценка временной сложности рекурсивного построения октодеревьев, равная $O(NH)$, то практическая временная сложность может быть оценена между $O(N)$ и $O(N \log N)$.

1.2 Декомпозиция алгоритма кластеризации

Кластеризация точек-событий в трёх- или четырёх-мерном пространстве необходима для некоторых задач пространственно-временной фильтрации и для выявления «сгущений» событий перед геометрическим моделированием выявляемых образований. Задача кластеризации ставится при изначально неизвестном количестве кластеров, критерии принадлежности — допустимое максимальное расстояние D между соседними точками кластера, расстояние может быть эвклидовым, или же неэвклидовым с построением неизотропной метрики на основе тензоров сейсмического момента [1]). В качестве основы выбран алгоритм кластеризация точек на регулярной сети [3], легко модифицирующийся для случаев трех- и четырех- измерений и для различных видов используемой метрики. Алгоритм позволяет выделять связные компоненты типа сгущений или лент, которые наиболее пригодны для распознавания разломов или зон трещиноватости. Однопроцессорная реализация без фрагментации данных имеет временную сложность от $O(N^2 \log N)$ до $O(N^3)$ [4], где N — количество точек-событий в облаке.

Модификация алгоритма заключается в следующих этапах:

1. Проводится разбиение исходного облака точек-событий рекурсивным механизмом построения окто- или 16- дерева, пока количество точек в области пространства, отвечающей порождённому узлу, не будет превышать заданное значение Z или очередной шаг разбиения не приведёт к заданному предельно малому размеру области.

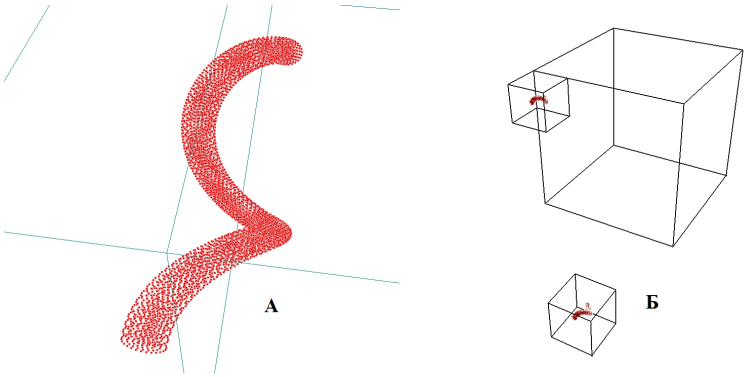


Рис. 1: А. Укрупнённый фрагмент тестового множества, попадающий на несколько соседних листьев. Б. Пример порожденных вершин октодерева для теста. Имеются 2 ограничивающие точки (на нижней плоскости ближе к наблюдателю и вдали от наблюдателя) для формирования требуемых размеров октодерева

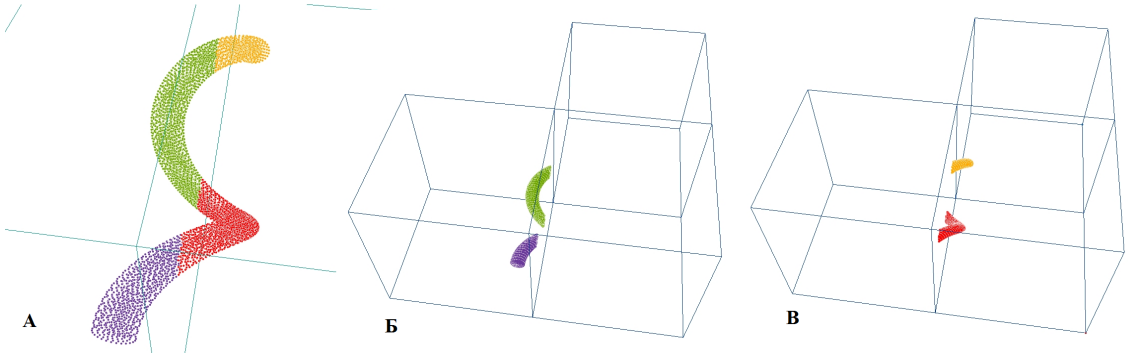


Рис. 2: Выполненные раздельно для соседствующих вершин кластеризации фрагмента облака

2. Исходным однопроцессорным алгоритмом осуществляется кластеризация для каждой вершины.

3. Проводится иерархическая склейка элементов разбиения вместе с полученными в них кластерами.

Для лучшей наглядности иллюстрирование будем производить для трёх-мерного случая. В качестве иллюстраций приводятся операции с тестовыми данными, которые представляют собой подмножества точек, закрученные в спирали (рис.1 А).

В октодереве первоначальная область с облаком точек выбирается кубической формы, чтобы избежать вырожденных случаев кластеризации «почти плоских» областей. Таким образом при начальном разбиении области, соответствующие узлам и вершинам, имеют кубическую форму. На рис. 1 Б приведён пример фрагмента разбиения с областями для вершин различного размера.

На рис. 2 приводится пример раздельной кластеризации фрагмента облака, образующего один кластер в исходном облаке, попадающего в области двух соседствующих вершин. При кластеризации каждому кластеру присваивается локальный номер, который сохраняется в структуре данных точки-события. Точки одного кластера объединяются в однонаправленный список.

Склейка фрагментов является наиболее сложной частью алгоритма. Производится последовательное укрупнение элементов дерева, начиная с вершин первоначального дерева. На каждом шаге производится склейка двух однотипных по размеру элементов, при этом дерево модифицируется и образуется новая вершина.

Есть 2 типа шагов: 1) поиск в дереве всех вершин, соседствующих с пустыми областями, однотипными по форме и размеру и укрупнение дерева путём операций слияния всех пар соседних областей. На рис. 3 приводится пример процесса такого укрупнения некоторой области дерева до приведения к соседству только вершин. 2) поиск и операции слияния всех пар однотипных по размеру вершин с одновременной склейкой соседствующих по границам кластеров. Результатом слияния пар элементов дерева является прямоугольный параллелепипед, необязательно куб. Шаги типов 1) и 2) попеременно чередуются, пока не останется един-

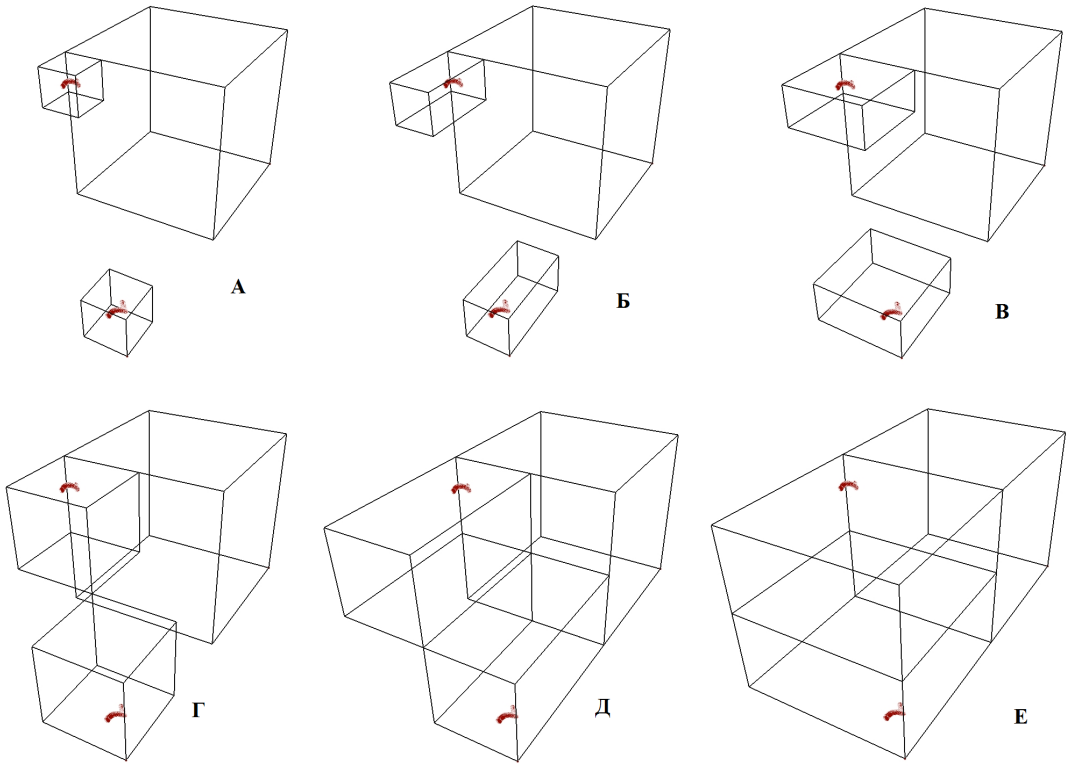


Рис. 3: Укрупнение элементов дерева в случае соседства вершин и пустых узлов

ственная вершина дерева, которая, собственно, и содержит все кластеры, включая склеенные. Внутренние циклы шагов как типа 1), так и типа 2) могут распараллеливаться.

Будем обозначать множество точек большой латинской буквой, например P, а точку в этом множестве как Pi, кластер, к которому принадлежит точка и номер этого кластера, как $k(Pi)$. При построении дерева возникают множества точек, содержащиеся в прямоугольной области пространства, соответствующее вершине при построении и модификации дерева. Обозначим как P множество точек, соответствующее первой из вершин, участвующих в склейке, и как R — множество точек, соответствующее второй вершине.

Алгоритм склейки кластеров соседних вершин шага типа 2) осуществляется следующим образом.

- Обеспечивается уникальность номеров кластеров в объединяемом множестве кластеров путём увеличения номеров кластеров в R на значение количества кластеров в P.
- Выбирается в отдельный массив G вершины из R, которые отстоят от границы с областью P не более чем на D (рис. 4 А).
- Области R и P сливаются в одну область S с объединением списочных структур.
- Для каждой точки Pi области P производится просмотр отобранных граничных точек из G. Если встретится точка Gj такая, что расстояние от Pi до Gj меньше D, список кластера $k(Gj)$ подсоединяется к списку кластера $k(Pi)$.

Легко видеть, что после просмотра всех точек из P все кластеры S, расстояние между какими либо точками которых меньше D, будут объединены (рис. 4 Б, В). На рис. 4 А и Б приведены иллюстрации результатов склейки для теста большого размера, на рис. 5 В — для реального облака.

Сложность по времени алгоритма необходимо оценивать по-отдельности для трёх этапов.

Выше мы оценили сложность построения дерева для облака точек-событий микросейсмических данных между $O(N)$ и $O(N \log N)$.

Кластеризация точек одной вершины немодифицированным алгоритмом оценивается от $O(M^2 \log M)$ до $O(M^3)$ [4], где M — число точек в области одной вершины, а время последовательной кластеризации

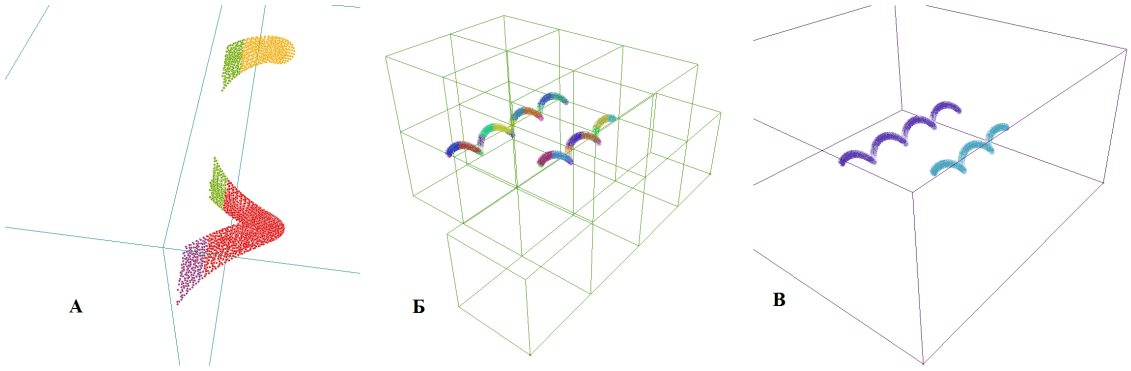


Рис. 4: А. Склейка кластеров соседствующих вершин: зелёным цветом окрашено множество приграничных точек, которое просматривается для точек области левой вершины. Б. Кластеры, прошедшие кластеризацию раздельно для вершин (цветом выделены результаты локальной кластеризации). В. Кластеры исходного облака, полученные в результате нескольких шагов иерархической склейки

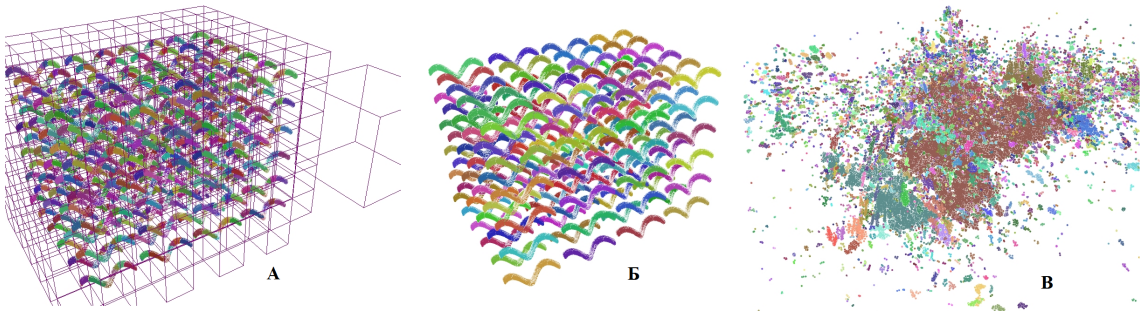


Рис. 5: А. Большой тест (облако в 770 000 точек, 426 вершин) после фрагментарной кластеризации. Б. Облако после иерархической склейки. В. Реальное облако (452 тысячи точек), кластеризованное модифицированным алгоритмом.

точек всех вершин примерно линейно зависит от числа вершин. Поскольку M ограничена константой Z, то сложность этапа кластеризации можно оценить в $O(N)$, если M всегда меньше Z. Однако это условие достигается не всегда, так как оно может конфликтовать с условием минимального размера ребра порождаемых областей, соответствующих вершинам. Практическая оценка показывает временную сложность этапа кластеризации точек всех вершин примерно в $O(N \log N)$. Алгоритм этапа кластеризации организуется в один цикл просмотра всех вершин дерева и легко распараллеливается.

Сложность по времени алгоритма этапа иерархической склейки точно оценить непросто. При операции склейки имеет место цикл по всем точкам множества P, с перебором для каждой точки множества G и проверкой условия близости. Время может быть $O(1)$ в случае пустого множества G и $O(Mg)$ в случае непустого G, где g — количество точек в G. Очевидно, в среднем $g << M$. Есть ещё внешние циклы с перебором вершин, что для прохода для верхнего уровня иерархии с числом вершин V даёт верхнюю оценку $O(MgV) = O(Ng)$, поскольку значение MV близко N. Примерно такую же оценку даёт цикл склейки каждого уровня иерархии дерева. Экспериментальные оценки показали, что сложность по времени не выше $O(N \log N)$. Заметим, что этап иерархической склейки распараллеливается не так эффективно, как этап кластеризации, так как по мере объединения вершин дерева объём точек в их областях увеличивается, а количество вершин уменьшается. В целом можно оценить сложность по времени всего модифицированного алгоритма от $O(N)$ до $O(N \log N)$.

Были проведены сравнительные эксперименты. Использовался компьютер с процессором Intel Xeon W3565 (создаётся до 8 потоков), 3.2 Ггц, память 12 Гбт. Время выполнения немодифицированной программы для 770 000 точек тестового множества составило 12635 сек. (3.5 часа). Модифицированная программа с распараллеливанием выполнилась за 20 сек. (выигрыш в 632 раза). При этом кластеризация выполнялась 9 секунд, склейка — 11 секунд. Модифицированная программа без распараллеливания выполнилась за 60 сек. (выигрыш в 211 раз). При этом кластеризация выполнялась 44 секунды, склейка — 16 секунд. Таким

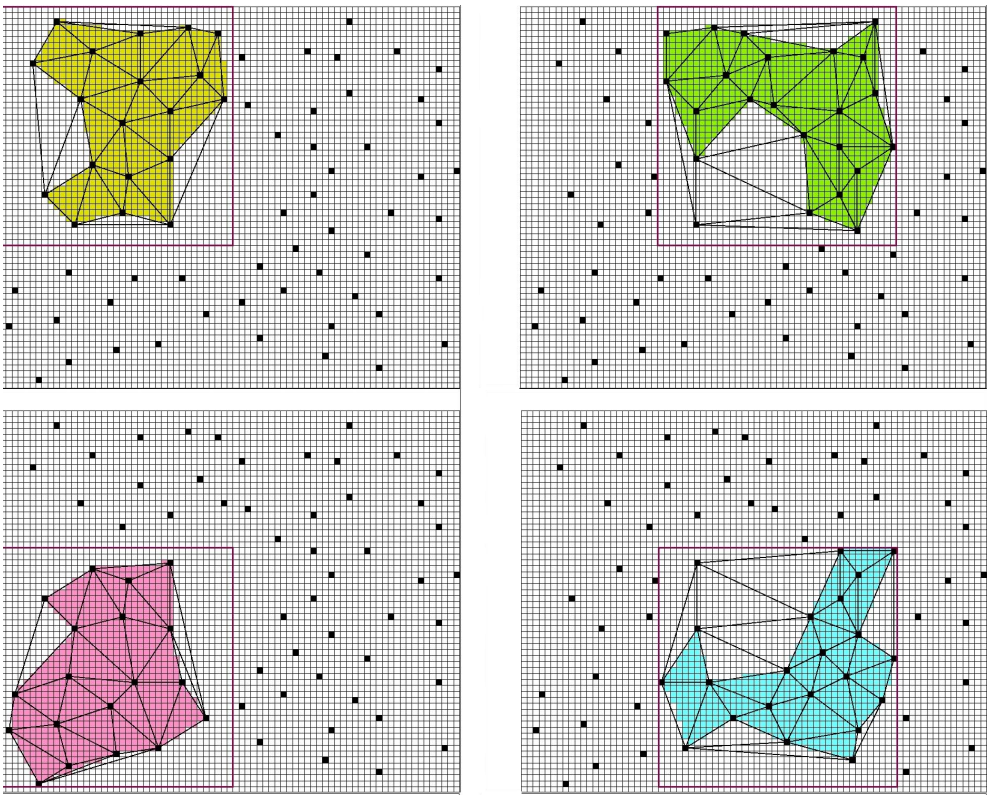


Рис. 6: Двумерная иллюстрация декомпозиции функции тетраэдризации

образом, если немодифицированный алгоритм может выполняться за приемлемое время (несколько десятков минут) для множества не более 50–100 тысяч точек, то модифицированный — для десятков миллионов точек.

2 Декомпозиция и распараллеливание воксельного алгоритма

В геоинформатике известен подход дополнения друг другом векторных и растровых представлений двухмерных объектов. В трехмерных задачах микросейсмики аналогичный приём по сочетанию векторных и воксельных представлений пространства точек-событий показывает свою результативность [1]. Покажем подход создания воксельного алгоритма на примере задачи моделирования объёмных тел.

В векторном случае построение моделей тел обычно опирается на тетраэдризацию Делоне с последующим исключением тетраэдров, у которых радиус описанной сферы превышает заданный параметр U, для моделирования невыпуклых тел с внутренними полостями. Тетраэдризация имеет те же проблемы, что и алгоритм кластеризации: трудности по декомпозиции данных для распараллеливания, квадратичный или кубический рост вычислительной сложности от количества точек. Используемый алгоритм тетраэдризации имеет сложность $O(N^2 \log N)$. Переход от векторного представления облака точек-событий к воксельному и обратно позволяет решить эти проблемы.

При вокселизации исходных данных создаётся воксельное пространство, где воксели могут иметь определенные свойства: «событие» (то есть содержат исходную точку-событие), «фон», «тело» и др. В алгоритме по воксельному пространству перемещается маска кубической формы, с ребром равным или больше $2U$. Шаг движения маски равен U. Маска движется построчно и послойно. Тетраэдризация проводится только в пределах текущего положения маски для вокселов типа «событие». Тетраэдры, удовлетворяющие выбранному выше критерию, заполняются вокселями со значением «тело». Результаты заполнения тетраэдров для очередного положения маски накладываются на перекрывающиеся области ранее вычисленных положений маски с приоритетом вокселей со значением «тело» или «событие». При этом размер маски и шаг её перемещения в результате наложения позволяют корректно определять свойства вокселей. На рис. 6 приведена

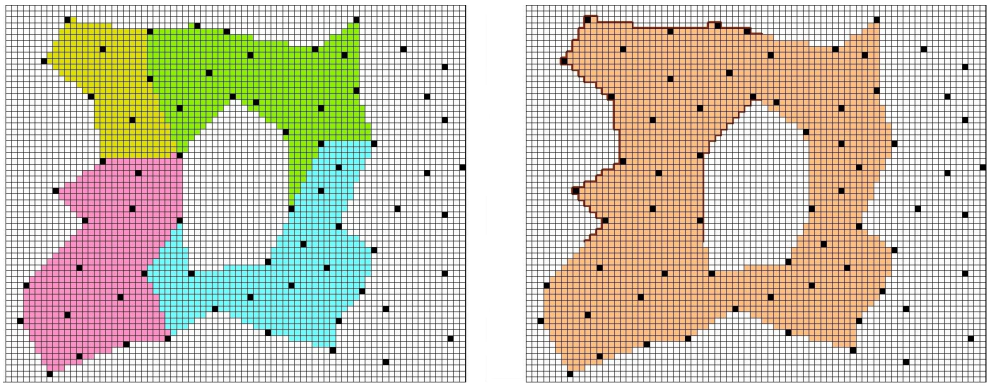


Рис. 7: Слияние с наложением тетраэдизованных зон воксельного пространства

двумерная иллюстрация четырёх шагов алгоритма: по два шага для двух соседних строк движения маски, на рис. 7 — слияние тетраэдризованных участков.

Таким образом, вокселизация позволяет разделить данные на независимое ограниченные зоны, что обеспечивает на практике близкий к линейному рост временной сложности, а также выполнять операции локальной тетраэдризации параллельно.

Заключение

В работе показан опыт кардинального ускорения некоторых функций для программ обработки данных микросейсмического мониторинга. При декомпозиции алгоритма кластеризации, очевидно, основной эффект достигается благодаря не распараллеливанию, а раздельному выполнению алгоритма на ограниченных множествах точек. Повышение быстродействия алгоритмов, в том числе с использованием подходов декомпозиции и вокселизации, позволяет ставить задачу исполняемости программного обеспечения в режиме реального времени не только на суперкомпьютерах, но и на многоядерных рабочих станциях. Это необходимо, например, для постоянно действующего микросейсмического мониторинга.

Список литературы

[1] Ерохин Г.Н., Алсынбаев К.С., Брыксин В.М., Савеленко В.В., Строков В.И., Козлов А.В., Козлов М.В. Алгоритмы интерпретации и визуализации результатов обработки данных постоянно действующего мониторинга месторождений углеводородов // Труды международной научной конференции «Марчуковские научные чтения — 2017», издательство Института вычислительной математики и математической геофизики Сибирского отделения РАН (Новосибирск), Новосибирск, 2017, с.289–295.

[2] Computational Geometry: Algorithms and Applications. — Springer Science and Business Media, 2008. — P. 259. — ISBN 978-3-540-77973-5.

[3] Кластеризация точек на регулярной сети, http://habrahabr.ru/post/138185/

[4] Алсынбаев К.С. Алгоритмы определения тел объёмных объектов в трехмерном нерегулярном облаке точек // Вестник БФУ им. И.Канта, 2015 г., № 10. С. 159–165. ISSN 2223-2095, e- ISSN 2310-3698.

Алсынбаев Камил Салихович — к.т.н., зав. лаб. НИИ прикладной информатики и математической геофизики Балтийского федерального университета имени И.Канта;
e-mail: KAlsynbaev@kantiana.ru.
Дата поступления — 30 апреля 2019 г.

О РЕШЕНИИ ОБРАТНЫХ ЗАДАЧ ЧИСЛЕННОЙ ОПТИМИЗАЦИИ КОМПОЗИТНЫХ КОНСТРУКЦИЙ

Е. В. Амелина[1], С. К. Голушко[1,2]

[1] *Новосибирский государственный университет, 630090, Новосибирск*
[2] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

УДК 539.3
DOI: 10.24411/9999-016A-2019-10005

Предложен и реализован эффективный подход к решению задач численной оптимизации композитных конструкций. Решена задача условной оптимизации многослойного композитного бака высокого давления при заданных ограничениях на прочность и внутренний объем конструкции. Параметрами проектирования являлись толщина стенки, форма бака и угол укладки армирующих волокон. Достоверность полученных решений подтверждена решением ряда тестовых оптимизационных задач и решением прямых задач расчета напряженно-деформированного состояния баков с найденными параметрами. Показано, что при решении широкого класса оптимизационных задач можно использовать математические модели композитных конструкций, построенные на основе соотношений классической теории оболочек Кирхгофа — Лява и уточненной теории Тимошенко.

Ключевые слова: композит, оболочка, конструкция, моделирование, оптимизация, напряженно-деформированное состояние.

Введение

Металлокомпозитные баки высокого давления являются перспективным типом конструкций ракетно-космической техники, обеспечивая их повышенную прочность и минимальный вес. Внутреннее давление является основной нагрузкой для таких баков. Анализ возможностей применения гибридных баков, состоящих из металлических сосудов (лейнеров), обеспечивающих герметичность конструкции, и композитных силовых оболочек, обеспечивающих прочность и жесткость, является весьма актуальным и перспективным направлением исследований [1, 2].

К настоящему времени можно выделить несколько подходов к решению задач оптимального и рационального проектирования композитных конструкций [3]– [11]. Наиболее распространенным является подход, при котором используются уравнения безмоментной теории оболочек и "нитяная" модель композиционного материала (КМ) [12, 13]. При таком подходе реализация безмоментного состояния конструкции по существу не обеспечивается, а только предполагается. Вопрос о том, будет ли спроектированная оболочка действительно безмоментной при безмоментных краевых условиях остается в таких работах открытым.

Другой подход связан с обоснованным упрощением исходной постановки задачи трехмерной теории упругости и использованием различных вариантов двумерных теорий оболочек [3, 14, 15].

Именно такой подход использован в данной работе. Для проверки достоверности получаемых решений использованы варианты теории оболочек Тимошенко [16] и Андреева-Немировского [17], учитывающие поперечные сдвиги с различной степенью точности. Достоверность результатов численной оптимизации обеспечена сравнением с известными частными аналитическими решениями.

1 Постановка задачи

Рассмотрим многослойный композитный сосуд давления, находящийся в равновесии под действием внутреннего давления. Необходимо определить параметры композиционного материала и конструкции, обеспечивающие выполнение следующих требований:

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 18-29-18029).

ISBN 978-5-901548-42-4

$$V \geq V_0, \quad P \geq P_0, \quad M \leq M_0, \tag{1}$$

где V — внутренний объем, P — внутреннее давление, M — масса сосуда; V_0, P_0, M_0 — некоторые заданные значения.

Будем решать задачу минимизации (или максимизации) одного из функционалов, определяющих критерий оптимальности, при фиксации оставшихся двух в виде ограничений из (1).

В общем случае математическая постановка задач оптимизации конструкций включает выбор оптимизируемого функционала, формулировку определяющих уравнений, а также ограничений на функции состояния и управляющие параметры.

Будем рассматривать бак как жестко защемленную на краях оболочку. С учетом наличия плоскости симметрии в середине бака, достаточно рассчитывать и проектировать только одну его половину. Вид нагружения и закрепления позволяет рассматривать задачу, как задачу осесимметричного деформирования.

Оболочка задается вращением образующей $r = r(\theta)$ относительно оси oy (рис. 1), где r — текущий радиус оболочки, θ — угол между нормалью к поверхности оболочки и осью вращения, который изменяется в пределах $[\theta_0; \theta_1]$.

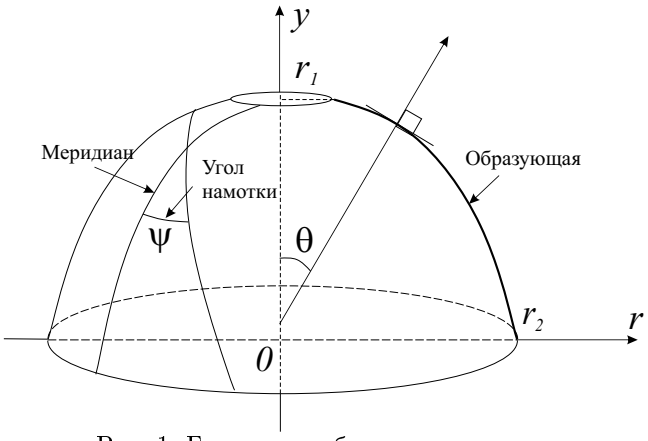


Рис. 1: Геометрия оболочки вращения

Для решения прямых задач расчета напряженно-деформированного состояния (НДС) многослойных композитных баков и верификации решений задач оптимизации будем использовать классическую теорию оболочек Кирхгофа — Лява [18] и уточненные теории Тимошенко [16] и Андреева — Немировского [17]. Физические соотношения опишем с помощью структурных моделей КМ [3].

Для решения обратных задач оптимизации будем использовать критерий минимума веса композитной оболочки:

$$M = 2\pi \int_{\theta_0}^{\theta_1} r R_1 h d\theta \left[\rho_m(1-\omega_r) + \rho_r \omega_r\right] \to \min, \tag{2}$$

где ρ_m, ρ_r — плотности материалов связующего и армирующих волокон, ω_r — объемное содержание арматуры.

В качестве управляющих функций выберем: главный радиус кривизны $R_1(\theta)$, толщину оболочки $h(\theta)$, угол спирального армирования $\psi(\theta)$.

Искомое решение должно удовлетворять ограничению на внутренний объем оболочки

$$\pi \int_{\theta_0}^{\theta_1} r^2 R_1 \sin\theta d\theta = V_0 \tag{3}$$

и требованию ее неразрушения

$$\max\{bs_r, bs_m\} \leq 1, \tag{4}$$

где bs_r, bs_m — интенсивности напряжений в связующем материале и армирующих волокнах, определяемым по рассчитанному НДС бака по одной из структурных моделей КМ.

Будем использовать следующие ограничения на функции управления:

$$0 \leq \psi \leq 90^\circ, \quad h_0^* \leq h \leq h_1^*, \quad R_0^* \leq R_1 \leq R_1^*, \tag{5}$$

где $h_0^*, h_1^*, R_0^*, R_1^*$ — константы.

При изготовлении композитных оболочек вращения широкое распространение получил метод непрерывной намотки по геодезическим линиям. В этом случае справедлива формула Клеро:

$$r \sin \psi(r) = C, \tag{6}$$

где C — константа, определяемая из условия на экваторе оболочке, а толщина оболочки определяется соотношением

$$h(r) = h_R \frac{R \cos \psi_R}{r \cos \psi(r)}, \tag{7}$$

которое имеет особенность у кромки отверстия, где угол армирования должен быть равен 90^o. Формула (7) применяется при $r \geq r_1 + r_\omega$, где величина r_ω равна ширине армирующей ленты. Соотношение, определяющее толщину стенки бака в итоге принимает вид:

$$h(r) = \begin{cases} h_R \dfrac{R \cos \psi_R}{r_\omega \cos \psi(r_1 + r_\omega)}, & r \leq r_1 + r_\omega; \\ h_R \dfrac{R \cos \psi_R}{r \cos \psi(r)}, & r \geq r_1 + r_\omega. \end{cases} \tag{8}$$

2 Прямая задача расчета НДС многослойного бака

Проблема определения НДС композитного бака приводит к необходимости решения краевых задач для жестких систем дифференциальных уравнений. Такие системы являются плохо обусловленными, а в структуре их решений имеются решения с ярко выраженным характером погранслоев. Численные расчеты были выполнены методами сплайн-коллокации и дискретной ортогонализации, реализованными в пакетах COLSYS [19] и GMDO [20]– [24] соответственно. Эти вычислительные инструменты показали высокую точность и эффективность при решении задач механики композитных пластин и оболочек [3].

Решение ряда прямых задач при различных значениях параметров проектирования и последующий анализ показали, что решение задач численной оптимизации можно осуществлять с использованием достаточно простых соотношениях теорий оболочек Кирхгофа–Лява или Тимошенко, для которых трудоемкость решения существенно (на порядок) меньше, чем с использованием уточненных теорий (например, теории Андреева – Немировского). На рис. 2 показано влияние угла спирального армирования ψ на интенсивности напряжений в связующем материале bs_m (пунктир) и арматуре bs_r (сплошная линия). Результаты, полученные с использованием теории Кирхгофа – Лява, обозначены линиями без дополнительных символов, теории Тимошенко — с $\triangle$, теории Андреева – Немировского — с $\square$.

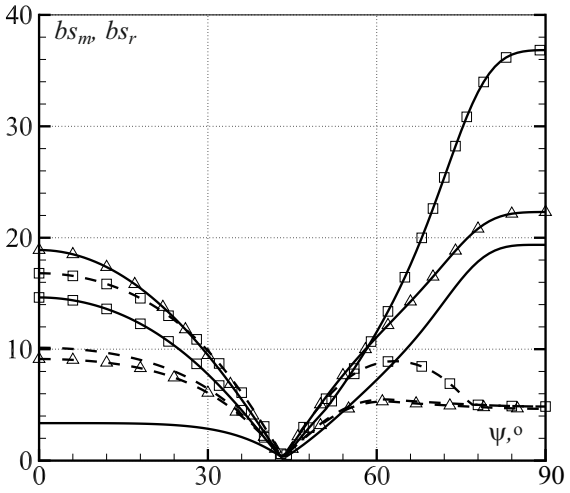


Рис. 2: Влияние угла спирального армирования на НДС композитного сосуда давления

Используя угол армирования в качестве параметра проектирования можно существенно повысить запас прочности бака: разница между наилучшим и наихудшим вариантами может отличаться более чем в 20 раз по интенсивностям напряжений в связующем bs_m и арматуре bs_r. Варианты с армированием, близким к окружному являются неудачными. В этом случае в зоне закрепления возникают значительные поперечные деформации, сопротивление которым оказывает, в основном, слабый материал матрицы, в то время как арматура практически не работает.

3 Обратная задача оптимизации многослойного бака

Рассмотрим тестовую задачу совместной оптимизации формы и толщины безмоментной осесимметричной изотропной оболочки вращения минимального веса, которая имеет аналитическое решение [29]. Здесь и далее параметры оболочки: $V_0 = 3\text{м}^3$, $r_1 = 0.5\text{м}$, $r_2 = 1\text{м}$, $q = 10^8\text{Па}$. В качестве ограничения по прочности использовано соотношение

$$\max\{|\sigma_{11}|, |\sigma_{22}|\} \le \sigma^*, \tag{9}$$

где σ^* — постоянная прочности материала.

Из уравнений равновесия безмоментной оболочки можно легко получить аналитические выражения для усилий

$$T_1 = \frac{T_{01} + \pi q(r^2 - r_1^2)\sqrt{1 + r'^2}}{2\pi r}, \qquad T_2 = qr\sqrt{1 + r'^2} + \frac{T_1 r r''}{1 + r'^2}. \tag{10}$$

Рассмотрим случай, когда $\sigma_{11} = \sigma^*$. Соотношения для радиуса и толщины оптимальной оболочки в декартовых координатах примут вид [29]:

$$r(x) = \sqrt{\alpha x + \beta x^2 + r_1^2}, \qquad h(x) = \frac{q}{2\sigma^*}\sqrt{\alpha(1+\beta)x + \beta(1+\beta)x^2 + \frac{\alpha^2}{4} + r_1^2}, \tag{11}$$

где α, β зависят от геометрических параметров конструкции.

Выберем в качестве управляющих функций главный радиус кривизны R_1 и толщину h, а в качестве переменной интегрирования — угол поворота нормали к образующей оболочки θ.

Поставленная задача является задачей условной оптимизации, включающей прямые (параллелепипедные) ограничения на управления и траекторные ограничения, наложенные в конце интервала изменения независимой переменной. Одним из распространенных подходов к решению таких задач является метод последовательной безусловной оптимизации. Он заключается в формировании свертки терминальных функционалов и многократном решении однокритериальной задачи с использованием методов оптимизации [25]. В данной работе для свертки использовался метод модифицированной функции Лагранжа.

Для решения вспомогательных задач безусловной оптимизации проводилась дискретизация управляющих функций на сетке, не связанной с сеткой решения прямой задачи, и, таким образом, искалось решение невыпуклой задачи конечномерной оптимизации [26]. Для решения задачи безусловной оптимизации был использован программный комплекс OPTCON-A [27, 28].

На рис. 3 приведено сравнение полученных численного и аналитического решений при различных значения параметра α (сплошные кривые — аналитическое решение, пунктирные — численное). В табл. 1 приведены числовые значения погрешностей.

Отличие между расчетным и точным значением массы составило 2 %, т.е. чувствительность функционала массы оболочки к изменению главного радиуса кривизны невелика.

Другая тестовая задача заключалась в поиске проекта бака минимальной массы в рамках полной моментной постановки на основе уравнений теории оболочек Кирхгофа – Лява. Значения растягивающих усилий на краях оболочки являлись искомыми параметрами, а перерезывающие усилия и моменты равнялись нулю. Максимальные относительные погрешности по компонентам искомых параметров составили для $\theta_0 = 2.3\%$, $\theta_1 = 0.8\%$, $h = 6.3\%$, $R_1 = 2.9\%$, $T_{10} = 4.0\%$, $T_{02} = 1.2\%$. Отклонение по значению целевой функции (весу) составило 0.4%.

На основании проведенных экспериментов можно заключить, что предложенный подход к численному решению оптимизационных задач является эффективным. Кроме того показано, что оптимальные решения, полученные в безмоментной постановке, могут остаться оптимальными и в более общей моментной постановке.

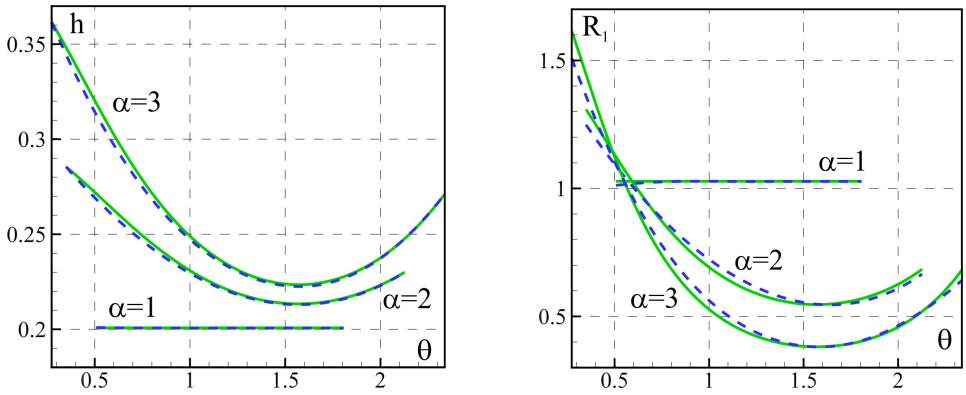


Рис. 3: Влияние параметра геометрии α на вид оптимальных решений

Таблица 1: Максимальные относительные погрешности численных решений

Функция проектирования	Максимальная относительная разность		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$
h	0.003%	1%	2%
R_1	0.003%	4%	17%

Далее были исследованы несколько различных постановок задачи оптимального проектирования. Основное их отличие заключалось в выборе функций проектирования, либо в виде функций, либо в виде констант. Дополнительно был рассмотрен проект поиска бака оптимальной формы, образованного непрерывной намоткой по геодезическим линиям.

В таб. 2 приведено сравнение проектов баков, полученных при решении оптимизационных задач. Приняты следующие обозначения: 1 — ψ , 2 — h, 3 — R_1; после "F" идут номера функций управления, которые задавались как функции, после "C" — функции управления — константы. Сокращением "Geod" обозначен проект с непрерывной намоткой по геодезическим линиям.

Таблица 2: Сравнение различных проектов многослойных баков

Проект	Масса, кг	M/M_{C123}	$\psi_{\max} - \psi_{\min}$	$h_{\max}/h_{\min}$	$R_{1_{\max}}/R_{1_{\min}}$
F123	16.04	72.5%	4	1.08	1.62
F12C3	16.02	72.4%	12	1.55	1.00
F13C2	16.05	72.5%	2	1.00	1.94
F23C1	17.01	76.9%	0	4.74	3.14
F1C23	20.23	91.4%	15	1.00	1.00
F2C13	16.42	74.2%	0	1.83	1.00
F3C12	16.07	72.6%	0	1.00	1.83
Geod	19.09	86.2%	85	9.95	5.42
C123	22.13	100.0%	0	1.00	1.00

Масса оптимизированного бака, полученного непрерывной намоткой волокон по геодезическим линиям (Geod) составила 19 кг, однако за счет выбора оптимальной формы бака (проект F3C12) можно добиться снижения веса до 16 кг.

Отметим, что проект с геодезической непрерывной намоткой обладает перепадом толщины стенки бака почти в 10 раз с резким градиентом вблизи места закрепления. Проект F123 более предпочтителен, перепад значений параметров не превышает 10% по толщине и 4.5% по углу армирования. Представляют интерес проекты с постоянной толщиной стенки бака (F13C2, F3C12), показывающие возможность получения конструкций с массой близкой к минимальной за счет управляющих функций R_1 и ψ.

Для проверки решения оптимизационной задачи полученные параметры управления были подставлены в прямую задачу расчета НДС конструкции.

Среди полученных решений можно выделить проект F123, НДС которого является практически безмоментным, а арматура почти равнонапряженной. В этом проекте влияние поперечных сдвигов минимально.

Подставив это решение в задачу расчета НДС в рамках различных вариантов теорий оболочек, обнаружим, что все три теории дают близкие результаты при расчете интенсивностей напряжений в связующем материале и арматуре.

Заключение

Разработан эффективный подход к оптимизации многослойных композитных сосудов давления, позволяющий строить проекты конструкций, удовлетворяющих заданным требованиям и обладающих рядом дополнительных ценных свойств: малым изменением параметров проектирования, напряженным состоянием близким к безмоментному и с почти равнонапряженной арматурой.

Полученные решения оптимизационных задач верифицированы решением прямых задач расчета НДС баков с найденными параметрами в рамках математических моделей, построенных на основе классической и уточненных теорий. Показано, что при решении оптимизационных задач можно использовать относительно простые математические модели облочек, построенные на основе соотношений теорий Кирхгофа — Лява и Тимошенко.

Список литературы

[1] Thesken J. C., Murthy P. L. N., Phoenix S. L. et al. A theoretical investigation of composite overwrapped pressure vessel (COPV) mechanics applied to NASA full scale tests. NASA/TM–2009—215684, 2009.

[2] Амелина Е. В., Буров А. Е., Голушко С. К., Лепихин А. М., Москвичев В. В., Юрченко А. В. Расчетно-экспериментальная оценка прочности металлокомпозитного бака высокого давления // Вычислительные технологии. 2016. Т. 21, № 5. С. 3–21.

[3] Голушко С. К., Немировский Ю. В. Прямые и обратные задачи механики упругих композитных пластин и оболочек вращения. Москва: ФИЗМАТЛИТ. 2008.

[4] Голушко С. К., Немировский Ю. В. Обзор и анализ подходов к проблеме рационального проектирования армированных оболочек. Красноярск. 1988. — 31 с. (Препринт ВЦ СО АН СССР; № 16).

[5] Голушко С. К., Немировский Ю. В. Об одном подходе к проектированию осесимметричных оболочек с равнонапряженной арматурой // Численные методы решения задач теории упругости и пластичности / Материалы X Всес. конф. Новосибирск: ИТПМ СО РАН, 1988. С. 58–64.

[6] Голушко С. К., Немировский Ю. В. Построение проектов армированных оболочечных конструкций минимального веса // Вычислительные проблемы механики / Межвуз. сб. Красноярск: Изд-во: КГУ, 1989. С. 117–130.

[7] Голушко С. К. Две задачи рационального проектирования армированных оболочек вращения. Красноярск. 1986. С. 8–14. (Препринт ВЦ СО АН СССР; № 1).

[8] Голушко С.К. Проектирование тонкостенных оболочек с равнонапряженной арматурой. Красноярск. 1987. С. 29–32. (Препринт ВЦ СО АН СССР; № 1).

[9] Golushko S.K. Direct and inverse problems in mechanics of composite shells // Proc. of The Sixth Japan-Russia Joint Symposium on Computational Fluid Dynamics. Nagoya: Nagoya University, 1998. P. 125–130.

[10] Голушко С.К., Немировский Ю.В., Одновал С.В. Расчет и рациональное проектирование композитных оболочек вращения // Динамика сплошной среды. 1998. № 113. С. 39–44.

[11] Golushko S. K. Direct and inverse problems in the mechanics of composite plates and shells // Computational Science and High Performance Computing. Russian-German Advanced Research Workshop, Novosibirsk, Russia. Germany: Springer-Verlag, 2005. P. 205–228.

[12] Stadler W., Krishnan V. Natural structural shapes for shells of revolution in the membrane theory of shells // Structural Optimization. 1989. V. 1. P. 19–27.

[13] Vasiliev V. V. Composite pressure vessels: design, analysis, and manufacturing. Blacksburg, Virginia, USA: Bull Ridge Publishing. 2009.

[14] Liang C.C., Chen H. W., Wang C. H. Optimum design of dome contour for filament-wound composite pressure vessels based on a shape factor // Compos. Struct. 2002. V. 58, iss. 4. P. 469–482.

[15] Zu L., Koussios S., Beukers A. Shape optimization of filament wound articulated pressure vessels based on non-geodesic trajectories // Compos. Struct. 2010. V. 92, iss. 2. P. 339–346.

[16] Григоренко Я. М., Василенко А. Т. Задачи статики анизотропных неоднородных оболочек. Москва: Наука. 1992.

[17] Андреев А. Н., Немировский Ю. В. Многослойные анизотропные оболочки и пластины: Изгиб, устойчивость, колебания. Новосибирск: Наука. 2001.

[18] Новожилов В. В. Теория тонких оболочек. Л.: Судпромгиз. 1951.

[19] Ascher U., Christiansen J., Russel R. D. Collocation software for boundary value ODEs // ACM. Trans. on Math. Software. 1981. V. 7, iss. 2. P. 209–222.

[20] Голушко С. К., Горшков В. В., Юрченко А. В. О двух численных методах решения многоточечных нелинейных краевых задач // Вычислительные технологии. 2002. Т. 7, № 2. С. 24–34.

[21] Голушко С. К., Морозова Е. В., Юрченко А. В. О численном решении краевых задач для жестких систем дифференциальных уравнений // Вестник КазНУ. Серия: Математика, механика, информатика. 2005. № 2. С. 12–26.

[22] Голушко С. К., Горшков В. В., Юрченко А. В. О двух численных методах расчета сопряженных композитных конструкций // В кн.: Современные методы математического моделирования природных и антропогенных катастроф. Тезисы докладов V науч. конференции. 1999. С. 49–55.

[23] Golushko S. K., Yurchenko A. V. Solution of boundary value problems in mechanics of composite plates and shells // Russian Journal of Numerical Analysis and Mathematical Modelling. 2010. V. 25, iss. 1. P. 27–55.

[24] Голушко С. К., Идимешев С. В., Семисалов Б. В. Методы решения краевых задач механики композитных пластин и оболочек. Новосибирск: КТИ ВТ СО РАН, 2014. 131 с.

[25] Бертсекас Д. Условная оптимизация и метод множителей Лагранжа. Москва: Радио и связь. 1987.

[26] Евтушенко Ю. Г. Методы решения экстремальных задач и их применение в системах оптимизации. Москва: Наука. 1982.

[27] Горнов А. Ю. Вычислительные технологии решения задач оптимального управления. Новосибирск: Наука. 2009.

[28] Амелина Е. В., Голушко С. К., Горнов А. Ю., Юрченко А. В. Задачи проектирования многослойных гибридных сосудов давления // В сб.: Безопасность и живучесть технических систем. Материалы и доклады: в 3-х томах. 2015. С. 34–40.

[29] Баничук Н. В. Оптимизация осесимметричных мембранных оболочек // Прикладная математика и механика. 2007. № 1. Р. 19–27.

Амелина Евгения Валерьевна — к.ф.-м.н., ст. науч. сотр. Новосибирского государственного университета;
e-mail: amelina.evgenia@gmail.com;
Голушко Сергей Кузьмич — д.ф.-м.н., профессор Новосибирского государственного университета;
глав. науч. сотр. Института вычислительных технологий СО РАН;
e-mail: s.k.golushko@gmail.com.
Дата поступления — 30 апреля 2019 г.

АСПЕКТЫ ВЫЧИСЛЕНИЯ ДЕТЕРМИНАНТОВ МАТРИЦ И ИХ РАСПРЕДЕЛЕНИЙ ПО ОПРЕДЕЛИТЕЛЯМ

Н. А. Антипин [1,2]

[1] *Московский технологический университет (МИРЭА), 119454, Москва*
[2] *Московский физико-технический институт, 141701, Москва*

УДК 519.688
DOI: 10.24411/9999-016A-2019-10006

В статье рассматриваются данные, полученные с помощью алгоритмов, разработанных с использованием встроенных типов данных в стандартных языках программирования. Приводится распределения матриц из 0 и 1 по детерминантам для 8-го порядка — пересчетом определителей всех и для порядка с 9-го по 11-ый точечно (используется метод Монте–Карло и ищутся лишь отдельные детерминанты, без указания количества для данного порядка). А также рассматриваются обобщения распределений на матрицы более больших классов элементов. Результаты данных вычислений могут быть полезны при передачи информации в открытых сетях, а также в естественных теоретических науках.

Ключевые слова: распределения определителей матриц данного порядка, функции распределений определителей матриц, встроенные типы данных.

Введение

Данная работа основана на результатах вычислений распределений определителей матриц и последующем обобщении этих данных. В первой секции представлена таблица с распределением матриц из 0 и 1 8-го порядка, а также некоторые пояснения как оно вычислялось. В следующей секции приведены обобщения некоторых функций на случай матриц из кольца вычетов с более крупным модулем, а также уточнения для распределений матриц из 0 и 1 произвольного порядка.

1 Вычисление распределений определителей

Вычисление распределений матриц из нулей и единиц по определителям порядок, которых выше 7 представляет собой не совсем ординарную задачу, так как в дополнение к аппаратным и алгоритмическим требованиям появляется трудно распознаваемая ошибка, связанная с погрешностью вычислений детерминантов матриц этого порядка. Встроенные типы данных не могут выполнять вычисления с достаточной точностью и нулевой элемент будет распознан неверно, но в результате мы все равно получим детерминант равный целому числу по размеру не сильно отличающийся от правильного значения, что отражено в таблице 1 [3] символом $\geq$. Для её устранения требуется искать ненулевой элемент больший этой погрешности.

В таблице 1 представлено распределение для ненулевых базисных матриц (4 на 8) [1]. Чтобы получить результирующее распределение необходимо к количеству матриц с определителем равным 0 добавить 189292576 (количество нулевых матриц 4 на 8), умноженное на 2^{32}.

В приведенных ниже таблицах показаны точечные распределения модулей определителей, то есть для данного числа определяется существование равного ему определителя матрицы. Возможны и другие значения определителей особенно в матрицах большего порядка.

ISBN 978-5-901548-42-4

Таблица 1: Распределение модулей определителей для матриц 8-го порядка

0	1	2	3
747327888707194368 0	3766962568171582080	3107221856321587200	1128344550375409920
4	5	6	7
1062603277179171840	280558398045864960	39183998133030912 0	92717618729258880
8	9	10	11
148970450186127360	52121937402380160	49370839378882560	14214384381012480
12	13	14	15
33604944564925440	6076097931697920	866161820385792 0	4660876422921600
16	17	18	19
5455580275353600	1206477746611200	260134381785600 0	548565282316800
20	21	22	23
1183920922598400	420654153830400	358862424883200	113325986995200
24	25	26	27
354614638694400	62087327366400	73204159795200	39723123686400
28	29	30	31
49098143001600	9118971187200	28015323033600	3621252096000
32	33	34	35
11081266560000	2806470374400	2831160729600	757170892800
36	37	38	39
3116471500800	131681894400	592568524800	65840947200
40	42	44	45
589825152000	65840947200	76814438400	2743372800
48	56		
32920473600	391910400		

Таблица 2: Точечное распределение модулей определителей матриц 9-го порядка

{0 -102} 104 105 108 110 112 116 117 120 125 128 144

2 Обобщение на случай матриц с элементами из Z_m

Для квадратной матрицы порядка n > 2 с элементами из кольца вычетов Z_m $[6, 7, 9]$ $(k1, k2, ..., k(m-1))$ — соответственно количество элементов $1, 2, ..., (m-1)$ в данной матрице, $N = n^2$) определим функцию $Factor(C_N^{k1} * C_N^{k2} * \cdots * C_N^{k(m-1)})$, равную произведению отношения общего числа элементов в матрице к количеству различных сочетаний каждого из набора элементов кольца вычетов Z_m по общему количеству элементов в матрице на весовой коэффициент (смотри рис. 1 и [8]), характеризующий распределение определителей для данного набора $k1, k2, ..., k(m-1)$ (предполагается, что каждый элемент может находится в любом месте матрицы). В случае $m = 2$: k — количество единиц в матрице и $Factor(C_N^k)$ есть $m^N/C_N^k * Ves(k)$ $= 2^N/C_N^k * Ves(k)$. В данном случае $Factor(C_N^k)$, умноженный на отношение общего количества элементов к максимальному детерминанту для данного набора элементов есть модуль среднего детерминанта этих элементов. А логарифм $Ves(k)$ есть их средняя энтропия $(S(k) = ln(Ves(k)))$.

Обстоятельные работы по теории полей приведены в [9] упомянутые там работы Голомба дают еще одно применение для матриц — генерация псевдослучайных последовательностей. Некоторые оценки в случае по-

Таблица 3: Точечное распределение модулей определителей матриц 10-го порядка

{0-246} {248-252} {254-267} 270 272 273 275 276 279 280 282 284 285 288 291 296 297 312 315 320

Таблица 4: Точечное распределение модулей определителей матриц 11-го порядка

{0-600} 603 604 606 608 609 611 612 {614-616} {618-621} 624 625 628 630 632
{634-636} 639 640 642 648 650 651 655 659 {664-666} 672 680 689 696 704 728 735

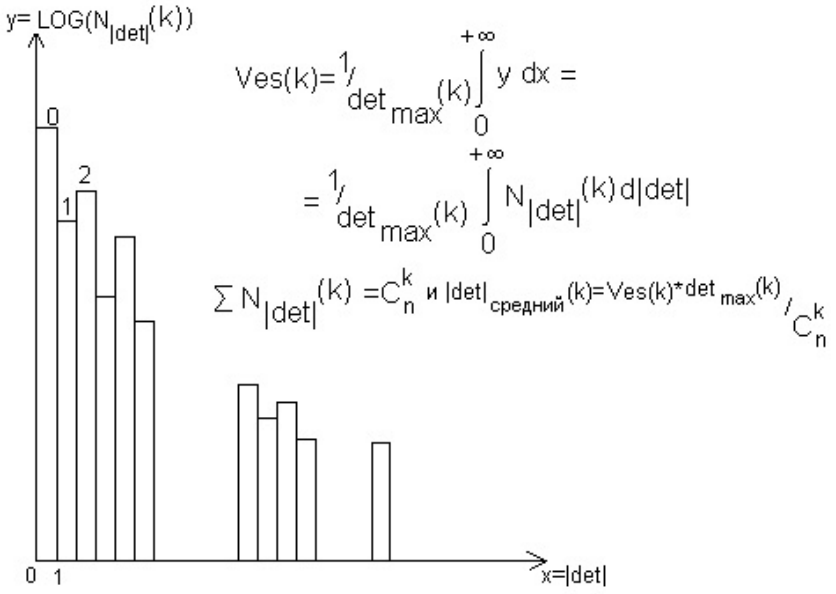


Рис. 1: Зависимость количества модулей определителя от их значения. А также основные функции, характеризующие эту зависимость

строения типа <Алёнушка> [1] для матриц 32-го порядка с элементами из Z_2 (аналогичные оценки должны быть справедливы и в случае, указанных там матриц большего порядка) приведены в [4], то есть период, полученной последовательности может быть меньше в несколько раз в случае Z_2^n. При построении рандомзирующих матриц, заполнение, как и в случае с инкапсуляцией данных, следует проводить матрицами с как можно большей энтропией (порядком). 8-уровневые схемы, показанные в [4] для инкапсуляции данных этим свойством не обладают. В [9] упоминается, что генерация с помощью матриц является громоздкой, но в случае матриц с элементами из Z_2 это далеко не так. Когда же научатся на процессорном уровне синхронно в параллельном режиме выполнять действия как показано в [2] с последующей ротацией модулей XOR, то это ещё в большей степени решит эту проблему.

Для матриц 32-го порядка, в достаточно приближенной степени, для отношения 1 к 0 равному (1:1) средний модуль детерминанта — $5*10^8$, для (3:1) — $5*10^7$, для (5:1) — 10^5. И при отклонении большем примерно 39% числа 1 отношения числа единиц к общему числу элементов в матрице из 0 и 1 — модуль детерминанта в среднем становится все меньше (не учитываются более тонкие аспекты, связанные с четностью модуля детерминанта (смотри распределение для $n = 8$)). Так что функция $Factor(C_n^k)$ не является линейной по отношению к модулю определителя. При $n = 64$ отношение $|det|_{\text{средний}}$ для $k1/n^2 = 0,5$ к $|det|_{\text{средний}}$ для $k2/n^2 = 0,75 \approx 10^{25}/10^{20} = 100000$.

По определению [5]

$$|det| = |a_{ik}| = \{\{a_{11}, ..., a_{1n}\}, \{a_{21}, ..., a_{2n}\}, ..., \{a_{n1}, ..., a_{nn}\}\} = \sum_j (-1)^{t(j)} a_{1j1} * a_{2j2} * ... * a_{njn}.$$

Для матриц из 0 и 1 в случае наличия в матрице $n * n$ — n единиц из этого определения вытекает $|det|_{max} = 1$, а в случае n нулей $|det|_{max} = n-1$. Так что $|det|_{\text{средний}}(k)$ не является симметричной функцией относительно вертикальной оси проведенной по уровню 0,5 абсциссы.

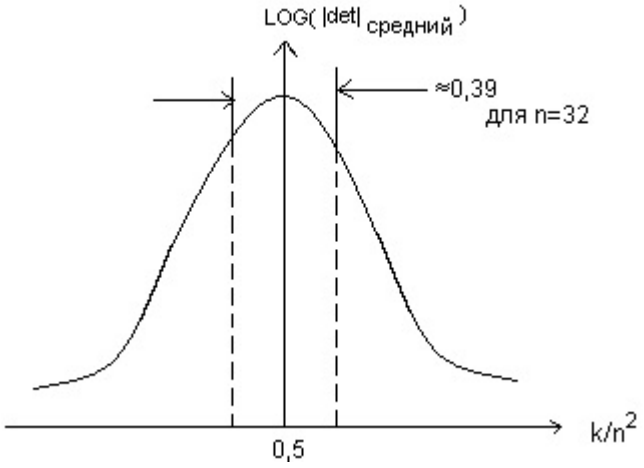


Рис. 2: Схематичная зависимость $|det|_{\text{средний}}$ от k

Заключение

В работе показаны распределения матриц из 0 и 1 по детерминантам для 8-го порядка — пересчетом определителей всех и для порядка с 9-го по 11-ый точечно, а также некоторые аспекты самих вычислений и соответствующие им соотношения. Обсуждаются применения матриц в методах генерации псевдослучайных последовательностей.

Список литературы

[1] Антипин Н. А. Аспекты шифра Хилла для разреженных матриц Альманах мировой науки, 2016 г, № 10-1(13), с. 7–13.

[2] Антипин Н. А. Модификации шифра Хилла для матриц 32-го порядка // Тенденции развития науки и образования: Сборник научных трудов, 31 июля 2015 г, с. 13–19, часть 1, ISBN 978-5-9906987-6-5.

[3] Антипин Н. А. Разреженные матрицы из нулей и единиц // Перспективы развития науки и образования: Сборник научных трудов, 30 июня 2017 г, с. 6–10, ISBN 978-5-9500654-2-2.

[4] Антипин Н. А. Теоретические и практические аспекты матриц из нулей и единиц // Альманах мировой науки, 2018 г, № 1(21), с. 6–35.

[5] Бугров Я. С., Никольский С. М. Элементы линейной алгебры и аналитической геометрии. 3-е изд., испр. и доп. — М.: Наука, 2013. 222 с.

[6] Виноградов И. М. Основы теории чисел: Учебное пособие. 12-е изд., стер. — СПб.: Издательство “Лань”, 2009. 176 с.

[7] Глухов М. М., Круглов И. А., Пичкур А. Б., Черемушкин А. В. Введение в теоретико-числовые методы криптографии: Учебное пособие. — СПб.: Издательство “Лань”, 2011. 400 с.

[8] Фихтенгольц Г. М. Основы математического анализа. Часть 1, 2. 4-е изд. — СПб.: Издательство “Лань”, 2002.

[9] Crandall R., Pomerance C. Prime Numbers. A Computational Perspective. — М.: URSS, 2011. — 664 p. ISBN 978-5-453-00016-6.

Антипин Николай Александрович — магистр Московского технологического университета;
e-mail: myantip@yandex.com.
Дата поступления — 30 апреля 2019 г.

МОДЕЛИРОВАНИЕ ДИНАМИКИ ГАЗОСОДЕРЖАНИЯ НЕФТИ В ПРОЦЕССЕ РАЗРАБОТКИ МЕСТОРОЖДЕНИЙ

А. В. Базовкин

ООО "СамараНИПИнефть", 443010, Самара

УДК 622.276.346.2+004.942+51-74
DOI: 10.24411/9999-016A-2019-10007

Предлагается метод моделирования изменения содержания попутного нефтяного газа в добываемом сырье. Метод основан на моделировании расхода попутного нефтяного газа на объектах подготовки, снабжённых системами измерения количества газа. Изменение газосодержания пластовой нефти, а также изменение доли попутного нефтяного газа в добываемом сырье моделируется на основе 0-мерной гидродинамической модели с учётом отношения подвижностей нефтяной и газовой фаз.

Ключевые слова: газосодержание нефти, газовый фактор, попутный нефтяной газ, узел учёта газа.

Введение

При разработке нефтяных месторождений дополнительно добываемым сырьем является попутный нефтяной газ (ПНГ), содержащий лёгкие углеводороды, а также азот и иные примеси. В зависимости от термобарических условий в пласте и компонентного состава нефти начальное газосодержание может варьироваться в широком диапазоне — от 0 до 1100 м3/т. Иногда в процессе разработке нефтяных месторождений пластовое давление снижается ниже давления насыщения нефти, что вызывает процесс выделения газа в пласте. С ростом газонасыщенности внутри порового объёма пласта начинается фильтрация газа к забоям добывающих скважин, что может вызвать резкий рост добычи ПНГ. Данный эффект является нежелательным с точки зрения разработки нефтяных месторождений как по причине ускоренного снижения пластового давления, так и по причине увеличения вязкости дегазированной нефти и снижения её подвижности. В этой связи задача моделирования динамики газосодержания нефти представляет интерес для нефтяной промышленности. Также значительный интерес представляет и смежная задача моделирования динамики добычи ПНГ. Её практический смысл состоит в уточнении плановых показателей для целей бизнес-планирования, в том числе с целью соблюдения действующего законодательства в отношении доли полезного использования ПНГ [1].

Кратко опишем процесс добычи ПНГ. При разработке нефтяных залежей сырьё по скважинам поднимается на поверхность Земли, где по системе сборных трубопроводов транспортируется на объекты подготовки нефти. На объектах подготовки добываемая продукция подвергается сепарации на нефть, воду и выделившийся из нефти попутный нефтяной газ. Полученные в результате сепарации объёмы ПНГ измеряются системой измерения количества газа (СИКГ). Показания СИКГ являются наиболее достоверным источником информации о динамике добычи ПНГ. Достоинством данных замеров является то, что они обычно ведутся в круглосуточном режиме и производятся после глубокой сепарации добываемой жидкости. Недостатком же таких замеров является то, что на объект подготовки нефти может поступать продукция одновременно с нескольких месторождений, насчитывающих, иногда, десятки пластов, которые характеризуются различными свойствами нефти. В таком случае становится затруднительным воспроизвести динамику добычи ПНГ отдельного пласта на основе замеров СИКГ. С целью решения данной проблемы в настоящей работе расход газа на СИКГ моделируется на основе 0-мерных гидродинамических моделей отдельных пластов. Так на основе моделирования динамики замеров СИКГ делается попытка описать процессы изменения газосодержания нефти в недрах Земли. Проблема прогнозирования уровней добычи ПНГ подробнее рассмотрена в работе [2].

ISBN 978-5-901548-42-4

1 Исходные уравнения

Пусть имеется некоторая установка подготовки нефти (УПН), на которой производится сепарация продукции на воду, нефть и ПНГ, причём уровни добычи последнего измеряются на СИКГ. Будем предполагать, что для пластов, добыча с которых поступает на УПН, известны все необходимые начальные геолого-физические характеристики (включая свойства нефти, сжимаемость породы и т.д.). Добыча ПНГ за n-ый календарный месяц, поступающая на УПН, равна

$$Q_g^n = \sum_i \alpha_i^n q_{oi}^n GOR_i^n \tag{1}$$

где α_i^n — доля добычи нефти i-го пласта в n-ый месяц, поступающая на рассматриваемую УПН, q_{oi}^n — объёмная добыча нефти i-го пласта в n-ый месяц, GOR_i^n — газонефтяной фактор i-го пласта в n-ый месяц. Газонефтяной фактор определяется как отношение

$$GOR_i^n = \frac{q_{gi}^n}{q_{oi}^n}$$

где q_{gi}^n — объёмная добыча ПНГ i-го пласта в n-ый месяц. Суммирование в (1) ведётся по всем пластам, добыча с которых поступает на рассматриваемую УПН. Тогда проблему моделирования динамики добычи ПНГ на УПН можно сформулировать следующим образом:

$$|Q_g^n - Q_{gm}^n| \to \min \tag{2}$$

где Q_{gm}^n — объёмная добыча газа, замеряемая на УПН в n-ый месяц. Очевидно, что при разработке нефтяной залежи при забойном давлении выше давления насыщения газонефтяной фактор равен начальному газосодержанию нефти. Моделирование изменения газонефтяного фактора при снижении пластового давления ниже давления насыщения представляет основную сложность при решении задачи (2).

1.1 Модель пласта

Для моделирования динамики газонефтяного фактора при давлении ниже давления насыщения используется соотношение, являющееся следствием уравнения Дарси [3, с. 99]

$$GOR_i^n = \frac{q_{gi}^n}{q_{oi}^n} = \frac{q_{gfi}^n + q_{gsi}^n}{q_{oi}^n} = \frac{k_{rg} B_o(p) \mu_o}{k_{ro} B_g(p) \mu_g} + R_s(p) \rho_{o0} \tag{3}$$

где q_{fgi}^n — объёмная добыча газа, фильтрующегося к забоям скважин в свободном виде, q_{sgi}^n — объёмная добыча газа, фильтрующегося к забоям скважин в растворённом виде, k_{ro}, k_{rg} — относительные фазовые проницаемости нефти и газа, μ_o, μ_g — коэффициенты динамической вязкости нефти и газа, B_o, B_g — объёмные коэффициенты нефти и газа, R_s — газосодержание нефти, ρ_{o0} — плотность нефти в поверхностных условиях, p — давление внутри порового пространства, называемое также пластовым давлением. $B_o(p)$ рассчитывается на основе эмпирической зависимости Стендинга [4, с. 79]. $B_g(p)$ определяется на основе обобщённого уравнения Менделеева-Клайперона с использованием эмпирической зависимости Брилла-Бегса-Стендинга для коэффициента сверхсжимаемости [5, с. 78]. Зависимость газосодержания нефти от пластового давления описывается на основе полуэмпирической зависимости [4, с. 79]

$$R_s(p) = R_{sb} \left(\frac{p}{p_b} \right)^{1{,}204}, \quad p < p_b$$

где p_b — давление насыщения, R_{sb} —газосодержание при давлении насыщения.

Для определения динамики газонефтяного фактора на основе выражения (3) необходимо произвести расчёт динамики пластового давления. Для этой цели может использоваться как простейшая 0-мерная модель, так и более сложные двумерные или трёхмерные гидродинамические модели. Для создания и расчёта последних на практике обычно используются коммерческие комплексы программ (в качестве примера можно привести такие программы как T-Navigator [6], Eclipse [7], Tempest [8], Stars [9]), или собственные программные разработки нефтяных компаний [10, 11]. Многие из таких программ позволяют не только решать задачи трёхмерной фильтрации, но позволяют моделировать термические процессы, а также учитывать многокомпонентность состава пластового флюида. Однако в силу того, что построение пространственных

гидродинамических моделей требует существенных трудозатрат, а в большей степени по причине того, что расчёт таких моделей занимает продолжительное время (характерное время расчёта гидродинамических моделей измеряется часами) в настоящей работе для описания динамики пластового давления используются 0-мерные модели. Отметим, что при решении задачи (2) часто необходимо моделирование динамики газонефтяного фактора одновременно для нескольких пластов, поэтому в случае использования пространственных гидродинамических моделей расчёт функционала в левой части (2) мог бы занять часы и даже десятки часов, при том, что решение задачи (2) подразумевает многократное вычисление данного функционала.

0-мерные динамические модели нефтяной или газовой залежи в геологии приято называть "моделью материального баланса" данной залежи. При этом обычно применяется упрощённая алгебраическая форма уравнений (см. напр. [12,13]). Нередко залежи нефти или газа окружены в пласте значительными объёмами воды, влияние которых следует учитывать при моделировании динамики пластового давления. Ниже дан вывод системы дифференциальных уравнений, описывающих динамику пластового давления для системы "нефтяной пласт" — "водоносный пласт". Данная модель содержит две ячейки, между которыми возможен переток воды, величина которого определяется разностью давлений между ячейками, а также специальным коэффициентом, характеризующим продуктивность водоносного пласта. Исходя из того, что поровый объём равен объёмам заполняющих его флюидов можем записать:

$$V_p\left(p(t)\right) = \frac{m_o(t) + m_{gs}(t)}{\rho_o\left(p(t)\right)} + \frac{m_w(t)}{\rho_w\left(p(t)\right)} + \frac{m_g(t)}{\rho_g\left(p(t)\right)} \tag{4}$$

где V_p — поровый объём, m_o, m_g, m_{gs}, m_w — масса нефти, свободного газа, растворённого газа и воды соответственно, ρ_o, ρ_g, ρ_w — плотность нефти, свободного газа и воды в пластовых условиях, t — время. Основное уравнение материального баланса описывает изменение порового объёма пласта при изменении пластового давления. После дифференцирования (4) получаем

$$\frac{dV_p}{dp}\frac{dp}{dt} = \left[\frac{1}{\rho_o}\frac{d(m_o + m_{gs})}{dt} - \frac{m_o + m_{gs}}{\rho_o^2}\frac{d\rho_o}{dp}\frac{dp}{dt}\right] + \left[\frac{1}{\rho_w}\frac{dm_w}{dt} - \frac{m_w}{\rho_w^2}\frac{d\rho_w}{dp}\frac{dp}{dt}\right] + \left[\frac{1}{\rho_{gf}}\frac{dm_g}{dt} - \frac{m_g}{\rho_g^2}\frac{d\rho_g}{dp}\frac{dp}{dt}\right]. \tag{5}$$

Коэффициенты сжимаемости для компонентов флюида

$$c_k(p) = -\frac{1}{V_k}\frac{dV_k}{dp} = -\frac{1}{m_k/\rho_k}\frac{d\left(m_k/\rho_k\right)}{dp} = \frac{1}{\rho_k}\frac{d\rho_k}{dp}, \quad k \in \{o, g, w\}$$

Коэффициент сжимаемости породы [14, с. 70]:

$$c_f(p) = \frac{1}{V_p}\frac{dV_p}{dp},$$

Тогда (5) можно переписать в виде

$$c_f V_p \frac{dp}{dt} = \left[\frac{1}{\rho_o}\frac{d\left(m_o + m_{gs}\right)}{dt} - \frac{m_o + m_{gs}}{\rho_o}c_o\frac{dp}{dt}\right] + \left[\frac{1}{\rho_w}\frac{dm_w}{dt} - \frac{m_w}{\rho_w}c_w\frac{dp}{dt}\right] + \left[\frac{1}{\rho_g}\frac{dm_g}{dt} - \frac{m_g}{\rho_g}c_g\frac{dp}{dt}\right]$$

с учётом уравнения (4) и с учётом того, что $m_{gs} = m_o R_s(p)\rho_{g0}$, после перегруппировки слагаемых получим

$$\left[\left(c_f + c_o\right)\frac{m_o\left(1 + R_s(p)\rho_{g0}\right)}{\rho_o} + \left(c_f + c_w\right)\frac{m_w}{\rho_w} + \left(c_f + c_g\right)\frac{m_g}{\rho_g} - \frac{m_o\rho_{g0}}{\rho_o}\frac{dR_s}{dp}\right]\frac{dp}{dt} = \\ = \frac{1 + R_s(p)\rho_{g0}}{\rho_o}\frac{dm_o}{dt} + \frac{1}{\rho_w}\frac{dm_w}{dt} + \frac{1}{\rho_g}\frac{dm_g}{dt}. \tag{6}$$

Плотности компонентов флюида в пластовых условиях ρ_k выражаются через соответствующие значения плотностей в поверхностных условиях ρ_{k0} соотношениями:

$$\rho_k(p) = \frac{\rho_{k0}}{B_k(p)}, \quad k \in \{o, g, w\},$$

С учётом этого уравнение материального баланса (6) окончательно можно записать в виде

$$\left[(c_f + c_o)\,\frac{m_o\,(1 + R_s(p)\rho_{g0})}{\rho_{o0}}B_o + (c_f + c_w)\,\frac{m_w}{\rho_{w0}}B_w + (c_f + c_g)\,\frac{m_g}{\rho_{g0}}B_g - \frac{m_o\rho_{g0}}{\rho_{o0}}B_o\frac{dR_s}{dp}\right]\frac{dp}{dt} = \\ = \frac{1 + R_s(p)\rho_{g0}}{\rho_{o0}}B_o\frac{dm_o}{dt} + \frac{1}{\rho_{w0}}B_w\frac{dm_w}{dt} + \frac{1}{\rho_{g0}}B_g\frac{dm_g}{dt}. \tag{7}$$

Аналогичное выражение для водоносного пласта примет вид:

$$(c_f + c_w)\,m_{wA}\frac{dp_A}{dt} = \frac{dm_{wA}}{dt}, \tag{8}$$

где m_{wA}, p_A — масса воды и давление в водоносном пласте соответственно. Для описания притока воды из водоносного пласта используется простая модель фильтрации:

$$\frac{dm_{wA}}{dt} = -k_e\,(p_A - p) \tag{9}$$

где k_e — проводимость между водоносным и нефтяным пластами. Связь между накопленной добычей и текущими запасами нефти, воды и газа в пласте задаётся уравнениями

$$m_o(t) = m_{oi} - m_{op}(t) \tag{10}$$

$$m_w(t) = m_{wi} - m_{wp}(t) + m_{winj}(t) + (m_{wAi} - m_{wA}(t)) \tag{11}$$

$$m_g = m_{gi} + (m_{oi}R_{si} - m_o(t)R_s(p))\,\rho_{g0} - m_{gp}(t) \tag{12}$$

где m_{oi}, m_{wi}, m_{gi} — начальная масса в пласте нефти, воды и свободного газа соответственно, m_{wAi} — начальная масса воды в водоносном пласте, m_{op}, m_{wp}, m_{gp} — накопленная добыча нефти, воды и свободного газа соответственно, m_{winj} — накопленная закачка воды в пласт, p_i — начальное пластовое давление, R_{si} — начальное газосодержание нефти. Уравнения (7)-(12) образуют замкнутую систему относительно неизвестных p, p_A, m_o, m_w, m_g, m_{wA}. Как отмечалось, зависимости величин B_o, B_w, B_g, R_s от пластового давления p предполагаются известными.

1.2 Преобразование уравнений

Для преобразования уравнения (8) введём замену переменных

$$\varphi(p_A) = exp\left[\int_{p_{Ai}}^{p_A}(c_f(x) + c_{wA}(x))\,dx\right] \tag{13}$$

тогда

$$\frac{d\varphi(p_A)}{dp_A} = (c_f(p_A) + c_w(p_A))\,\varphi(p_A) \tag{14}$$

и уравнение (8) можно переписать в виде:

$$\frac{1}{\varphi(p_A)}\frac{d\varphi(p_A)}{dt} = \frac{1}{m_{wA}}\frac{dm_{wA}}{dt}.$$

Интегрируя последнее выражение по времени получим:

$$\ln\varphi(p_A) = \ln m_{wA} - \ln m_{wAi},$$

откуда

$$\varphi(p_A)m_{wAi} = m_{wA}. \tag{15}$$

С учётом (15) и (14) уравнение (9) примет вид:

$$m_{wAi}\,(c_f(p_A) + c_w(p_A))\,\varphi(p_A)\frac{dp_A}{dt} = -k_e\,(p_A - p)$$

или

$$\frac{dp_A}{dt} = \frac{-k_e\left(p_A - p\right)}{m_{wAi}\left(c_f(p_A) + c_w(p_A)\right)\varphi(p_A)}. \tag{16}$$

Уравнение (11) с учётом (15) примет вид:

$$m_w(t) = m_{wi} - m_{wp}(t) + m_{winj}(t) + m_{wAi}\left(1 - \varphi(t)\right). \tag{17}$$

Уравнение (12) с учётом (10) примет вид:

$$m_g = m_{gi} + \left(m_{oi}R_{si} - \left[m_{oi} - m_{op}(t)\right]R_s(p)\right)\rho_{g0} - m_{gp}(t). \tag{18}$$

В свою очередь из уравнений (10), (11), (9), (12) следует:

$$\frac{dm_o}{dt} = -\frac{dm_{op}}{dt}, \tag{19}$$

$$\frac{dm_w}{dt} = -\frac{dm_{wp}}{dt} + \frac{dm_{winj}}{dt} + k_e\left(p_A - p\right), \tag{20}$$

$$\frac{dm_g}{dt} = -\frac{dm_{gp}}{dt} - m_{oi}\rho_{g0}\frac{dR_s}{dt} + \rho_{g0}\left(m_{op}\frac{dR_s}{dt} + R_s\frac{dm_{op}}{dt}\right). \tag{21}$$

С учётом (19), (20), (21) уравнение (7) примет вид

$$\frac{dp}{dt} = \frac{1}{H(p, p_A, t)}\left[G(p,t) + \frac{B_w}{\rho_{w0}}k_e(p_A - p)\right], \tag{22}$$

где

$$H(p, p_A, t) = \Bigg[(c_f + c_o)\frac{\left(m_{oi} - m_{op}(t)\right)\left(1 + R_s(p)\rho_{g0}\right)}{\rho_{o0}}B_o + \left(c_f + c_w\right)\frac{m_{wi} - m_{wp}(t) + m_{winj}(t)\left(1 - \varphi(p_A)\right)}{\rho_{w0}}B_w +$$
$$\left(c_f + c_g\right)\frac{m_{gi} + m_{oi}R_{si}\rho_{g0} - m_{gp}(t) + \left(m_{oi} - m_{op}(t)\right)R_s(p)\rho_{g0}}{\rho_{g0}}B_g +$$
$$-\frac{\left(m_{oi} - m_{op}(t)\right)\rho_{g0}}{\rho_{o0}}B_o\frac{dR_s}{dp} + B_g\left(m_{oi} - m_{op}(t)\right)\frac{dR_s}{dp}\Bigg],$$

$$G(p,t) = -\left(1 + R_s(p)\rho_{g0}\right)\frac{B_o}{\rho_{o0}}\frac{m_{op}}{dt} + \frac{B_w}{\rho_{w0}}\left(-\frac{dm_{wp}}{dt} + \frac{dm_{winj}}{dt}\right) - \frac{B_g}{\rho_{g0}}\frac{m_{gp}}{dt} + B_gR_s\frac{m_{op}}{dt}.$$

Уравнения (16), (22) образуют замкнутую систему и решаются численно.

2 Численный метод

Для краткости обозначим правые части уравнений (16), (22) соответственно через $f_1(p, p_A)$ и $f_2(p, p_A, t)$. Для численного решения уравнений (16) и (22) используется разностная аппроксимация вида

$$\frac{p_A^{k+1} - p_A^k}{\tau} = f_1\left(\frac{p^k + p^{k+1}}{2}, \frac{p_A^k + p_A^{k+1}}{2}\right) \tag{23}$$

$$\frac{p^{k+1} - p^k}{\tau} = f_2\left(\frac{p^k + p^{k+1}}{2}, \frac{p_A^k + p_A^{k+1}}{2}, k\tau\right) \tag{24}$$

где p^k, p_A^k — значения сеточных функций на k-ом временном шаге. Данная разностная схема реализуется следующим образом. Предполагается, что $p|_{t=0} = p_A|_{t=0}$ является известной величиной. Пусть найдены

значения функций на k-ом временном шаге p^k, p_A^k. Значения функций на шаге $k+1$ находятся итеррационно. Значение $p_A^{k,j}$ на j-ой итерации определяется из уравнения

$$\frac{p_A^{k,j} - p_A^k}{\tau} = f_1\left(\frac{p^k + p^{k,j-1}}{2}, \frac{p_A^k + p_A^{k,j}}{2}\right) \quad j = 1, 2, 3, ... \tag{25}$$

при этом $p^{k,0} = p^k$. Значение $p^{k,j}$ находится из уравнения

$$\frac{p^{k,j} - p^k}{\tau} = f_2\left(\frac{p^k + p^{k,j}}{2}, \frac{p_A^k + p_A^{k,j}}{2}, k\tau\right) \quad j = 1, 2, 3, ... \tag{26}$$

Итерации (25), (26) выполняются либо наперёд заданное максимальное число раз, либо до выполнения условия

$$|p^{k,j} - p^{k,j+1}| \le \varepsilon.$$

Представленные выше уравнения вида

$$\frac{f - f^k}{\tau} = F\left(\frac{f^k + f}{2}\right)$$

решаются методом Ньютона:

$$f^{m+1} = f^m - \frac{\left(f^m - f^k\right) - \tau F\left(\dfrac{f^m + f^k}{2}\right)}{\left(1 - \dfrac{\tau}{2}F'\right)},$$

где значение производной F' находится численно. Заметим, что если связь между нефтяным и водоносным пластами отсутствует ($k_e = 0$), то при реализации (25) будет выполняться только одна итерация. Использование неявной аппроксимации уравнений (16), (22) обеспечивает устойчивость разностной схемы во всём физически корректном диапазоне параметров задачи (7)-(12).

3 Расчет газонефтяного фактора

Если предполагать, что из нефтяного пласта ведётся добыча только попутного нефтяного газа и прорывы природного газа из газовых шапок отсутствуют, то становится возможным моделирование динамики добычи ПНГ на основе уравнения (3). При этом возможны два сценария расчёта по модели материального баланса:

1. Исторические значения добычи газа m_{gp} известны и достоверны. Рассчёт динамики добычи газа на основе формулы (3) производится только для прогнозного периода.

2. Исторические значения добычи газа m_{gp} либо ставятся под сомнение, либо неизвестны вовсе. В таком случае исходная система уравнений (7)–(12) дополняется уравнением (3).

Поскольку нам требуется решить задачу (2), поэтому будем считать, что мы имеем дело со вторым сценарием. Уравнение (3) аппроксимируется следующим образом:

$$\frac{m_{gp}^{k+1} - m_{gp}^k}{\tau} = \frac{\rho_{g0}}{\rho_{o0}} \frac{m_{op}^{k+1} - m_{op}^k}{\tau} \left[R_s\left(\frac{p^k + p^{k+1}}{2}\right)\rho_{o0} + \frac{k_{rg} B_o \left(\dfrac{p^k + p^{k+1}}{2}\right)\mu_o}{k_{ro} B_g \left(\dfrac{p^k + p^{k+1}}{2}\right)\mu_g} \right] \tag{27}$$

Использование уравнения (3) в модели материального баланса возможно на основе введения дополнительного цикла внешних итераций с целью нахождения m_{gp}^{k+1} в соответствии с уравнением (27).

Результаты расчётов динамики пластового давления, а также динамики газонефтяного фактора, выполненные на основе изложенной модели материального баланса, см. в [2].

Заключение

Для моделирования динамики изменения газосодержания пластовой нефти в процессе разработки предлагается подход, основанный на моделировании расхода попутного нефтяного газа на установках подготовки нефти, снабжённых узлами учёта. Моделирование газонефтяного фактора для пластов, разрабатываемых со снижением пластового давления ниже давления насыщения, предлагается осуществлять на основе 0-мерных гидродинамических моделей (модель материального баланса). В работе представлен вывод исходных уравнений и предложен способ их численного решения.

Отметим, что в силу своей грубости в некоторых случаях модель материального баланса не сможет корректно описать динамику пластового давления и, как следствие, динамику газонефтяного фактора. Однако при решении задачи (2) использование более сложных моделей практически не представляется возможным.

Результаты моделирования расхода газа для реальных узлов учёта, представленные в работе [2], демонстрируют, что на основе моделей материального баланса во многих случаях можно получить результаты, качественно согласующиеся с фактическими замерами.

Список литературы

[1] Постановление Правительства Российской Федерации от 8 ноября 2012 № 1148 // Собрание законодательства РФ. 2012. №47. Ст. 6499.

[2] Кожин В.Н., Базовкин А.В., Середа И.А., Амиров А.А. Адаптивная методика прогноза добычи газа и газовых факторов в условиях недостаточности замеров // Нефтепромысловое дело. 2019, №3, С. 45-51.

[3] Dake L.P. The practice of reservoir engineering. Elsevier 1994.

[4] Sanni M. Petroleum Engineering: Principles, Calculations and Workflows. John Wiley & Sons 2018.

[5] Beggs H.D. Production optimization. Using NODAL analysis. OGCI and Petroskills Publications, Tulsa 1991

[6] http://rfdyn.ru/technology/

[7] https://sis.slb.ru/products/eclipse/

[8] http://roxar.ru/software/tempest/

[9] https://www.cmgl.ca/stars/

[10] Бадыков И.Х., Байков В.А., Борщук О.С. Программный комплекс “РН-КИМ” как инструмент гидродинамического моделирования залежей углеводородов // Недропользование XXI век. 2015. № 4 (54). С. 96-103.

[11] Филиппов Д.Д., Васекин Б.В., Митрушкин Д.А. Гидродинамическое моделирование сложнопостроенных коллекторов на динамической адаптивной PEBI-сетке // PROНЕФТЬ. Профессионально о нефти. 2017. № 4. С. 48-53

[12] Мангазеев П.В. Практические рекомендации по решению производственных задач в нефтяной промышленности: учебное пособие. Томск: Центр подготовки и переподготовки специалистов нефтегазового дела Томского политехнического университета, 2013.

[13] Havlena, D., Odeh, A. S. The material balance as an equation of a straight line. // Journal of Petroleum Technology. 1963. V. 15, iss. 8. P. 896–900.

[14] Гиматудинов Ш.К. Физика нефтяного и газового пласта. М.: Недра, 1971.

Базовкин Андрей Владимирович — к.ф.-м.н., гл. спец. ООО "СамараНИПИнефть";
e-mail: BazovkinAV@samnipineft.ru.
Дата поступления — 20 мая 2019 г.

УЧЕТ ИНФОРМАЦИИ В ДИСКРЕТНО-НЕПРЕРЫВНЫХ СИСТЕМАХ УРАВНЕНИЙ ТИПА УРЫСОНА

В. А. Белозуб[1], М. Г. Козлова[1]

[1] *ФГАОУ ВО «Крымский федеральный университет им. В. И. Вернадского», 295007, Симферополь*

УДК 519.7
DOI: 10.24411/9999-016A-2019-10008

Рассматриваются системы нелинейных дискретно-непрерывных уравнений типа Урысона. Дискретность соответствует дискретным моментам времени измерения и дискретности точек наблюдения. Интегральный оператор содержит множитель, соответствующий характеру модели отражения от поверхности и дельтообразное ядро, зависящее от времени отражения импульса от исследуемой поверхности (двойное расстояние от точки наблюдения до поверхности, деленное на скорость). Решение таких систем является некорректной задачей. Регуляризирующие алгоритмы основаны на учете информации о решении, модели дистанционного зондирования, параметрах ядра. Учитываются: гладкость искомого решения, его положительность, монотонность, наличие точек экстремума, асимптотические свойства интегрального оператора типа Урысона в зависимости от параметров дельтаобразного ядра. Линеаризация и замена уравнений близкими, их дискретизация, позволяет строить приемлемые по сложности итерационные алгоритмы. Проведены численные эксперименты, подтверждающие теоретические выкладки.

Ключевые слова: нелинейные дискретно-непрерывные уравнения типа Урысона, задачи восстановления изображений, регуляризирующие алгоритмы.

Введение

Разработка устойчивых, регуляризирующих алгоритмов связана с адекватным использованием априорной и другой информации о решении, модели, способах получения данных косвенных измерений. В работе рассматриваются модели в виде дискретно-непрерывных аналогов нелинейных интегральных уравнений типа свертки первого рода. Разнообразие регуляризирующих алгоритмов определяется наличием доступной информации и возможностью ее использования. Тем самым формируется интеллектуальная система, базовыми элементами которой являются ультрасистемы (по терминологии А. В. Чечкина [1]), преобразователи семантической информации и ультраоператоры.

Дискретно-непрерывные системы уравнений моделируют процесс сканирования поверхности антенными устройствами, характер сигнала и его отражение от сканируемой поверхности. Такие нелинейные системы алгебраических уравнений являются аналогами нелинейных уравнений типа Урысона 1-го рода [2, 8].

Синтез дискретных и дискретно-непрерывных моделей — аналогов непрерывных — выявил ряд проблем. Процесс дискретизации (и квантования) существенно зависит от зондирующего сигнала, его дельтаобразности и искомого решения (поверхности и характера отражения от поверхности). Особенность состоит в применении асимптотических формул для интеграла с большими параметрами и учета известной информации о решении (априорной, прецедентной и др.). Таким образом, это может быть композиция дискретных адаптивных моделей, допускающих реоптимизацию по прецедентам. Дискретная модель должна обеспечивать проведение вычислительных экспериментов по решению прямой задачи, а также обратной, которая является некорректной. При выборе дискретно-непрерывных моделей предпочтение отдается простейшим прямой и обратной задачам.

ISBN 978-5-901548-42-4

1 Дискретно-непрерывные модели уравнений типа Урысона

К дискретным, дискретно-непрерывным, интегральным нелинейным уравнениям типа Урысона приводит широкий класс задач моделирования косвенных измерений, дистанционного зондирования, сейсмо- и геологоразведки и др. В качестве примера рассмотрим одномерный случай определения формы геологической аномалии и ее характеристик (плотность, проводимость и др.) по результатам измерений на поверхности Земли.

Требуется определить плотность аномальной области $\rho(x), a \leqslant x \leqslant b$, находящейся на глубине H, ограниченной поверхностью $h(x), a \leqslant x \leqslant b$ по изменениям вертикальной составляющей $u(x)$ — силы тяжести на поверхности в точке (x, H) (см. рис. 1).

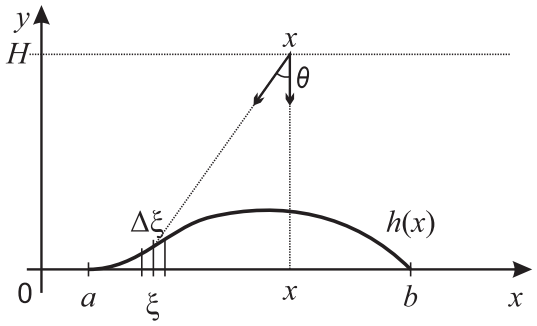


Рис. 1: Модель измерения характеристик аномалии

Масса элемента «объема» в точке ξ будет равна $\rho(\xi)\Delta\xi h(\xi)$ или $\rho(\xi)\Delta\xi$, если не зависит от $h(\xi)$. Расстояние от точки $(\xi, h(\xi))$ до точки измерения (x, H) составляет

$$\left[(x-\xi)^2 + (H - h(\xi))^2\right]^{\frac{1}{2}}$$

Изменение силы тяжести описывается законом тяжести Ньютона $f = \gamma\frac{m}{r^2}$ с гравитационной постоянной γ. Для вертикальной составляющей

$$\Delta f_v(x) = \Delta u(x) = \frac{\gamma\rho(\xi)\Delta\xi h(\xi) cos\theta}{(x-\xi)^2 + (H-h(\xi))^2} = \gamma\frac{h(\xi)\rho(\xi)\Delta\xi[H-h(\xi)]}{[(x-\xi)^2 + (H-h(\xi))^2]^{\frac{3}{2}}}.$$

Откуда получаем следующее интегральное уравнение (если x не дискретно) типа Урысона

$$\gamma\int_a^b \frac{\rho(\xi)h(\xi)[H-h(\xi)]d\xi}{[(x-\xi)^2 + (H-h(\xi))^2]^{\frac{3}{2}}} = u(x), \quad a \leqslant x \leqslant b. \tag{1}$$

В предположении, что форма аномалии, находящейся на глубине H, нас не интересует, а масса элемента объема в $\Delta\xi$ равна $\rho(\xi)\Delta\xi$, получим уравнение

$$\gamma H\int_a^b \frac{\rho(\xi)d\xi}{[(x-\xi)^2 + H^2]^{\frac{3}{2}}} = u(x), \quad a \leqslant x \leqslant b. \tag{2}$$

Здесь неизвестной функцией является только плотность $\rho(\xi)$ (и может быть H, a, b), в отличие от уравнения (1), в котором неизвестны две функции $h(\xi)$ и $\rho(\xi)$. Вариант уравнения (1), когда масса элемента $\Delta\xi$, сосредоточенного возле точки ξ, определяется как $\rho(\xi)\Delta\xi$ и не зависит от $h(\xi)$ запишем в виде:

$$Af \equiv \gamma\int_a^b \frac{\rho(\xi)[H-h(\xi)]d\xi}{[(x-\xi)^2 + (H-h(\xi))^2]^{\frac{3}{2}}} = u(x), \quad a \leqslant x \leqslant b, \quad f = (\rho, h). \tag{3}$$

Задача решения уравнений первого рода (1)–(3) является некорректно поставленной [3]. Если через $k(x, y)$ обозначить ядро интегрального оператора A (3)

$$k(x, y) = \frac{\gamma y}{[x^2 + y^2]^{\frac{3}{2}}}, \tag{4}$$

тогда уравнение (3) $Af = u$ можно записать в виде

$$Af \equiv \int_a^b k(x-\xi, H-h(\xi))\rho(\xi)d\xi = u(x), \quad a \leqslant x \leqslant b. \tag{5}$$

Если измерения проводятся в дискретных точках $x_i, i = \overline{1,m}$, то получаем дискретно-непрерывные системы уравнений типа свертки

$$(Af)_i = \int_a^b k(x_i-\xi, H-h(\xi))\rho(\xi)d\xi = u(x_i) \equiv u_i, \quad a \leqslant x_i \leqslant b, \quad i = \overline{1,m}. \tag{6}$$

Дискретный вариант уравнения (6) представляет собой нелинейную алгебраическую систему уравнений:

$$\sum_{j=1}^{n} A_j k_{i-j}(h_j)\rho_j = u_i, \quad i = \overline{1,m}, \tag{7}$$

$$k_{i-j}(h_j) = k(x_i-\xi_j, H-h(\xi_j)), \quad \rho_j = \rho(\xi_j), \quad u_i = u_i(x).$$

На примере уравнения (5) можно показать, что невозможно восстановить две функции $h(\xi), \rho(\xi)$ из одного уравнения только по информации о правой части $u(x)$. Поэтому необходимо иметь систему таких уравнений или восстанавливать только некоторые характеристики искомых функций (точки экстремума, участки монотонности и др.), используя априорную, прецедентную или иного рода информацию о решении $f = (h, \rho)$ и специфике модели. Можно поставить задачу нахождения одной из функций h или ρ в предположении, что другая задана.

Пусть заданы приближения функции $h(\xi)$ в виде последовательности функций $h_n(\xi), n = 0, 1, 2, ...,$, где $h_0(\xi) = 0$, тогда $n-$е приближение для $\rho(\xi)$ будем находить в результате решения интегрального уравнения типа (5) (или системы дискретно-непрерывных уравнений (6)):

$$\int_a^b k(x-\xi, H-h_n(\xi))\rho(\xi)d\xi = u(x), \quad a \leqslant x \leqslant b. \tag{8}$$

Приближенное решение $\rho(\xi)$ используется для построения нового приближения для $h(\xi)$:

$$\int_a^b k(x-\xi, H-h(\xi))\rho_n(\xi)d\xi = u(x), \quad a \leqslant x \leqslant b.$$

В результате получаем решения $f_n = (h_n(\xi), \rho_n(\xi))$.

Естественно, решение находится в результате применения регуляризирующих алгоритмов. Для $h_n(\xi)$ обоснование для дискретного аналога следует из работ Васина В.В. (см. для алгоритма Левенберга-Марквардта работы [5, 6]).

Другим примером является задача восстановления поверхности по результатам измерения с помощью антенных устройств времени отражения сигнала (импульса) от поверхности $h(\xi)$, направленного вертикально вниз:

$$\tau \equiv z(\xi) = \frac{2}{c}\left[(x-\xi)^2 + (H-h(\xi))^2\right]^{\frac{1}{2}},$$

где c — скорость распространения импульса, τ — удвоенное время (до поверхности и обратно).

В связи с тем, что интегрируется отраженный сигнал с весовой функцией $k(x, \xi)$ (ядро, диаграмма направленности, аппаратная функция измерительного прибора), а импульс имеет определенную форму $n(t)$, получаем уравнение

$$Az \equiv \int_a^b m(s)k(x-s)n(t-z(s))ds = u(x,t), \quad a \leqslant x \leqslant b, \quad c \leqslant t \leqslant d, \tag{9}$$

где $m(s)$ характеризует отражение от поверхности, $n(t)$ — дельтаобразная функция. Функция m может зависеть от z и z'.

Таким образом, в качестве модельных будем рассматривать дискретно-непрерывные нелинейные операторы следующего вида

$$(Az)_i \equiv \int_a^b m(s)n(t_i - z(s))ds, \quad c \leqslant t_i \leqslant d, \quad i \in M \subset \mathbb{N}$$

$$(Az)_i \equiv \int_a^b k(x-s)n(t_i - z(s))ds, \quad c \leqslant t_i \leqslant d, a \leqslant x \leqslant b$$

$$(Az)_i \equiv \int_a^b m(s)k(x-s)n(t_i - z(s))ds, \quad c \leqslant t_i \leqslant d, a \leqslant x \leqslant b.$$

В линеаризованных уравнениях используются производные по z, которые имеют вид

$$A'(z)_i h \equiv -\int_a^b m(s)n'(t_i - z(s))h(s)ds, \quad c \leqslant t_i \leqslant d,$$

$$A'(z)_i h \equiv -\int_a^b k(x-s)n'(t_i - z(s))h(s)ds, \quad c \leqslant t_i \leqslant d,$$

$$A'(z)_i h \equiv -\int_a^b m(s)k(x-s)n'(t_i - z(s))h(s)ds, \quad c \leqslant t_i \leqslant d.$$

Для рассматриваемого класса моделей можно предъявить иерархию математических моделей и множество сценариев измерения, сбора и использования информации для корректного восстановления искомых величин.

Например, в задаче восстановления характерных точек решения в пространстве гладких функций для системы интегро-дифференциальных уравнений первого рода по двум измерениям

$$\int_a^b \varphi_k(x, h(x), h'(x))exp\left(-\lambda\left(t - \frac{2}{c}R_k(h)\right)^2\right)dx = g_k(t), \quad k = 1, 2, \tag{10}$$

$$R_k(h) = \sqrt{(H - h(x))^2 + (x - a_k)^2}, \tag{11}$$

где $\varphi_k > 0$ — достаточное число раз непрерывно дифференцируемая, $\lambda >> 1$ — фиксированный параметр, c, H, a_1, a_2 — фиксированные числа. При этом $H >> 1, R_k(h) \neq 0$. Предполагается существование решения $h(x)$ системы (10)–(11) при заданных $g_k(t), k = 1, 2$.

Для выяснения поведения интегрального оператора (10) в зависимости от большого параметра $\lambda >> 1$ используем метод Лапласа [4]

$$q(t) = \int_a^b f(x)exp(-\lambda(t - z(x)))dx.$$

Предполагается, что для фиксированного t существуют корни $x_1, x_2, ..., x_p$ уравнения $z(s) = t$ и все они являются простыми, причем ни одно из этих точек не является граничной точкой $[a, b]$. Тогда имеет место следующее асимптотическое представление

$$q(t) = \lambda^{-\frac{1}{2}}\sqrt{\pi}\left(\sum_{k=1}^{p}\frac{f(x_k)}{z'(x_k)} + o(\lambda^{-1})\right).$$

Если для выбранного t существует единственный кратный корень $\widehat{x}$, уравнение $z(x) = t$ кратности 1, так что $z'(\widehat{x}) = 0, z''(x) \neq 0$, тогда

$$q(t) = \lambda^{-\frac{1}{4}}\left(\frac{f(\widehat{x})}{|z'(\widehat{x})|}\frac{1}{2}\Gamma\left(\frac{1}{4}\right) + o(\lambda^{-\frac{1}{2}})\right),$$

где $\Gamma(x)$ — гамма-функция.

Если $f(x)$ — положительная, непрерывная медленно меняющаяся функция, то из оценок следует, что функция $q(t)$ принимает максимальное значение в точках, лежащих в окрестности порядка $\lambda^{-\frac{1}{4}}$ точек $\widehat{t}$, для которых существует стационарная точка $\widehat{x}$ такая, что $z(\widehat{x}) = \widehat{t}$. Таким образом, экстремальные точки правой части определяют экстремальные точки решения интегрального уравнения.

Перейдем к рассмотрению системы (10). Согласно анализу поведения интегралов, будем полагать, что $t \in \Omega(t)$, т.е. промежуток изменения t, на котором рассматривается система, содержит промежуток $[m_k, M_k]$ где $m_k = min_{[a,b]} \frac{2}{c} R_k(h)$, $M_k = max_{[a,b]} \frac{2}{c} R_k(h), k = 1, 2$.

Заметим, что при сформулированных условиях относительно H, стационарные точки $h(x)$ близки к стационарным точкам функции $R_k(h)$. Пусть $h_1, h_2, ..., h_m$ — упорядоченное по убыванию множество значений функции h в ее стационарных точках $x_1, x_2, ..., x_m$. Предположим, что расстояние a_1, a_2 от $[a, b]$ имеет порядок $H >> 1$, тогда последовательность

$$R_k(h_i) = \frac{2}{c}\sqrt{(M - h_i(x))^2 + (x_i - a_k)^2}, \tag{12}$$

является упорядоченной по возрастанию. Характерными точками решения будем считать стационарные точки, которые определяются из (12) по измеренным функциям $g_k(t)$.

Алгоритм предполагает упорядочение по возрастанию точек максимума $g_k(t) : t_1^k, t_2^k, ..., t_p^k, k = 1, 2$ решение системы $R_k(h_i) = t_i^k$ относительно $(h_i, x_i), h(x_i) = h_i$. В использовании схемы алгоритма без регуляризации наблюдается практическая неустойчивость.

Информация о монотонности искомого решения $z(s)$ позволяет всести задачу решения нелинейного уравнения Урысона 1 рода к линейному уравнению типа свертки. Соответствующие результаты можно найти в работе [8, 9]. Учет разнообразной информации уже требует разработки итнтеллектуализированной системы обработки данных [7].

Заметим, что данный подход приводит к различным сценариям нахождения искомых величин по результатам косвенных измерений с учетом разнообразной информации. Здесь естественным является подход, основанный на совместной работе нескольких программных агентов, передающих информацию друг другу.

Для обоснования восстановления решений исследуемых уравнений с помощью, например, двух программных агентов, рассмотрим простую схему решения одного уравнения $Az = u$ на основе решения близкого ему уравнения $\widetilde{A}\widetilde{z} = \widetilde{u}$. Задача сводится к двум близким экстремальным задачам

$$M^\alpha[z] = \alpha||z||^2_{W_2^1} + ||Az - u||^2_{L_2}$$

$$\widetilde{M}^\alpha[\widetilde{z}] = \widetilde{\alpha}||\widetilde{z}||^2_{W_2^1} + ||\widetilde{A}\widetilde{z} - \widetilde{u}||^2_{L_2}$$

Им эквивалентны уравнения

$$Bz \equiv \alpha(-z'' + z) + A^*Az = A^*u \equiv g,$$
$$\widetilde{B}\widetilde{z} \equiv \widetilde{\alpha}(-\widetilde{z}'' + \widetilde{z}) + \widetilde{A}^*\widetilde{A}\widetilde{z} = \widetilde{A}^*\widetilde{u} \equiv \widetilde{g}$$

или алгебраические уравнения

$$\mathbb{K}_\alpha Z = G, \quad \widetilde{\mathbb{K}}_\alpha \widetilde{Z} = \widetilde{G}$$

с параметрами регуляризации $\alpha, \widetilde{\alpha}$. Для уравнений вида $Bz = g$ и $\widetilde{B}\widetilde{z} = \widetilde{g}$ справедлива

Теорема. *Пусть выполнены следующие условия:*

1. $\exists Y_0 \subset Y, \quad g \in Y, \quad \widetilde{g} \in Y, \quad mo\ (g - \widetilde{g}) \in Y_0$;

2. Уравнение $\widetilde{B}z_0 = \widetilde{g}$ *имеет единственное решение;*

3. $B - \widetilde{B} : X \to Y_0$;

4. $\exists \widetilde{B}^{-1} : Y_0 \to X_0$;

5. Выполняется условие $||\widetilde{B}^{-1}(B - \widetilde{B})|| < 1$.

Тогда итерационный процесс $z_{n+1} = z_n\widetilde{B}^{-1}[Bz_n - g]$ *— сходится в* X_0. *Все элементы* z_n *обладают свойством* $z^* - z_n \subset X_0$. *Справедлива формула для погрешности:*

$$||z_n - z^*|| \leqslant \frac{||\widetilde{B}^{-1}(\widetilde{B} - B)||^n}{1 - ||\widetilde{B}^{-1}(\widetilde{B} - B)||}||\widetilde{B}^{-1}g||_{X_0}.$$

Здесь $X = Y = L_2, X_0 = W_2^1$.

Для метода регуляризации Лаврентьева $F(t) \equiv Az - u + \alpha(z - z_m)$ итерационная процедура построения z_{n+1} приближения $z_{n+1} = z_n + h$ приводит к алгоритму нахождения h из линейного уравнения

$$[\alpha I + (A'z_n)]h - F(z_n) = 0$$

или

$$z_{n+1} = z_n - \gamma[\alpha I + (A'z_n)]^{-1}[Az_n - u + \alpha(z_n - z_m)].$$

В качестве близких уравнений следует рассматривать уравнения

$$\mathbb{K}h \equiv [\alpha I + (A'z_n)]h = Az_n - u + \alpha(z_n - z_m) \equiv g,$$

$$\widetilde{\mathbb{K}}\widetilde{h} \equiv [\widetilde{\alpha} I + (\widetilde{A}'\widetilde{z}_n)]\widetilde{h} = \widetilde{A}\widetilde{z}_n - \widetilde{u} + \widetilde{\alpha}(\widetilde{z}_n - \widetilde{z}_m) \equiv \widetilde{g},$$

обеспечив $||\widetilde{\mathbb{K}}^{-1}(\mathbb{K} - \widetilde{\mathbb{K}})|| < 1$, можем найти решение через решение близкого уравнения, более простого по своей структуре, или решение, которое уже найдено для различных уровней погрешности (прецедентная информация).

Для модифицированного варианта метода Левенберга-Марквардта [5, 6] решения нелинейных уравнений в частном случае имеет вид

$$z_{n+1} = z_n - \gamma[\widetilde{A}'(z_n)^*\widetilde{A}'(z_n) + \alpha]^{-1}[\widetilde{A}'(z_n)^*(Az_n - u_\delta) + \alpha(z_n - \widetilde{z}_m)].$$

Здесь $\widetilde{z}_m$ — начальное приближение для искомого решения.

Заключение

В работе получены следующие результаты: приведен набор модельных дискретно-непрерывных моделей, применяемых в прикладных задачах восстановления изображений. Получены регуляризирующие алгоритмы, учитывающие различную априорную информацию. Показана возможность использования асимптотических методов и многоагентных подходов.

Список литературы

[1] Чечкин А. В. Математическая информатика. — М.: Наука, 1991. — 416 с.

[2] Lukyanenko V. A., Kozlova M. G., Hazova U. A. Some Tasks for Integral Equations of Urison's Type // Proceedings of the International Conference «Integral Equations – 2010», 25-27 August 2010. — Lviv, Ukraine: PAIS, 2010. — P. 80-84.

[3] Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г. Регуляризирующие алгоритмы и априорная информация. — М.: Наука, 1983. — 200 с.

[4] Федорюк М. В. Метод перевала. — М.: Наука, 1974. — 350 с.

[5] Васин В. В., Пересторонина Г. Я. Метод Левенберга–Марквардта и его модифицированные варианты для решения нелинейных уравнений с приложением к обратной задаче гравиметрии// Тр. ИММ УрО РАН, 17, № 2. — 2011. — С. 53–61.

[6] Васин В. В. Метод Левенберга–Марквардта для аппроксимации решений нерегулярных операторных уравнений// Автоматика и телемеханика. — 2012. — Т. 73, выпуск 3. — С. 28–38.

[7] Белозуб В. А., Козлова М. Г. Интеллектуализация обработки данных в некорректных задачах дискретной оптимизации // Международная конференция «XXVII Крымская Осенняя Математическая Школа-симпозиум по спектральным и эволюционным задачам» (КРОМШ-2016): сборник тезисов. — Симферополь: ООО ФОРМА, 2016. — С. 132.

[8] Белозуб В. А., Козлова М. Г., Лукьяненко В. А. Уравнения типа Урысона в задачах восстановления точек поверхности // Соболевские чтения. Международная школа-конференция: Тез. докл. / Под ред. Г. В. Демиденко. — Новосибирск: Изд-во Института математики, 2017. — С. 35-36.

[9] Белозуб В. А., Козлова М. Г. Некоторые алгоритмы восстановления решений типа Урысона // Материалы IX международной конференции «Современные методы и проблемы теории операторов и гармонического анализа и их приложения IX» (г. Ростов-на-Дону, 22—25 апреля 2019 г.) / Под ред. А. В. Гиля; [Электронный ресурс]. — Ростов н/Д: Издательство Ростовского отделения Российской инженерной академии, 2019. — С. 126—127.

Белозуб Владимир Антонович — аспирант кафедры информатики,
факультет математики и информатики
Таврической академии ФГАОУ ВО «КФУ им. В. И. Вернадского»;
e-mail: disstroier@mail.ru;
Козлова Маргарита Геннадьевна — к.ф.-м.н., доцент кафедры информатики,
факультет математики и информатики
Таврической академии ФГАОУ ВО «КФУ им. В. И. Вернадского»;
e-mail: art-inf@mail.ru.
Дата поступления — 1 июня 2019 г.

ОБ АЛГОРИТМАХ ОТЫСКАНИЯ НАИМЕНЬШЕГО РАССТОЯНИЯ МЕЖДУ СИСТЕМАМИ ТОЧЕК В ПРОСТРАНСТВЕ

И. А. Блатов[1], Е. В. Китаева[2]

[1] *Поволжский государственный университет телекоммуникация и информатики, 443010, Самара*

[2] *Самарский национальный исследовательский университет имени академика С.П. Королева, 443086, Самара*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10009

Рассматривается задача определения минимума сумм квадратов расстояний между двумя системами точек в пространстве. Необходимость решения такой задачи возникает в компьютерной химии и при распознавании образов. Грубый алгоритм имеет факториальную сложность. Предлагается алгоритм, позволяющий найти решение с заданной точностью ε за $O(n^3/\varepsilon^{3/2})$ арифметических операций.

Ключевые слова: распознавание образов, совмещение объектов, венгерский алгоритм, алгоритм Кабша.

Введение

Пусть в пространстве заданы две системы точек M и N, каждая из которых содержит n точек. Рассматривается задача отыскания такого положения этих систем в пространстве и паросочетания их вершин, для которых сумма квадратов попарных расстояний $\rho(M,N)$ между точками M и N была бы минимальной. В случае фиксированного паросочетания данная задача решается за линейное ($O(n)$) количество операций алгоритмом Кабша [1, 2]. Грубый поиск оптимального паросочетания требует перебора $n!$ различных положений. В случае фиксированного положения в пространстве систем и известен венгерский алгоритм [3] отыскания оптимального паросочетания за $O(n^3)$ операций. Однако неизвестен [4] алгоритм общего назначения, позволяющий за полиномиальное время найти оптимальное паросочетание и соответствующее $\rho(M,N)$. Различные варианты алгоритмов, сочетающих метод Кабша и венгерский метод, рассматривались в [5–7] (см. также библиографию в [6, 7]). Однако это либо эвристические алгоритмы, не гарантирующие отыскание глобального минимума, либо алгоритмы, использующие взаимозависимость и специальные свойства сравниваемых структур, либо экспоненциально или факториально сложные алгоритмы. Авторы [7]) утверждают, что предложенный ими алгоритм Go-PERMDIST имеет полиномиальную сложность, но не приводят точных оценок вычислительной сложности. В настоящей статье для решения данной задачи рассматриваются три алгоритма общего назначения, основанные на алгоритмах Кабша и венгерском алгоритме, позволяющие найти решение поставленной задачи с заданной точностью $O(n^3/\varepsilon^{3/2})$ за арифметических операций. Проведен сравнительный экспериментальный анализ этих алгоритмов. Первый (грубый) алгоритм представляет собой сочетание венгерского алгоритма с перебором по узлам ε-сети в пространстве параметров изометрий в R^3. Второй алгоритм (комбинированный) отличается от первого последовательным выполнением венгерского алгоритма и алгоритма Кабша вместо одного лишь венгерского алгоритма. Третий алгоритм (итерационный) сочетает в себе венгерский алгоритм, алгоритм Кабша и итерационный метод покоординатного спуска в пространстве параметров изометрий в R^3. Проведен сравнительный экспериментальный анализ этих алгоритмов. Показана высокая эффективность второго алгоритма по сравнению с первым и третьим.

Расчет проводился на семействах атомных решеток из баз данных МНИЦТМ (Самара).

ISBN 978-5-901548-42-4

1 Постановка задачи и алгоритмы

Пусть $M = \{\vec{M}_i(x_i, y_i, z_i), 1 \le i \le n\}$, $N = \{\vec{N}_i(x_i, y_i, z_i), 1 \le i \le n\}$ — две совокупности точек (радиус-векторов) в R^3,заданных координатами относительно прямоугольной декартовой системы координат $OXYZ$. Через I обозначим единичную матрицу третьего порядка. Если $w = (w_1, w_2, w_3)$ вектор в R^3, то $\| w \| = \sqrt{w_1^2 + w_2^2 + w_3^2}$ обозначает его евклидову норму. через $C, C_1, C_2 \cdots$ будем обозначать положительные константы, не зависящие от n. Если это не вызывает недоразумений, то одним и тем же символом C могут обозначаться разные константы. Для некоторой величины f зависящей от n, запись $f = O(n)$ означает, что $f \le Cn$.

Рассмотрим задачу отыскания вектора $\nu = (\nu_1, \nu_2, \nu_3)$, ортогональной матрицы третьего порядка $U = \{u_{ij}\}$ $(UU^T = I)$ и перестановки $P = (p_1, p_2, \cdots, p_n)$ первых n натуральных чисел, для которых величина $E = \frac{1}{n}\sum_{i=1}^{n} \| UM_{p_i} - N_i - \nu \|^2$ минимальна. Систему M будем называть подвижной системой, а N — неподвижной системой. Сделаем следующие предположения.

A1. Центры тяжести $\frac{1}{n}\sum_{i=1}^{n} M_i$ и $\frac{1}{n}\sum_{i=1}^{n} N_i$ систем M и N совмещены сдвигом в начале координат $OXYZ$.

В предположении **A1** рассматриваемая задача имеет вид

$$E = \frac{1}{n}\sum_{i=1}^{n} \| UM_{p_i} - N_i \|^2 \to \min_{P,U} \tag{1}$$

при условиях

$$UU^T = I, \quad P = (p_1, \cdots, p_n), \quad p_i \in [1, n], \quad p_i \ne p_j \ \text{ при } \ i \ne j \tag{2}$$

A2. Оси тензоров инерции систем систем M и N совмещены ортогональным преобразованием с осями декартовой системы $OXYZ$.

Замечание 1. *В случае ограничения $p_i = i$, $i = 1, \cdots, n$, задача (1)–(2) решается алгоритмом Кабша [1,2], результат действия которого обозначим $AK(M, N)$. Положение UM подвижной системы, соответствующее результату действия алгоритма Кабша, обозначим $argminAK(M, N)$.*

Замечание 2. *В случае, когда положение подвижной системы в пространстве фиксировано $U = I$ задача (1)–(2) решается венгерским алгоритмом [3], результат действия которого обозначим $AV(M, N)$. Саму перенумерованную венгерским алгоритмом систему обозначим $argminAV(M, N)$.*

Перейдем к определению алгоритмов. В соответствии с [8] определим матрицу U в виде

$$U(\phi, \psi, \theta) = U_i(\phi, \psi, \theta) = P_i A(\phi, \psi, \theta), \quad i = 1, 2, \tag{3}$$

где

$$A = \begin{pmatrix} \cos\theta + (1-\cos\theta)x^2 & (1-\cos\theta)xy - z\sin\theta & (1-\cos\theta)xz + y\sin\theta \\ (1-\cos\theta)yx + z\sin\theta & \cos\theta + (1-\cos\theta)y^2 & (1-\cos\theta)yz - x\sin\theta \\ (1-\cos\theta)zx - y\sin\theta & (1-\cos\theta)zy + x\sin\theta & \cos\theta + (1-\cos\theta)z^2 \end{pmatrix}, \tag{4}$$

$x = \cos\phi\cos\psi$, $y = \sin\phi\cos\psi$, $z = \sin\psi$, $w = (x, y, z)$ — направляющий вектор оси поворота. Здесь $\phi \in [-\pi/2, \pi/2]$, $\psi \in [-\pi/2, \pi/2]$ — углы, определяющие направляющий вектор w прямой l, проходящей через начало координат, $\theta \in [0, 2\pi]$ — угол поворота точки системы M вокруг прямой l против часовой стрелки, $P_1 = I$, P_2 — матрица отражения [8] относительно прямой l.

Зафиксируем натуральное m. Пусть $h = \pi/m$.

Алгоритм 1(примитивный)

A1.1. $E := AV(M, N)$;

A1.2. Для $i := 1 \cdots 2$ выполнить

Для $j := 0 \cdots m$ выполнить

Для $k := 0 \cdots m$ выполнить

Для $l := 0 \cdots m$ выполнить

начало

$\phi := -\frac{\pi}{2} + jh$, $\psi := -\frac{\pi}{2} + kh$, $\theta := 2lh$;

$E1 := AV(U_i(\phi, \psi, \theta)M, N)$;

если $(E1 < E)$ то $(E1 := E)$;

конец.

Алгоритм 2(комбинированный)

A1.1. $E := AK(argminAV(M,N),N)$;

A1.2. Для $i := 1 \cdots 2$ выполнить
Для $j := 0 \cdots m$ выполнить
Для $k := 0 \cdots m$ выполнить
Для $l := 0 \cdots m$ выполнить
начало
$\phi := -\frac{\pi}{2} + jh$, $\psi := -\frac{\pi}{2} + kh$, $\theta := 2lh$;
$E1 := AK(argminAV(U_i(\phi,\psi,\theta)M,N)$;
если $(E1 < E)$ то $(E1 := E)$;
конец.

Третий алгоритм представляет собой комбинацию алгоритма 2 и итерационного метода покоординатного спуска. Определим сначала алгоритм покоординатного спуска. Обозначим

$$E_1(\phi,\psi,i) = \min_{\theta\in[0,2\pi]} AV(U_i(\phi,\psi,\theta)M,N), E_2(\phi,\theta,i) = \min_{\psi\in[-\pi/2,\pi/2]} AV(U_i(\phi,\psi,\theta)M,N),$$

$$E_3(\psi,\theta,i) = \min_{\phi\in[-\pi/2,\pi/2]} AV(U_i(\phi,\psi,\theta)M,N).$$

Через $\theta = argminE_1(\phi,\psi,i)$, $\psi = argminE_2(\phi,\theta,i)$, $\phi = argminE_3(\psi,\theta,i)$, обозначим соответствующие значения аргументов, при которых достигается минимум. Зафиксируем $\varepsilon \in (0,1]$, $\phi \in [-\pi/2,\pi/2]$, $\psi \in [-\pi/2,\pi/2]$. Рассмотрим следующий алгоритм.

начало
$\theta := 0$;
$F := AV(U_i(\phi,\psi,\theta)M,N)$;
1: Для $i := 1 \cdots 2$ выполнить
начало
$F_1(i) := E_1(\phi,\psi,i)$; $\theta_1 := argminE_1(\phi,\psi,i)$; $F_2(i) := E_2(\phi,\theta_1,i)$; $\psi_1 := argminE_2(\phi,\theta_1,i)$;
$F_3(i) := E_3(\phi_1,\theta_1,i)$; $\theta_1 := argminE_3(\phi,\theta_1,i)$;
конец
$F_4 := \min\limits_{i\in[1,2]} F_3(i)$;
если $|F_4 - F| > \varepsilon$ то ($\phi := \phi_1$, $\psi := \psi_1$, $\theta := \theta_1$, $F := F_4$; идти на 1);
$F := F_4$;
конец.

Результат действия данного алгоритма (число F) обозначим $F := DCM(\phi,\psi,\varepsilon)$.

Алгоритм 3 (итерационный)

A3.1. $E := AV(M,N)$;

A3.2. Для $j := 0 \cdots m$ выполнить
Для $k := 0 \cdots m$ выполнить
начало
$\phi := -\frac{\pi}{2} + jh$, $\psi := -\frac{\pi}{2} + kh$;
$E1 := DCM(\phi,\psi,\varepsilon)$; если $(E1 < E)$ то $(E1 := E)$;
конец.

2 Оценка скорости сходимости алгоритмов

В данном разделе мы докажем следующее утверждение.

Теорема. *Для отыскания значения E с точностью ε алгоритмами 1–3 достаточно совершить $O(n^3/\varepsilon^{3/2})$ арифметических операций.*

Доказательство проведем для алгоритма 1. Очевидно, что значение E для алгоритмов 2 и 3 не больше, чем для алгоритма 1, поэтому и для них теорема будет справедлива. Обозначим $AV(U_i(\phi,\psi,\theta)M,N) = f_i(\phi,\psi,\theta)$, $\quad i = 1,2$. Пусть минимум $E1$ функции $f_i(\phi,\psi,\theta)$ достигается при некотором паросочетании вершин и значениях параметров $\phi = \tilde{\phi}, \psi = \tilde{\psi}, \theta = \tilde{\theta}, i = \tilde{i}$.Обозначим через $\tilde{f}(\phi,\psi,\theta)$ функцию, значение

которой равно $AV(U_{\tilde{i}}(\phi,\psi,\theta)M,N)$ для этого фиксированного паросочетания P при всех ϕ, ψ, θ. Тогда, очевидно, $\min \tilde{f}(\phi,\psi,\theta) = \min\limits_{i=1,2} f_i(\phi,\psi,\theta)$ Но в силу своего определения (см. (3)–(4)) $\tilde{f}$ есть дважды дифференцируемая функция с ограниченными вторыми частными производными. Поэтому в точке минимума ее первый дифференциал $d\tilde{f}(\tilde{\phi},\tilde{\psi},\tilde{\theta}) = 0$. Тогда согласно формуле Тейлора

$$\tilde{f}(\phi,\psi,\theta) = E + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{\partial^2\tilde{f}}{\partial x^i \partial x^j}(\bar{\phi},\bar{\psi},\bar{\theta})(x_i-\tilde{x}_i)(x_j-\tilde{x}_j), \tag{5}$$

где обозначено $x_1=\phi$, $x_2=\psi$, $x_3=\theta$, $\bar{\phi}=\tilde{\phi}+\lambda(\phi-\tilde{\phi})$, $\bar{\psi}=\tilde{\psi}+\lambda(\psi-\tilde{\psi})$, $\bar{\theta}=\tilde{\theta}+\lambda(\theta-\tilde{\theta})$ для некоторого $\lambda\in[0,1]$. При этом

$$\frac{\partial^2\tilde{f}}{\partial x^i\partial x^j}(\phi,\psi,\theta) = \sum_{k=1}^{n}\frac{\partial^2\tilde{f}_k}{\partial x^i\partial x^j}(\phi,\psi,\theta), \tag{6}$$

где $\tilde{f}_k(\phi,\psi,\theta) = \frac{1}{n}\parallel U_{\tilde{i}}(\phi,\psi,\theta)M_{p_k} - N_k \parallel^2$. Поскольку $|\frac{\partial^2\tilde{f}_k}{\partial x^i\partial x^j}(\phi,\psi,\theta)| \le \frac{C}{n}$ то из (6) следует, что

$$|\frac{\partial^2\tilde{f}}{\partial x^i\partial x^j}(\bar{\phi},\bar{\psi},\bar{\theta})| \le C. \tag{7}$$

Поскольку в $O(1/m)$ — окрестности точки $(\bar{\phi},\bar{\psi},\bar{\theta})$ всегда найдется набор узлов (ϕ_j,ψ_k,θ_l), по которым идет перебор в алгоритме 1, то из (5)–(7) имеем $\min\limits_{j,k,l}\tilde{f}(\phi_j,\psi_k,\theta_l) - E \le \frac{C}{m^2}$. Учитывая, что для любых (ϕ,ψ,θ) будет $\min\limits_{i=1,2} f_i(\phi,\psi,\theta) \le \tilde{f}(\phi,\psi,\theta)$, получаем, что

$$\min_{i=1,2} f_i(\phi_j,\psi_k,\theta_l) - E \le \frac{C}{m^2}. \tag{8}$$

Поскольку каждое значение $f_i(\phi,\psi,\theta)$ находится за $O(n^3)$ арифметических операций, а нахождение минимума с точностью ε требует в силу (8) значения $m = O(\varepsilon^{-1/2})$, то общее число операций составит $O(n^3/\varepsilon^{3/2})$. Теорема доказана.

3 Результаты численных экспериментов

Рассматривались три семейства S_1, S_2, S_3 атомных решеток, интерпретируемых как совокупности точек в R^3. Первое семейство S_1 состоит из 19 структур, каждая из которых содержит 9 точек, второе S_2 из 12 структур, содержащих 18 точек каждая, третье семейство S_3 из 7 структур, содержащих 57 точек каждая. Вычислительный эксперимент для каждой из структур состоял в последовательном отыскании величины (1) для всевозможных пар структур семейств. Приближенное значение этой величины, найденное одним из алгоритмов 1, 2, 3, обозначим E_{appr}. В алгоритме 3 минимум на каждой итерации покоординатного спуска определялся методом золотого сечения.

Результаты счета представлены в трех таблицах. В каждой из таблиц для соответствующего семейства представлена зависимость абсолютной погрешности для алгоритмов 1, 2, 3 соответственно, а также время счета на компьютере INTEL(R) Core(TM) i5-2500 CPU @ 3,30GHz 3,60 GHz. При этом для $\Delta E = \max\limits_{S_i|E_{appr}-E_{exact}|}$ вычислялось с помощью алгоритма Кабша и полного перебора 9! возможных паросочетаний, а для S_2 и S_3 в качестве E_{exact} бралось значение алгоритма 2 для значения $m = 128$.

Заключение

1. Из результатов экспериментов видно, что алгоритм 2 имеет существенные преимущества перед алгоритмами 1 и 3. При сопоставимом с алгоритмом 1 времени счета, алгоритм 2, начиная с некоторого значения m, фактически, дает точное значение E, достигая на всех парах оптимального паросочетания перед выполнения алгоритма Кабша. Остаточная погрешность связана с погрешностями численных методов алгебры, используемых в алгоритме Кабша.

Таблица 1: Семейство S_1

m	4	8	16	32	64
Алгоритм 1	$3.15 \cdot 10^{-1}$ 2 сек.	$1.07 \cdot 10^{-1}$ 12 сек.	$5.59 \cdot 10^{-2}$ 1 мин 20 сек.	$2.79 \cdot 10^{-2}$ 9 мин. 39 сек.	$6.07 \cdot 10^{-3}$ 1 ч. 12 мин. 42 сек.
Алгоритм 2	$5.04 \cdot 10^{-2}$ 5 сек.	$1.45 \cdot 10^{-5}$ 25 сек.	$1.45 \cdot 10^{-5}$ 2мин. 54 сек.	$1.34 \cdot 10^{-5}$ 21 мин. 30 сек.	$1.27 \cdot 10^{-5}$ 2ч. 44 мин. 13 сек.
Алгоритм 3	$7.18 \cdot 10^{-2}$ 1 мин. 45 сек.	$5.86 \cdot 10^{-2}$ 5 мин. 20 сек.	$5.86 \cdot 10^{-2}$ 18 мин. 37 сек.	$5.64 \cdot 10^{-2}$ 1ч. 9 мин. 41 сек.	$5.64 \cdot 10^{-2}$ 4 ч. 36 мин. 2 сек.

Таблица 2: Семейство S_2

m	4	8	16	32	64
Алгоритм 1	$3.56 \cdot 10^{-1}$ 2 сек.	$2.40 \cdot 10^{-1}$ 14 сек.	$1.86 \cdot 10^{-1}$ 1 мин 38 сек.	$2.45 \cdot 10^{-2}$ 11 мин. 45 сек.	$1.38 \cdot 10^{-2}$ 1 ч. 29 мин. 22 сек.
Алгоритм 2	$8.26 \cdot 10^{-1}$ 3 сек.	$3.11 \cdot 10^{-2}$ 21 сек.	$3.71 \cdot 10^{-6}$ 2мин. 20 сек.	$3.38 \cdot 10^{-6}$ 17 мин. 43 сек.	$4.22 \cdot 10^{-7}$ 2ч. 14 мин. 49 сек.
Алгоритм 3	$8.55 \cdot 10^{-2}$ 1 мин. 10 сек.	$8.55 \cdot 10^{-2}$ 3 мин. 35 сек.	$8.55 \cdot 10^{-2}$ 12 мин. 27 сек.	$8.55 \cdot 10^{-2}$ 46 мин. 4 сек.	$8.55 \cdot 10^{-2}$ 2 ч. 57 мин. 38 сек.

Таблица 3: Семейство S_3

m	4	8	16	32	64
Алгоритм 1	$7.20 \cdot 10^{-1}$ 5 сек.	$5.02 \cdot 10^{-1}$ 1 мин. 31 сек.	$1.60 \cdot 10^{-1}$ 3 мин 35 сек.	$1.48 \cdot 10^{-1}$ 26 мин. 40 сек.	$1.83 \cdot 10^{-2}$ 3 ч. 33 мин. 51 сек.
Алгоритм 2	$3.80 \cdot 10^{-1}$ 6 сек.	$1.53 \cdot 10^{-1}$ 35 сек.	$8.10 \cdot 10^{-2}$ 4мин. 2 сек.	$1.74 \cdot 10^{-5}$ 30 мин. 8 сек.	$1.39 \cdot 10^{-6}$ 3ч. 46 мин. 49 сек.
Алгоритм 3	$1.28 \cdot 10^{-1}$ 2 мин. 28 сек.	$1.28 \cdot 10^{-1}$ 7 мин. 28 сек.	$1.15 \cdot 10^{-2}$ 25 мин. 25 сек.	$1.15 \cdot 10^{-2}$ 1ч. 24 мин. 40 сек.	$1.15 \cdot 10^{-2}$ 6 ч. 1 мин. 19 сек.

2. Алгоритм 3 демонстрирует отсутствие сходимости. Это связано с тем, что метод золотого сечения находит минимум только для унимодальных функций, а условие унимодальности не всегда выполняется. Однако применение более надежных методов отыскания одномерного минимума приведет к существенному возрастанию времени счета. Применение метода золотого сечения в итерационных алгоритмах показало хорошие результаты для отыскания максимума объема пересечения выпуклых многогранников в работе [9]. Однако для задач из настоящей статьи он оказался менее эффективен.

Список литературы

[1] Kabsch W. A solution of the best rotation to relate two sets of vectors // Acta Crystallographica. 1976. V. 32, P. 922–923.

[2] Kabsch W. A discussion of the solution for the best rotation to relate two sets of vectors // Acta Crystallographica. 1978. V. 34, P. 827–828.

[3] H.W. Kuhn. The Hungarian Method for the Assignment Problem // Naval Research Logistics. 2005. V. 52(1), P. 7–21.

[4] Ali Sadeghi, S. Alireza Ghasemi, Bastian Schaefer, Stephan Mohr, Markus A. Lill, Stefan Goedecker. Metrics for measuring distances in configuration spaces // The Journal of chemical physics. 2013. V. 139, 184118

[5] Allen, W. J.; Rizzo, R. C. Implementation Of The Hungarian Algorithm To Account For Ligand Symmetry And Similarity In Structure-Based Design // J. Chem. Inf. Model. 2014. V. 54, P. 518–529.

[6] Matthew Griffiths, Samuel P. Niblett, and David J. Wales. Optimal Alignment of Structures for Finite and Periodic Systems
J. Chem. Theory Comput. 2017. V. 13, P. 4914–4931

[7] Berhane Temelso, Joel M. Mabey, Toshiro Kubota, Nana Appiah-padi, George C. Shields. ArbAlign: A Tool for Optimal Alignment of Arbitrarily Ordered Isomers Using the Kuhn-Munkres Algorithm
J. Chem. Info. Model. 2017. V. 57 (5), P. 1045–1054

[8] Szymanski, John E. (1989). Basic Mathematics for Electronic Engineers:Models and Applications. Taylor & Francis. P. 154. ISBN 0278000681.

[9] Shevchenko A.H., Blatov I.A., Kitaeva E.V., Blatov V.A. Local coordination versus overall topology in crystal structures: deriving knowledge from crystallographic databases
Cristal Growth and Design. 2017. V. 17, №2. P. 774–785

Блатов Игорь Анатольевич — д.ф.-м.н., зав. кафедрой высшей математики Поволжского государственного университета телекоммуникаций и информатики;
e-mail: blatow@mail.ru;
Китаева Елена Викторовна — к.ф.-м.н., доцент кафедры дифференциальных уравнений и теории управления;
Самарский национальный исследовательский университет имени академика С.П. Королева;
e-mail: el_kitaeva@mail.ru.

ОПРЕДЕЛЕНИЕ ИЗМЕНЕНИЯ ПРОСТРАНСТВЕННО-ВРЕМЕННОГО СОСТОЯНИЯ ТЕХНОГЕННЫХ СИСТЕМ МЕТОДАМИ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Т. Ю. Бугакова

Сибирский государственный университет геосистем и технологий, 630108, Новосибирск

УДК 528.2:528.4
DOI: 10.24411/9999-016A-2019-10010

В совокупности геодинамические и техногенные системы представляют собой сложные системы, главной проблемой обеспечения безопасности которых, является невозможность полностью исключить риск возникновения чрезвычайных ситуаций в таких системах и необходимость свести этот риск к минимуму. Для контроля пространственно-временного состояния (ПВС) техногенных систем необходимы данные об их геометрических свойствах, как функциях времени. К ним относятся форма, размеры, положение в пространстве и другие свойства, характеризующие взаимное расположение множества элементов системы относительно внешней среды и относительно друг друга. Исходными данными для моделирования служат временные ряды координат множества контрольных точек исследуемой системы, полученные по результатам повторных циклов геодезических измерений. Выполнить непосредственное измерение таких свойств геодезическими средствами чаще всего не удаётся и поэтому для их определения в работе применяются методы численного моделирования.

Ключевые слова: пространственно-временное состояние, техногенная система, математические методы, чрезвычайные ситуации, геодезический контроль, риск, мультиагентная система.

Введение

Проблемы риска и производственной безопасности техногенных систем в настоящее время имеют большее значение. Особую актуальность они приобретают в связи с возникновением техногенных катастроф при строительстве и эксплуатации крупных инженерных сооружений, например, в атомной энергетике, химической промышленности, машиностроении, на транспорте и других отраслях хозяйственной деятельности. Причинами катастроф в основном являются неправильная эксплуатация человеко-машинных систем и различные геодинамические процессы и явления. Так как техногенная система (ТС) — это сложная, искусственно созданная человеком конструкция, которая работает в контакте с природной окружающей средой, то обеспечение безопасности техногенных систем основано на совокупности задач относящихся к различным областям знаний. К ним относятся задачи политического, социально-экономического, управленческого, технологического характера и т.д.

Одним из признаков возникновения техногенных катастроф является изменение состояния ТС. Состояние техногенных систем определяется множеством свойств. Поэтому производить оценку состояния ТС необходимо комплексно, привлекая к этому процессу множество специалистов из различных областей профессиональной деятельности (строителей, архитекторов, геодезистов и т.д.). Методом геодезического контроля можно оценить только пространственно-временное состояние (ПВС) техногенной системы, т.е. определить ее положение в пространстве относительно неподвижной системы координат и времени [1].

Для определения и оценки ПВС конструкций зданий и сооружений, применяются современные технологии геодезического контроля, такие как, методы нивелирования, лазерное сканирование, автоматизированные системы мониторинга, ГНСС и др. Существующие регламенты и инструкции по эксплуатации ТС определяют порядок, объем работ и периодичность геодезических измерений. Для выявления изменений ПВС

ISBN 978-5-901548-42-4

ТС, геодезические данные проходят определенную обработку при помощи специализированных программ, в которых заложены математические алгоритмы. Однако, анализ современных ГИС и систем проектирования, позволил сделать вывод о том, что среди программных продуктов, существующих на российском рынке, нет такого программного решения, который позволил бы системно (комплексно) подходить к проблеме определения, анализа и контроля пространственно-временных состояний ТС по геодезическим данным. Отсутствует автоматизированная информационно-аналитическая система обработки непрерывно-дискретного потока данных в режиме реального времени, соединяющая в единый алгоритм контроль (получение геодезических данных о ПВС ТС), оценку, анализ ПВС ТС (как в целом, так и конкретно по параметрам), прогнозирование сценариев возможного развития событий для своевременных управленческих решений и предупреждения чрезвычайных ситуаций. Отсутствуют интеллектуальные системы контроля ПВС ТС по геодезическим данным. Исследование современных методов обработки геодезических данных обнаружило отсутствие системной методологии, рассматривающей объект как систему, ориентирующей исследование на раскрытие целостности объекта, на выявление многообразных типов связей в нем и сведение их в единую теоретическую картину в целях комплексной оценки изменения пространственно-временного состояния ТС, установления причинно-следственных связей и прогнозирования сценариев развития динамических процессов.

Поэтому существует проблема регламентирующая отсутствие системной методологии определения пространственно-временного состояния ТС по геодезическим данным и отсутствие автоматизированной информационно-аналитической системы обработки непрерывно-дискретного потока данных в режиме реального времени. Разработка системной методологии и информационно-аналитической системы для комплексной непрерывной во времени оценки, анализа и прогноза пространственно-временного состояния техногенных систем дает возможность своевременно принимать управленческие решения для предотвращения чрезвычайных ситуаций при эксплуатации техногенных систем.

Таким образом, в настоящее время ситуация в геодезии характеризуется наличием объективного противоречия между новыми технологическими возможностями получения и обработки геодезических данных и отсутствием системной методологии увязывающей в единый информационно-аналитический комплекс определение, оценку и прогнозирование сценариев изменения ПВС ТС для непрерывного во времени контроля ПВС и своевременного принятия управленческих решений.

1 Методология исследований

Проблему отсутствия системной методологии определения ПВС ТС по геодезическим данным и отсутствие автоматизированной информационно-аналитической системы обработки непрерывно-дискретного потока данных в режиме реального времени позволяют решить мультиагентные технологии, которые реализуются на основе мультиагентных систем, применяемых там, где протекает большой поток информации и требуется принятие управленческих решений.

Мультиагентная система (МАС) основана на взаимодействии между собой интеллектуальных агентов, способных функционально общаться между собой и совместно принимать решения. Каждый агент включает в себя решение задачи анализа, прогноза и диагностики, выявления скрытых зависимостей и поддержки принятия оптимальных решений, на основе заложенных в него математических методов и алгоритмов обработки данных [2].

Основными преимуществами мультиагентных систем являются: параллельное выполнение операций, их способность к самоорганизации, управление знаниями, способность принимать распределенные управленческие решения, управлять ресурсами. На сегодняшний день существует множество мультиагентных систем, которые решают задачи в самых разных областях: поиск, электронная коммерция, оптимальное динамическое планирование производства и сбыта продукции, логистика, экономика, транспорт, моделирование, телекоммуникации, однако в геодезическом производстве мультиагентные системы не применяются.

Результатом достижения глобальной цели МАС в геодезии является определение пространственно-временного состояния ТС в реальном режиме времени, прогнозирование, подбор оптимальных математических методов определения ПВС ТС и применение интеллектуальных алгоритмов принятия решений [3].

На рис. 1 приведена обобщенная схема мультиагентного подхода к решению задачи определения пространственно-временного состояния ТС, состоящий из блока сбора данных (агент 1), блока аналитики (агент 2) и блока управления (агент 3).

Агент 1 представляет собой автоматизированную систему мониторинга, изображенную на рисунке 1. Она состоит из множества контрольных устройств, установленных в теле системы (датчиков), приемника

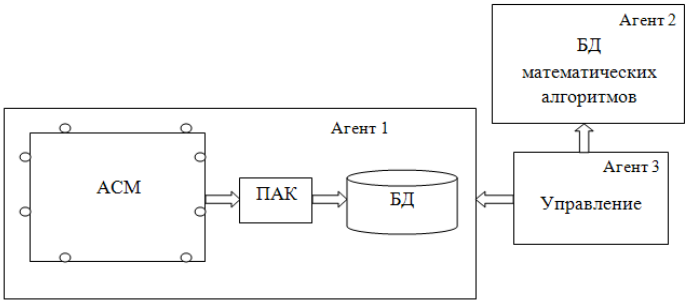


Рис. 1: Концептуальная модель мультиагентной системы (МАС)

и преобразователя сигналов (ПиПС), базы данных. Этап сбора данных предполагает поступление полной (необходимой для принятия решения) информации о техногенной системе и ее передачу в модуль управления. Данные, передаваемые контрольными устройствами, имеют k степеней свободы, что позволяет в дальнейшем установить корреляцию процессов и установить причинно-следственные связи.

Агент 2 содержит базу данных математических алгоритмов для решения задач определения пространственно-временного состояния ТС, и алгоритм выбора стратегии, в основе которого лежит функция эффективности применения того или иного алгоритма.

Агент 3 принимает управленческие решения, осуществляет обратную связь с ТС и осуществляет передачу информации пользователю. Функциями этого агента являются принятие решений о частоте дискретизации поступления данных от агента 1, декомпозиции ТС, определение ее структурных частей, требующих детального рассмотрения и выявления причины изменения ПВС, локализации мест деформации и установления причинно-следственных связей [4].

2 Численные методы определения изменения пространственно-временного состояния техногенных систем

В настоящее время существует многообразие программных продуктов для обработки геодезических данных. Основным инструментом, применяемым в большинстве программного обеспечения для определения ПВС ТС, являются математические методы. Разнообразие целей, уникальность ТС и особенности изменения ПВС часто требуют индивидуального подхода к обработке результатов геодезических измерений. При этом часто возникают трудности выбора оптимального математического метода для решения поставленной задачи. В связи с этим необходимо иметь базу математических алгоритмов и методику формирования оптимального математического алгоритма для конкретно поставленной цели.

Исходными данными для моделирования пространственно-временного состояния техногенных систем служат временные ряды координат $X_i, Y_i, Z_i, (i = 1, 2 \dots n)$ множества n геодезических контрольных точек, установленных в теле ТС, полученные по результатам повторных циклов геодезических измерений. Для контроля пространственно-временного состояния техногенных систем необходимы данные об их геометрических свойствах, как функциях времени. К ним относятся свойства, характеризующие взаимное расположение множества контрольных геодезических точек системы относительно внешней среды и относительно друг друга. Выполнить непосредственное измерение таких свойств даже современными техническими средствами чаще всего не удаётся и поэтому для их определения применяют методы численного моделирования.

Рассматривая отдельные геодезические контрольные точки или некоторые их множества, связанные заданными отношениями (как элементарные объекты) можно установить структуру, геометрические параметры которой будут являться свойствами ТС, характеризующими ее пространственное состояние. Возможность различного выбора элементарных объектов обеспечивает свободу в определении структуры. В результате на множестве исходных геодезических данных могут быть определены геометрические объекты не обязательно состоящие из конечного множества точек. Например, прямая, проходящая через две заданные точки, плоскость, содержащая три заданные точки, многоугольник, составленный из отрезков прямых и т.д. Это позволяет находить свойства объектов в виде различных геометрических признаков (рисунок 2):

- внешнюю конфигурацию, в которой отображается структура объекта (точка, линия, полоса, оболочка, стержень, слой) и размерность пространства состояний;
- количество и размерность связей со смежными элементами, иерархию связей;
- уравнения линий и поверхностей;
- числовые характеристики.

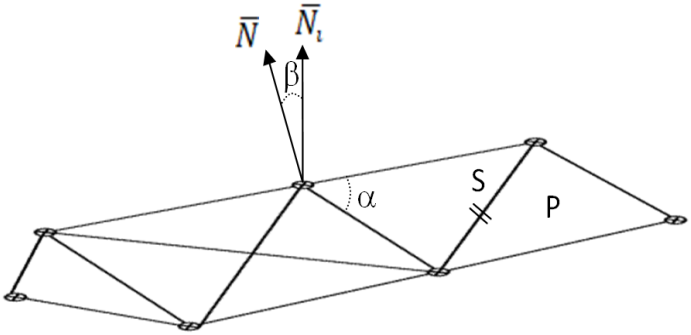


Рис. 2: Пример структуры, определенной на множестве контрольных геодезических точек с координатами X, Y, Z

На рис. 2 обозначено:

- $\overline{N}, \overline{N_i}$ — нормали, проведенные к плоскости конечных элементов;
- S — расстояние между контрольными точками;
- P — площади конечных элементов;
- α — углы между направлениями;
- β — углы между нормалями.

Приведенные параметры являются основными и простейшими геометрическими характеристиками ТС, которые могут быть получены по геодезическим данным. Кроме этих характеристик, существует множество других вариантов: условие принадлежности четырех точек одной и той же плоскости, угол между плоскостями, угол между прямой и плоскостью, расстояние от точки до прямой, проекция вектора на плоскость и т.д. Все они являются основными элементарными объектами, из которых можно, исходя из целей моделирования или конструктивных особенностей ТС, оценивать качественные свойства и вычислять значения геометрических характеристик [1].

Как правило, изменение ПВС ТС происходит в результате влияния внешних факторов, что приводит к изменению положения ТС относительно системы координат и (или) деформациям. Анализируя движения и деформации, можно судить об «опасности» состояния ТС и принимать необходимые меры для снижения риска и ущерба от возникновения чрезвычайных ситуаций. Особую опасность для эксплуатации ТС представляют деформации. Фактически деформации представляют собой движения частей системы относительно друг друга, сопровождающиеся изменениями формы и размеров. В геодезии существуют различные методы определения деформаций, в основе которых заложены алгоритмы вычисления приращений координат контрольных точек относительно первоначального состояния или относительно друг друга, а также вычисления параметров отклонения элементов конструкции ТС от принятых стандартов.

Применение численных математических методов обработки данных позволяют представить изучаемую ТС в виде математической модели, увязывающей математическими соотношениями всю систему в целом и, таким образом, определить ее форму. Форма и размеры систем определяются границей, отделяющей ее от внешней среды. Форма любой системы определяется набором интегральных и дифференциальных характеристик. Интегральными характеристиками, например, являются геометрические свойства всей системы — возможность ее представления одним геометрическим телом, его размеры, площадь поверхности,

объем занимаемого пространства, числовые значения инвариантных характеристик. Дифференциальными характеристиками системы служат направления касательных и нормалей к поверхностям и/или линиям, ограничивающим систему, их кривизна, площади частей поверхности и длины линий, охватывающих эти части и другие.

Для определения формы ТС, по координатам множества геодезических точек необходимо, в соответствии с целью, выбрать геометрический образ, который принимается в качестве модели формы, и определить требования (критерии), которым этот образ должен удовлетворять. После этого нужно оценить значения конечного числа параметров, необходимых для математического описания выбранного геометрического образа формы системы соответственно предъявляемым требованиям [5, 6].

На рисунках 3–7 приведены примеры математического описания геометрических образов формы системы.

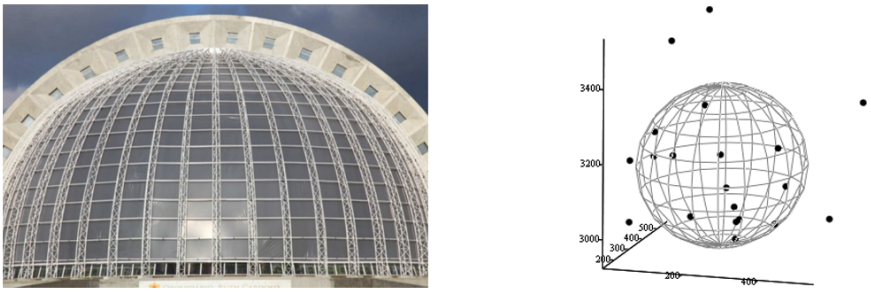


Рис. 3: Аппроксимация множества точек сферой

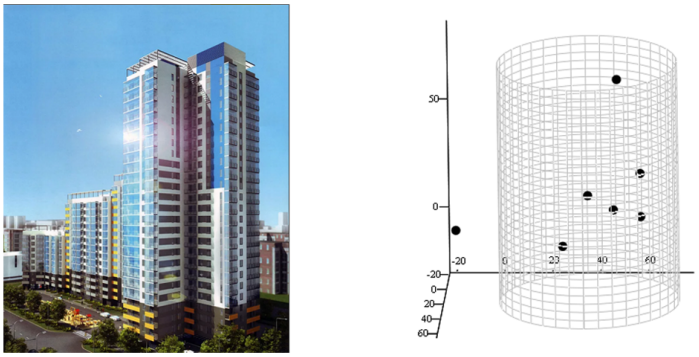


Рис. 4: Аппроксимация множества точек цилиндром

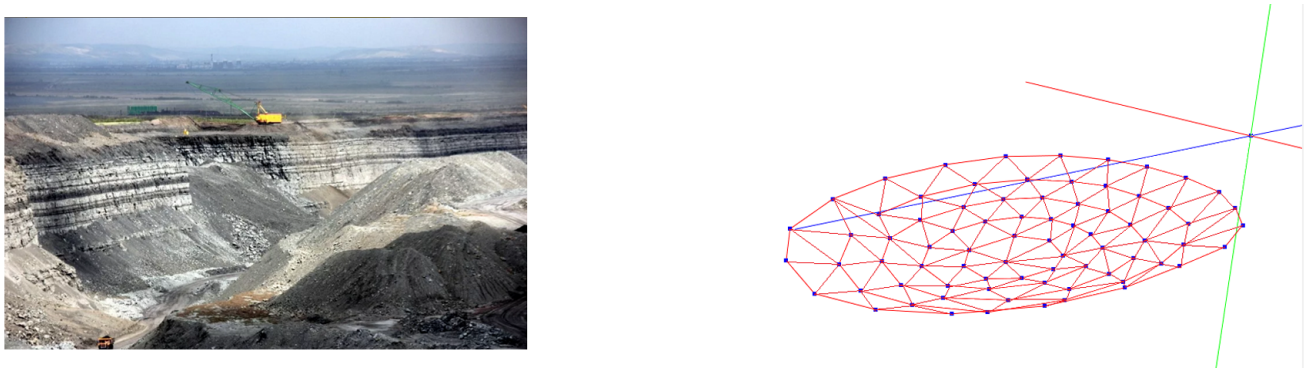


Рис. 5: Аппроксимация множества точек методом конечных элементов (метод Делоне)



Рис. 6: Аппроксимация множества точек спиралью в параметрическом представлении

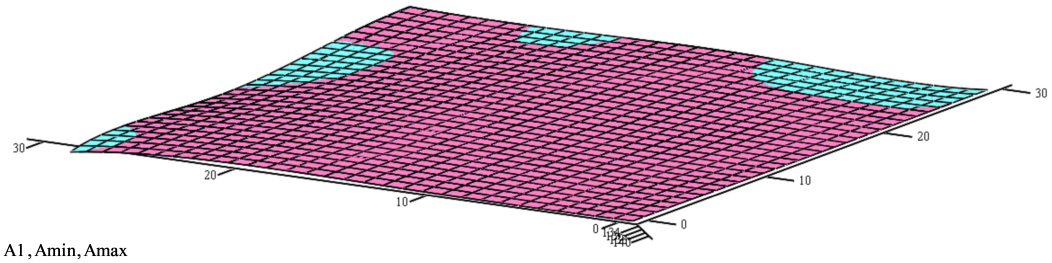


Рис. 7: Определение деформации фундамента инженерного объекта методом сплайн-интерполяции

Изменение формы ТС, как функции времени является явным признаком ее деформации. А математическая модель формы позволяет определить количественные показатели деформации ТС в любой точке системы, а не только в точках геодезического контроля [7].

3 Результаты

В результате работы была сформирована база математических алгоритмов определения ПВС ТС, основанных на численных методах моделирования, которая может быть использована при разработке мультиагентой системы определения пространственно-временного состояния техногенных систем. Приведенные методы математического моделирования позволят представить ПВС ТС в виде единой теоретической картины в целях комплексной оценки изменения пространственно-временного состояния ТС, установления причинно-следственных связей и прогнозирования сценариев развития динамических процессов.

Заключение

Изложенные математические методы определения пространственно-временного состояния ТС могут быть использованы в мультиагентных системах, а также в программных продуктах, предназначенных для обработки геодезических данных. Подбор оптимального математического решения при оценке ПВС ТС позволит своевременно принимать управленческие решения для предотвращения чрезвычайных ситуаций при эксплуатации техногенных систем и значительно снизить риск возникновения чрезвычайных ситуаций техногенного характера.

Результаты исследований найдут применение при оценке и анализе техногенного риска в процессе проектирования, строительства и эксплуатации инженерно-технических систем: зданий и сооружений, дорог, нефте- и газопроводов, при решении задач физической геодинамики, в картографии, экологии и геоинформатике, а также могут быть использованы для определения пространственно-временного состояния технических систем любой сложности.

Список литературы

[1] Бугакова Т. Ю., Вовк И. Г. Определение вращательного движения объекта по результатам многократных геодезических измерений [Текст] // Интерэкспо Гео-Сибирь — 2013: IX Междунар. науч. конгр., 15-26 апр. 2013, — Новосибирск: СГГА, 2013. — С.88–92.

[2] Бугакова Т.Ю., Кноль И.А., Шарапов А.А. Разработка аппаратно-программного комплекса для прогнозирования и определения оптимального варианта изменения пространственно-временного состояния техногенных объектов // Регулирование земельно-имущественных отношений в России: правовое и геопространственное обеспечение, оценка недвижимости, экология, технологические решения. 2018. Т. 2. С. 183–190.

[3] Бугакова Т.Ю., Шарапов А.А. Алгоритмы функционирования мультиагентной системы определения пространственно-временных состояний объекта // Интерэкспо ГЕО-Сибирь-2017. XIII Междунар. науч. конгр., 17–21 апреля 2017 г., Новосибирск. Т. 7. С. 3–7.

[4] Бугакова Т.Ю., Кноль И.А. Мультиагентное моделирование пространственно-временного состояния техногенных объектов. // Интерэкспо ГЕО-Сибирь-2017. XIII Междунар. науч. конгр., 17–21 апреля 2017 г., Новосибирск. Т. 7. С. 8–13.

[5] Бугакова Т.Ю., Шарапов А.А. Разработка прототипа мультиагентной системы контроля состояния техногенных объектов //Вестник СГУГиТ. — 2017. — Вып. 1 (22). С. 142–150.

[6] Бугакова Т.Ю., Кноль И.А., Интерактивный контроль пространственно-временного состояния техногенных объектов с применением технологии WEBGL// Вестник СГУГиТ. — 2016. — Вып. 4 (36). — С. 114–121.

[7] Бугакова Т.Ю., Басаргин А.А., Каленицкий А.И. Применение ГИС-технологий и методов математического моделирования для определения крена плиты фундамента инженерного сооружения // Вестник СГУГиТ. — 2018. — Вып. 2 (23). — С. 70–78.

Бугакова Татьяна Юрьевна — к.т.н., доцент, заведующий кафедрой прикладной информатики и информационных систем Сибирского государственного университета геосистем и технологий;
e-mail: Bugakova-tu@yandex.ru.
Дата поступления — 30 апреля 2019 г.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНОГО «ЯДЕРНОГО» АЛГОРИТМА И МЕТОДА ПОЛИГОНА ЧАСТОТ

Т. Е. Булгакова[1], А. В. Войтишек[1,2]

[1] *Новосибирский государственный университет, 630090, Новосибирск*

[2] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

УДК 519.676
DOI: 10.24411/9999-016A-2019-10011

В работе показано, что дополненный соображениями численной сеточной аппроксимации функций «ядерный» алгоритм приближения вероятностной плотности конструктивно совпадает с рандомизированным проекционно-сеточным функциональным численным алгоритмом многомерного аналога метода полигона частот для приближения решения уравнения Фредгольма второго рода. Из этого следует, что для «ядерного» алгоритма приближения вероятностной плотности можно использовать соображения теории условной оптимизации многомерного аналога метода полигона частот. В свою очередь, для развития теории конструирования и условной оптимизации рандомизированных проекционно-сеточных функциональных численных алгоритмов можно использовать соображения из теории «ядерных» оценок вероятностных плотностей.

Ключевые слова: рандомизированные проекционно-сеточные функциональные алгоритмы, многомерный аналог метода полигона частот, «ядерные» оценки вероятностных плотностей, численная аппроксимация функций

1. Многомерный аналог метода полигона частот

В последние годы (главным образом в новосибирской школе методов Монте-Карло) развивается теория рандомизированных функциональных алгоритмов (см., например, [1, 2]). Наиболее содержательные примеры применения этих алгоритмов связаны с приближением неизвестного решения $\varphi(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ интегрального уравнения Фредгольма второго рода

$$\varphi(\mathbf{x}) = \int k(\mathbf{x}', \mathbf{x})\varphi(\mathbf{x}')\, d\mathbf{x}' + f(\mathbf{x}) \quad \text{или} \quad \varphi = K\varphi + f, \tag{1.1}$$

на ограниченной области $X \subset \mathbb{R}^d$; здесь $k(\mathbf{x}', \mathbf{x})$ (ядро интегрального оператора K) и $f(\mathbf{x})$ (свободный член уравнения) — заданные функции.

Для приближения функции $\varphi(\mathbf{x})$ используем представления классической теории численной аппроксимации функций (см., например, [3]), имеющих общий вид

$$\varphi(\mathbf{x}) \approx L^{(M)}\varphi(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)}\chi^{(i)}(\mathbf{x}) \tag{1.2}$$

для некоторого специально выбранного набора базисных функций

$$\mathbf{\Xi}^{(M)} = \left\{\chi^{(1)}(\mathbf{x}), ..., \chi^{(M)}(\mathbf{x})\right\}, \tag{1.3}$$

Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2019-0002), а также при финансовой поддержке Совета по грантам президента РФ и государственной поддержке ведущих научных школ (код проекта НШ–5913.2018.1).

ISBN 978-5-901548-42-4

(вид этих функций определяет тип аппроксимации (1.2)) и коэффициентов

$$\mathbf{W}^{(M)} = \left\{ w^{(1)}, ..., w^{(M)} \right\}, \tag{1.4}$$

определяемых как функционалы от неизвестной приближаемой функции $\varphi(\mathbf{x})$.

В рандомизированных функциональных алгоритмах коэффициенты (1.4) вычисляются приближенно методом Монте-Карло $w^{(i)} \approx \tilde{w}^{(i)}(n_i)$ с числом испытаний n_i (в данной работе будет изучаться случай $n_1 = ... = n_M \equiv n$) и рассматривается приближение

$$\varphi(\mathbf{x}) \approx L^{(M)}\tilde{\varphi}(\mathbf{x}) = \sum_{i=1}^{M} \tilde{w}^{(i)}(n)\chi^{(i)}(\mathbf{x}). \tag{1.5}$$

В работах [4, 5] предложена новая (по сравнению с книгами [1, 2]) классификация рандомизированных функциональных алгоритмов для приближения решения уравнения (1.1): выделены сеточные, проекционные и проекционно-сеточные алгоритмы (тип метода определяется выбором базисных функций (1.3), коэффициентов (1.4) и приближения (1.5)). В этих же работах приведены соображения о том, почему сеточные и проекционные рандомизированные функциональные алгоритмы могут быть неэффективными (и даже нереализуемыми) при решении практически значимых задач, связанных с решением уравнений вида (1.1). Так, для теоретически привлекательного сеточного метода зависимых испытаний требуется гладкость ядра $k(\mathbf{x}', \mathbf{x})$ интегрального оператора K, а подавляющее большинство ядер из прикладных задач содержат интегрируемые особенности (вплоть до дельта-функций) и даже не могут быть явно вычислены. Сеточный метод сопряженных блужданий является крайне трудоемким из-за необходимости моделирования индивидуального набора траекторий соответствующих прикладных цепей Маркова для каждого узла $\mathbf{x}_i$ вводимой сетки

$$X^{(M)} = \{\mathbf{x}_1, ..., \mathbf{x}_M\} \tag{1.6}$$

в области X. Проекционные методы обладают достаточно очевидной численной неустойчивостью.

Проекционно-сеточные рандомизированные функциональные алгоритмы не обладают перечисленными недостатками. Для этих алгоритмов базисные функции (1.3) и коэффициенты $w^{(i)} = w^{(i)}\left(\boldsymbol{\varphi}^{(M)}\right)$; $\boldsymbol{\varphi}^{(M)} = \{\varphi(\mathbf{x}_1), ..., \varphi(\mathbf{x}_M)\}$ из (1.4) связаны с сеткой (1.6) таким образом, что обеспечивают как малость детерминированной компоненты погрешности

$$\delta_{det}^{(\mathbb{B}(X))} = \|\varphi - L^{(M)}\varphi\|_{\mathbb{B}(X)}$$

для используемого нормированного функционального пространства $\mathbb{B}(X)$, а также устойчивость приближения (1.5), определяемую относительной малостью (близостью к единице) константы Лебега $\tilde{L} = \sup_{\mathbf{x}\in X} \sum_{i=1}^{M} \left|\chi^{(i)}(\mathbf{x})\right|$ из соотношения

$$\delta_{stoch}^{(\mathbb{C}(X))} = \|L^{(M)}\varphi - L^{(M)}\tilde{\varphi}\|_{\mathbb{C}(X)} \le \tilde{L} \max_{i=1,...,M} \left| w^{(i)}\left(\boldsymbol{\varphi}^{(M)}\right) - w^{(i)}\left(\tilde{\boldsymbol{\varphi}}^{(M)}(n)\right)\right|$$

(см., например, [1, 2]); здесь $\tilde{\boldsymbol{\varphi}}^{(M)}(n) = \left(\tilde{\varphi}^{(\mathbf{x}_1)}(n), ..., \tilde{\varphi}^{(\mathbf{x}_M)}(n)\right)$, а $\tilde{\varphi}^{(\mathbf{x}_i)}(n)$ — монте-карловское приближение значения $\varphi(\mathbf{x}_i)$; $i = 1, ..., M$.

В этом случае приближения метода Монте-Карло

$$\tilde{\mathbf{W}}^{(M)} = \left\{ \tilde{w}^{(1)}(n), ..., \tilde{w}^{(M)}(n) \right\}$$

для коэффициентов (1.4) из соотношения (1.5) имеют вид

$$\tilde{w}^{(i)}(n) = w^{(i)}\left(\tilde{\boldsymbol{\varphi}}^{(M)}(n)\right), \text{ чаще всего } w^{(i)}\left(\tilde{\boldsymbol{\varphi}}^{(M)}(n)\right) = \tilde{\varphi}^{(\mathbf{x}_i)}(n).$$

В свою очередь, для получения значений $\tilde{\boldsymbol{\varphi}}^{(M)}(n)$ в рандомизированных функциональных *проекционно-сеточных* алгоритмах применяется следующая специальная технология (определяющая, в частности, отличие от *сеточных* функциональных алгоритмов). Выбираются финитные, одинаковые по форме для всех $\{\mathbf{x}_1, ..., \mathbf{x}_M\}$ функции (версии «ядерной» функции $\kappa^{(\mathbf{x})}(\mathbf{y})$ для различных значений параметра $\mathbf{x}$ — см. раздел 2 данной статьи)

$$\mathbf{K}^{(M)} = \left\{ \kappa^{(\mathbf{x}_1)}(\mathbf{y}), ..., \kappa^{(\mathbf{x}_M)}(\mathbf{y}) \right\}, \tag{1.7}$$

так же, как и базисные функции (1.3), связанные с сеткой (1.6), и такие, что

$$\int \varphi(\mathbf{y})\kappa^{(\mathbf{x}_i)}(\mathbf{y})\,d\mathbf{y} \approx \varphi(\mathbf{x}_i);\;\; i=1,...,M; \tag{1.8}$$

см. [1, 2, 6, 7].

Далее следует напомнить ставшие уже классическими (см., например, главу 4 из учебника [1]) соображения о том, что для приближенного вычисления линейных функционалов вида

$$I_h = \int \varphi(\mathbf{y})h(\mathbf{y})\,d\mathbf{y} \tag{1.9}$$

от решения $\varphi(\mathbf{x})$ уравнения (1.1) целесообразно использовать *основной оцениватель* (или *несмещенную монте-карловскую оценку по столкновениям*):

$$I_h = \mathbf{E}\zeta;\;\; \zeta = \sum_{m=0}^{N} Q^{(m)} h\left(\boldsymbol{\xi}^{(m)}\right), \tag{1.10}$$

где

$$\boldsymbol{\xi}^{(0)}, \boldsymbol{\xi}^{(1)}, ..., \boldsymbol{\xi}^{(N)} \tag{1.11}$$

является *прикладной цепью Маркова* (или однородной цепью Маркова, обрывающейся с вероятностью единица) с начальной плотностью $\pi(\mathbf{x})$ и переходной функцией $p(\mathbf{x}',\mathbf{x}) = r(\mathbf{x}',\mathbf{x}) \times \left[1 - p^{(a)}(\mathbf{x}')\right]$ (здесь $r(\mathbf{x}',\mathbf{x})$ — вероятностная переходная плотность, а $0 \le p^{(a)}(\mathbf{x}') \le 1$ обозначает вероятность обрыва траектории; соответственно, N — это случайный номер обрыва траектории). Случайные веса $\{Q^{(m)}\}$ из (1.10) определяются следующими рекуррентными соотношениями:

$$Q^{(0)} = \frac{f\left(\boldsymbol{\xi}^{(0)}\right)}{\pi\left(\boldsymbol{\xi}^{(0)}\right)};\;\; Q^{(m)} = Q^{(m-1)} \times \frac{k\left(\boldsymbol{\xi}^{(m-1)}, \boldsymbol{\xi}^{(m)}\right)}{p\left(\boldsymbol{\xi}^{(m-1)}, \boldsymbol{\xi}^{(m)}\right)};\;\; m=1,...,N. \tag{1.12}$$

Учитывая, что соотношения (1.8) имеют вид (1.9) получаем следующий ***рандомизированный проекционно-сеточный функциональный алгоритм***.

АЛГОРИТМ 1.1. *Моделируя n траекторий*

$$\boldsymbol{\xi}_j^{(0)}, \boldsymbol{\xi}_j^{(1)}, ..., \boldsymbol{\xi}_j^{(N_j)};\;\; j=1,...,n \tag{1.13}$$

прикладной цепи Маркова (1.15), *получаем значения*

$$\tilde{\varphi}^{(\mathbf{x}_i)}(n) = \frac{1}{n}\sum_{j=1}^{n}\sum_{m=0}^{N_j} Q_j^{(m)} \kappa^{(\mathbf{x}_i)}\left(\boldsymbol{\xi}_j^{(m)}\right);\;\; i=1,...,M;$$

здесь веса $\left\{Q_j^{(m)}\right\}$ *вычисляются по формулам вида* (1.12)*:*

$$Q_j^{(0)} = \frac{f\left(\boldsymbol{\xi}_j^{(0)}\right)}{\pi\left(\boldsymbol{\xi}_j^{(0)}\right)};\;\; Q_j^{(m)} = Q_j^{(m-1)} \times \frac{k\left(\boldsymbol{\xi}_j^{(m-1)}, \boldsymbol{\xi}_j^{(m)}\right)}{p\left(\boldsymbol{\xi}_j^{(m-1)}, \boldsymbol{\xi}_j^{(m)}\right)};\;\; j=1,...,n;\;\; m=1,...,N_j.$$

Затем приближаем функцию $\varphi(\mathbf{x})$ по формуле вида (1.5)*:*

$$\varphi(\mathbf{x}) \approx L^{(M)}\tilde{\varphi}(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)}\left(\tilde{\varphi}^{(\mathbf{x}_1)}(n), ..., \tilde{\varphi}^{(\mathbf{x}_M)}(n)\right)\chi^{(i)}(\mathbf{x}). \tag{1.14}$$

В работах [1, 2, 6, 7] приведены соображения *теории условной оптимизации*, подтверждающие целесообразность использования в качестве базисных функций (1.3) «абсолютно устойчивых» финитных функций *мультилинейной аппроксимации* (или аппроксимации Стренга — Фикса [8] с производящей функцией $\beta^{(1)}(u)$, являющейся B-сплайном первого порядка) на регулярной сетке с шагом h по каждой координате:

$$\chi^{(i)}(\mathbf{x}) = \beta^{(1)}\left(\frac{x^{(1)}}{h} - j_i^{(1)}\right) \times ... \times \beta^{(1)}\left(\frac{x^{(d)}}{h} - j_i^{(d)}\right);\quad \beta^{(1)}(u) = \begin{cases} u+1, \text{ при } -1 \le u \le 0; \\ -u+1, \text{ при } 0 \le u \le 1; \\ 0 \text{ иначе;} \end{cases} \tag{1.15}$$

$$\mathbf{x} = \left(x^{(1)}, ..., x^{(d)}\right) \;\; \mathbf{x}_i = \left(j_i^{(1)}h, ..., j_i^{(d)}h\right); \;\; j_i^{(k)} \;\; - \text{целые числа}; \quad i = 1, ..., M \tag{1.16}$$

(здесь множество X, на котором происходит приближение решения $\varphi(\mathbf{x})$ уравнения (1.1), представляет собой прямоугольный параллелепипед). Кроме того, «ядерную» функцию из соотношений (1.7), (1.8) предлагается выбирать в виде

$$\kappa^{(\mathbf{x})}(\mathbf{y}) = \begin{cases} \frac{1}{h^d} & \text{при} \;\; \mathbf{x} \in \Delta^{(\mathbf{x})}, \\ 0 & \text{иначе}, \end{cases} \tag{1.17}$$

где $\Delta^{(\mathbf{x})} = \left\{\mathbf{y} = \left(y^{(1)}, ..., y^{(d)}\right) : \; x^{(s)} - h/2 < y^{(s)} < x^{(s)} + h/2; \;\; s = 1, ..., d; \;\; \mathbf{x} = \left(x^{(1)}, ..., x^{(d)}\right)\right\}$.

При этом приближения аппроксимационных коэффициентов (1.4) имеют простейший вид

$$w^{(i)}\left(\tilde{\varphi}^{(\mathbf{x}_1)}(n), ..., \tilde{\varphi}^{(\mathbf{x}_M)}(n)\right) = \tilde{\varphi}^{(\mathbf{x}_i)}(n). \tag{1.18}$$

Алгоритм 1.1, в котором используются функции (1.15), (1.17) и аппроксимационные коэффициенты (1.18), назван в [1, 2, 6, 7] ***многомерным аналогом метода полигона частот***.

В разделе 2 данной работы показано, что описанный подход к построению проекционно-сеточного алгоритма 1.1 в определенной степени аналогичен построению *«ядерных» оценок вероятностных плотностей* (см., например, [9]). Особо отмечено, что в работах по теории «ядерных» оценок (в том числе в [9]), авторы, рассуждая о приближениях функций, неоправданно не включают в рассмотрение элементы теории численной аппроксимации функций (см., например, [3]). При добавлении этого упущенного элемента соответствующие «ядерные» приближения вероятностных плотностей по сути совпадают с алгоритмом 1.1.

2. Численная аппроксимация вероятностных плотностей с использованием «ядерных» оценок

В классической работе [9] рассматривается непараметрическая оценка вероятностной плотности распределения $P(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ по выборочным значениям $\left\{\hat{\boldsymbol{\xi}}_1, ..., \hat{\boldsymbol{\xi}}_n\right\} \subset \mathbb{R}^d$ из этого распределения вида

$$P(\mathbf{x}) \approx Z_n(\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} \kappa^{(\mathbf{x})}\left(\hat{\boldsymbol{\xi}}_j\right), \tag{2.1}$$

где $\kappa^{(\mathbf{x})}(\mathbf{y})$ — некоторая финитная параметрическая, одинаковая по форме для всех значений параметра $\mathbf{x}$, *«ядерная» функция*. Приближение (2.1) называется *«ядерной» оценкой плотности* $P(\mathbf{x})$. К слову, в работе [9] и других источниках (в частности, в разделе 2 работы [10]) слова «ядро», «ядерная» приводятся без кавычек, но в данной работе мы используем кавычки для того, чтобы различить по названиям функции $k(\mathbf{x}', \mathbf{x})$ (ядро интегрального уравнения (1.1)) и $\kappa^{(\mathbf{x})}(\mathbf{y})$.

При исследовании свойств приближения (2.1) существенно используется то обстоятельство, что, согласно закону больших чисел, для достаточно больших n выполнено

$$Z_n(\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} \kappa^{(\mathbf{x})}\left(\hat{\boldsymbol{\xi}}_j\right) \approx \mathbf{E}\kappa^{(\mathbf{x})}\left(\hat{\boldsymbol{\xi}}\right) = \int \kappa^{(\mathbf{x})}(\mathbf{y})P(\mathbf{y})\, d\mathbf{y}. \tag{2.2}$$

Определенным конструктивным недостатком теории «ядерных» оценок плотности и ее приложений (см., в частности, работы [9, 10]) является отсутствие обсуждения алгоритма практического (в первую очередь, численного, компьютерного) приближения функции $P(\mathbf{x})$ в целом, основанного на теории сеточной аппроксимации функций (см., например, [3]). Такой алгоритм мог бы выглядеть следующим образом.

Учитывая конечность выборки $\left\{\hat{\boldsymbol{\xi}}_1, ..., \hat{\boldsymbol{\xi}}_n\right\}$ можно рассмотреть ограниченное множество $X \subset \mathbb{R}^d$, которому принадлежат все выборочные значения, ввести в этой области сетку (1.6) и рассмотреть приближение вида (1.2) для функции $P(\mathbf{x})$:

$$P(\mathbf{x}) \approx L^{(M)}P(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)}\left(P(\mathbf{x}_1), ..., P(\mathbf{x}_M)\right)\chi^{(i)}(\mathbf{x}).$$

АЛГОРИТМ 2.1. *Вычисляем значения* $\tilde{P}^{(\mathbf{x}_i)}(n) = Z_n(\mathbf{x}_i)$; $i = 1, ..., M$ *по формулам вида* (2.1) *и приближаем функцию* $P(\mathbf{x})$ *по формуле вида* (1.14)*:*

$$P(\mathbf{x}) \approx L^{(M)}\tilde{P}(\mathbf{x}) = \sum_{i=1}^{M} w^{(i)}\left(\tilde{P}^{(\mathbf{x}_1)}(n), ..., \tilde{P}^{(\mathbf{x}_M)}(n)\right)\chi^{(i)}(\mathbf{x}). \tag{2.3}$$

Алгоритм 2.1 основан на аналогах соотношений (1.12)

$$\int P(\mathbf{y})\kappa^{(\mathbf{x}_i)}(\mathbf{y})\,d\mathbf{y} \approx P(\mathbf{x}_i); \quad i = 1, ..., M, \tag{2.4}$$

которые, в свою очередь, следуют из соотношений (2.1) и (2.2).

Сравнение алгоритмов 1.1 и 2.1 дает основной вывод данной работы.

ЗАМЕЧАНИЕ 2.1. *Основанный на подходах численной сеточной аппроксимации функций «ядерный» алгоритм* 2.1 *приближения вероятностной плотности* $P(\mathbf{x})$ *конструктивно совпадает с рандомизированным проекционно-сеточным функциональным алгоритмом* 1.1 *приближения решения* $\varphi(\mathbf{x})$ *уравнения Фредгольма второго рода* (1.1)*. Поэтому для алгоритма* 1.1 *вполне целесообразно применять новое название* ***рандомизированный «ядерный» функциональный алгоритм*** *для приближения решения уравнения* (1.1)*.*

Отличие алгоритмов 1.1 и 2.1 состоит только в разности форм монте-карловских оценок для приближенного вычисления функционалов вида (1.8) и (2.4) (что, в свою очередь, связано с определенным различием приближаемых функций $\varphi(\mathbf{x})$ и $P(\mathbf{x})$). Отличие также состоит в том, что в задаче приближения плотности $P(\mathbf{x})$ выборка $\left\{\hat{\boldsymbol{\xi}}_1, ..., \hat{\boldsymbol{\xi}}_n\right\}$ считается *заданной* (при этом число выборочных значений n, как правило фиксировано и не может быть увеличено), а для функции $\varphi(\mathbf{x})$ число n моделируемых траекторий (1.13) прикладной цепи Маркова (1.11) может варьироваться.

Суть замечания 2.1 в определенной степени отражена в разделе 2 работы [10], но необходимые детали использования теории сеточной аппроксимации функций в этой работе не приводятся.

В связи с основным выводом замечания 2.1 можно сформулировать следующие соображения.

ЗАМЕЧАНИЕ 2.2. *Для «ядерного» алгоритма* 2.1 *приближения вероятностной плотности* $P(\mathbf{x})$ *можно использовать соображения теории условной оптимизации проекционно-сеточного функционального алгоритма* 1.1 *из работ* [1, 2, 6, 7].

Так, по аналогии с работами [1, 2, 6, 7] можно рекомендовать использование функций (1.15), (1.17) и аппроксимационных коэффициентов вида (1.18), т. е.

$$w^{(i)}\left(\tilde{P}^{(\mathbf{x}_1)}(n), ..., \tilde{P}^{(\mathbf{x}_M)}(n)\right) = \tilde{P}^{(\mathbf{x}_i)}(n), \tag{2.5}$$

в алгоритме 2.1. Это позволяет, в частности, получить выражения для условно-оптимальных параметров алгоритма 2.1 вида

$$M_{opt} = \left[\frac{\hat{H}_1[(2\nu+1)d+4]}{(2\nu+1)d}\right]^{d/2}\gamma^{-d/2},\ n_{opt} = \frac{\hat{H}_2^2\hat{H}_1^{d/2}\,[(2\nu+1)d+4]^{2+d/2}}{16[(2\nu+1)d]^{d/2}} \times (2\ln M_{opt} - \ln\ln M_{opt} + \hat{H}_3) \times \gamma^{-2-d/2} \tag{2.6}$$

для заданного уровня погрешности γ и специально подобранных констант $\hat{H}_1$, $\hat{H}_2$, $\hat{H}_3$ и ν.

Учитывая, что базис (1.15) является моделируемым, можно также предложить использовать нормированную функцию $L^{(M)}\tilde{P}(\mathbf{x})$ из соотношения (2.3) в качестве плотности для численного моделирования дополнительных выборочных значений $\{\boldsymbol{\xi}_j\}$, близких по распределению к значениям $\left\{\hat{\boldsymbol{\xi}}_j\right\}$ из соотношения (2.1), с использованием соответствующей версии метода суперпозиции (см. разделы 17, 18 книги [11]).

ЗАМЕЧАНИЕ 2.3. *Для развития теории конструирования и условной оптимизации проекционно-сеточного функционального алгоритма* 1.1 *можно использовать соображения из теории «ядерных» оценок вероятностных плотностей из работы* [9].

В последнем замечании речь идет прежде всего о варьировании и оптимальном выборе «ядерной» функции $\kappa^{(\mathbf{x})}(\mathbf{y})$. В работе [9] рассматриваются возможности выбора этой функции в виде

$$\kappa^{(\mathbf{x})}(\mathbf{y}) = \prod_{s=1}^{d}\frac{1}{h^{(s)}(n)}\hat{\kappa}^{(s)}\left(\frac{x^{(s)} - y^{(s)}}{h^{(s)}(n)}\right), \tag{2.7}$$

где положительные (вообще говоря, зависящие от величины выборки n) числа $h^{(s)}(n) > 0;\ s = 1, ..., d$ определяют носитель («коэффициенты размытости») «ядерной» функции $\hat{\kappa}^{(\mathbf{x})}(\mathbf{y})$, а ограниченные в совокупности четные ($\hat{\kappa}^{(s)}(y) = \hat{\kappa}^{(s)}(-y)$) функции $\hat{\kappa}^{(s)}(y)$ имеют единичный «второй момент» и конечные «m-е моменты»:

$$\int_{-\infty}^{+\infty} y^2 \hat{\kappa}^{(s)}(y)\, dy = 1, \quad \int_{-\infty}^{+\infty} y^m \hat{\kappa}^{(s)}(y)\, dy < \infty; \quad m > 2. \tag{2.8}$$

Содержательные результаты по оптимизации приближения (2.1) с «ядерной» функцией (2.7) получаются лишь для случая

$$h^{(1)}(n) = ... = h^{(d)}(n) \equiv \hat{h}(n); \quad \hat{\kappa}^{(1)}(y) = ... = \hat{\kappa}^{(d)}(y) \equiv \hat{\kappa}(y)$$

(см., в частности [9]); случай различных $\{h^{(s)}(n)\},\ \{\hat{\kappa}^{(s)}(y)\};\ s = 1, ..., d$ практически не изучен.

Для функции (1.17) имеем

$$\hat{h}(n) = h; \quad \hat{\kappa}(y) = \begin{cases} 1 & \text{при} \quad |y| \le 1/2, \\ 0 & \text{при} \quad |y| > 1/2, \end{cases} \tag{2.9}$$

т. е. «коэффициент размытости» совпадает с шагом равномерной сетки (1.16), а функция $\hat{\kappa}(y)$ является кусочно-постоянной.

В работе [9] в рамках асимптотического (при $n \to \infty$) подхода с помощью разложения функции $P\left(x^{(1)} + \hat{h}(n)y^{(1)}, ..., x^{(d)} + \hat{h}(n)y^{(d)}\right)$ по всем переменным в точке $\mathbf{x} = \left(x^{(1)}, ..., x^{(d)}\right)$ в ряд Тейлора и минимизации относительной глобальной среднеквадратической ошибки

$$\left(\hat{\delta}^{(L_2)}\right)^2 = \frac{\int \mathbf{E}\left[P(\mathbf{x}) - Z_n(\mathbf{x})\right]^2 d\mathbf{x}}{Q}; \quad Q = \int P^2(\mathbf{x})\, d\mathbf{x}$$

получено асимптотическое приближение

$$\left(\hat{\delta}^{(L_2)}\right)^2 \sim \frac{n^{-1}\hat{h}^{-d}(n) I^d + (1/4)\hat{h}^4(n) J}{Q}; \ I = \int_{-\infty}^{+\infty} \hat{\kappa}^2(y)\, dy; \ J = \int .. \int \left[\sum_{l=1}^{d} \frac{\partial^2 P\left(x^{(1)}, .., x^{(d)}\right)}{\partial \left(x^{(l)}\right)^2}\right]^2 dx^{(1)}..dx^{(d)}. \tag{2.10}$$

Минимизация полученного приближения величины $\left(\hat{\delta}^{(L_2)}\right)^2$ и выбор функции $\hat{\kappa}(y)$ из соображений минимизации величины I из (2.10) с соблюдением условий (2.8) дают оптимальное значение «коэффициента размытости» и оптимальную форму функции $\hat{\kappa}(y)$:

$$\hat{h}_{opt}(n) = \left(\frac{d \times I^d}{n \times J}\right)^{1/(d+4)}, \quad \hat{\kappa}_{opt}(y) = \begin{cases} \frac{3}{4\sqrt{5}} - \frac{3y^2}{20\sqrt{5}} & \text{при} \quad |y| \le \sqrt{5}, \\ 0 \quad \text{при} \quad |y| > \sqrt{5}. \end{cases} \tag{2.11}$$

Вопрос о том, насколько целесообразно использование параметров (2.11) в сочетании с шагом h сетки (1.6) вида (1.16) в алгоритме 2.1 (в частности, получится ли выигрыш по сравнению со случаем, когда параметр $\hat{h}(n)$ и функция $\hat{\kappa}(y)$ имеют вид (2.9)), требует отдельного подробного исследования.

Заключение

В работе показано, что при подключении теории численной аппроксимации функций можно получить «конструктивную» версию «ядерной» оценки для приближения вероятностной неизвестной плотности $P(\mathbf{x})$ по данным выборочным значениям (алгоритм 2.1). Эта конструкция аналогична рандомизированному проекционно-сеточному методу приближения решения $\varphi(\mathbf{x})$ интегрального уравнения Фредгольма второго рода (1.1) (т. е. алгоритму 1.1). Если дополнительно выбрать мультилинейный аппроксимационный базис (1.15), «ядерную» функцию (1.17) и аппроксимационные коэффициенты (2.5) (т. е. рассмотреть версию алгоритма 2.1, соответствующую многомерному аналогу метода полигона частот), то можно получить выражения (2.6) для условно-оптимальных параметров алгоритма 2.1.

Отмечено также, что можно назвать рандомизированный проекционно-сеточный алгоритм 1.1 *функциональным «ядерным» алгоритмом* и попытаться развить его теорию с помощью оптимизации выбора «ядерной» функции (2.7) подобно тому, как это сделано в работе [9].

Список литературы

[1] Михайлов Г. А., Войтишек А. В. Численное статистическое моделирование. Методы Монте-Карло. М.: Изд. центр «Академия», 2006.

[2] Войтишек А. В. Функциональные оценки метода Монте-Карло. Новосибирск: НГУ, 2007.

[3] Бахвалов Н. С. Численные методы. М.: Наука, 1975.

[4] Войтишек А. В. Классификация и возможности практического применения рандомизированных функциональных численных алгоритмов решения интегральных уравнений Фредгольма второго рода // Математический анализ. Итоги науки и техники. Серия: Современная математика и ее приложения. Тематические обзоры. 2018. Т. 155. С. 3–19.

[5] Войтишек А. В. Разработка и оптимизация рандомизированных функциональных численных методов решения практически значимых интегральных уравнений Фредгольма второго рода // Сибирский журнал индустриальной математики. 2018. Т. 21. № 2 (74). С. 32–45.

[6] Shkarupa E. V., Voytishek A. V. Convergence of discrete-stochastic numerical procedures with independent or weakly dependent estimators at grid nodes // Journal of Statistical Planning and Inference. 2000. V. 85. P. 199–211.

[7] Войтишек А. В., Головко Н. Г., Шкарупа Е. В. Оценка погрешности многомерного аналога метода полигона частот // Сибирский журнал вычислительной математики. 2002. Т. 5, № 1. С. 11–24.

[8] Марчук Г. И., Агошков В. И. Введение в проекционно-сеточные методы. М.: Наука, 1981.

[9] Епанечников В. А. Непараметрическая оценка многомерной плотности вероятности // Теория вероятностей и ее применения. 1969. Т. 14. Вып. 1. С. 156–161.

[10] Михайлов Г. А. Рандомизированные алгоритмы метода Монте-Карло для задач со случайными параметрами (метод «двойной рандомизации») // Сибирский журнал вычислительной математики. 2019. Т. 22, № 2. С. 187–200.

[11] Войтишек А. В. Дополнительные сведения о численном моделировании случайных элементов. Новосибирск: НГУ, 2007.

Булгакова Татьяна Евгеньевна — старший преподаватель Специализированного учебно-научного центра Новосибирского государственного университета;
e-mail: tatyana.bulgakova@gmail.com
Войтишек Антон Вацлавович — д.ф.-м.н., профессор, ведущий научный сотрудник Института вычислительной математики и математической геофизики СО РАН;
профессор Новосибирского государственного университета;
e-mail: vav@osmf.sscc.ru.
Дата поступления — 30 апреля 2019 г.

АЛГОРИТМЫ МЕТОДА МОНТЕ-КАРЛО ДЛЯ МОДЕЛИРОВАНИЯ АВТОТРАНСПОРТНОГО ПОТОКА СО СЛУЧАЙНЫМИ ПАРАМЕТРАМИ

А. В. Бурмистров[1,2], М. А. Коротченко[1]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.642
DOI: 10.24411/9999-016A-2019-10012

В данной работе авторами предложены новые алгоритмы метода Монте-Карло для численной оценки вероятностных моментов линейных функционалов от решения уравнения больцмановского типа со случайными параметрами. Уравнение такого типа возникает в кинетической модели автотранспортного потока с выделенными ускорениями. Для возможности учета разных типов взаимодействующих автомобилей, а также случайных параметров, которые описывают навыки и поведение водителя конкретного автомобиля, модель усовершенствована в рамках подхода, ранее разработанного авторами.

Ключевые слова: кинетическая модель, весовое моделирование цепи Маркова, двойная рандомизация.

Введение

Настоящая работа изучает влияние ряда параметров, в том числе случайных, на моделирование автотранспортного потока (АТП) в рамках кинетической модели с выделенными ускорениями [1]. Актуальность данного исследования обусловлена постоянным увеличением объема транспортных и пассажирских перевозок, связанных, в том числе, с увеличением количества личных и общественных транспортных средств. Это приводит к перегруженности дорог, большим пробкам, увеличению количества аварий и, как следствие, к необходимости оптимизации транспортной системы с учетом закономерностей её функционирования.

На протяжении нескольких десятилетий исследованием транспортных систем занимались ученые из различных областей науки и сегодня имеется обширная литература по теме изучения и моделирования АТП. Представленные в литературе модели можно разбить на три класса (подробнее обзор см. в [1]): микроскопические, мезоскопические и макроскопические. Кроме того, различают детерминированные и стохастические модели. Кинетические модели являются мезоскопическими: они оперируют информацией об индивидуальном автомобиле, а именно о его параметрах на микро уровне. В результате моделирования модели этого класса дают возможность получить данные о поведении АТП в целом, то есть о макрохарактеристиках потока. Такие модели рассматривают поток автомобилей как газ, состоящий из взаимодействующих частиц, в котором каждая частица соответствует автомобилю. Уравнения, используемые в моделях данного типа, аналогичны уравнениям газовой кинетики, в частности, уравнению Больцмана. Однако, в отличие от уравнения Больцмана, в случае АТП нет аналогов законам сохранения момента и энергии при парных взаимодействиях.

В кинетических моделях выделяют два основных типа взаимодействий: ускорение и торможение. Кроме того, возможность обгона обычно вводят с помощью вероятности, которая зависит от плотности автомобилей на многополосной дороге. Выбор между перечисленными видами взаимодействий зависит как от абсолютных, так и от относительных величин скорости и ускорения взаимодействующих автомобилей.

В выбранной авторами модели изучается связь количества автомобилей (аналог концентрации вещества в газовой динамике) и частоты изменения ускорений автомобилями (аналог скорости реакции). Подобно

Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2019-0002) и при частичной финансовой поддержке Российского фонда фундаментальных исследований (проекты 17-01-00698, 18-01-00356, 18-01-00599).

ISBN 978-5-901548-42-4

взаимодействиям молекул с парными скачкообразными изменениями скоростей в газовой динамике, в описываемой модели использование водителями педалей управления рассматривается как акт кинетических парных взаимодействий автомобилей, приводящих к скачкообразным изменениям ускорений.

1 Кинетическая модель с выделенным ускорением

Одним из основных предположений в кинетических моделях является возможность учитывать только парные взаимодействия частиц. Другими словами, среднее время между двумя взаимодействиями автомобилей должно быть намного больше времени самого взаимодействия. Если результатом взаимодействия автомобилей становится скачкообразное изменение скорости (как в газовой динамике), то предположение о парных взаимодействиях остается обоснованным только при достаточно неплотных потоках, что подтверждается наблюдениями. Поэтому в [2] была предложена модель, в которой ускорение внесено в число фазовых координат, описывающих состояние автомобиля (наряду с пространственной координатой и скоростью, которые традиционно используются в кинетических моделях), т. к. именно эта координата скачкообразно меняется при взаимодействии автомобилей. Такая модификация фазового пространства дает возможность распространить кинетическую модель на более широкий класс АТП (полученная модель достаточно адекватно описывает не только случай частично затрудненного движения, но и более плотные АТП) и была обусловлена несколькими причинами. Во-первых, как показали измерения, относительное изменение ускорения намного меньше, чем у других процессов в потоке (скорости, относительного расстояния и др.). Во-вторых, ускорение — это самая естественная переменная, которой управляет каждый водитель путем нажатия на педали управления своим транспортным средством. Кроме того, существенное предположение об "автомобильном хаосе" дополнительно подтверждается тем, что изменения ускорения зависят не только от типа автомобиля, но также от поведения и навыков каждого конкретного водителя, что вносит дополнительную случайность в модель. Это предположение используется в кинетических моделях АТП и является аналогичным предположению о "молекулярном хаосе" в газовой динамике.

В пространственно-однородном случае в рамках модели, предложенной в [2], одночастичная плотность $f \equiv f(a, v, t)$ распределения автомобилей с ускорением a и скоростью v удовлетворяет интегро-дифференциальному уравнению Больцмановского типа

$$\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial v} = \int\limits_{\bar{a},\bar{v},a'} \left[\Sigma(a' \to a|v, \bar{a}, \bar{v}, \mathbf{m}_f(t))f(a', v, t) - \Sigma(a \to a'|v, \bar{a}, \bar{v}, \mathbf{m}_f(t))f\right] f(\bar{a}, \bar{v}, t)\, \mathtt{d}\bar{a}\, \mathtt{d}\bar{v}\, \mathtt{d}a'. \qquad (1)$$

Здесь $\bar{a}$ и $\bar{v}$ — соответственно, ускорение и скорость автомобиля, находящегося непосредственно перед "текущим" автомобилем. Между этими двумя автомобилями и происходит взаимодействие. Далее будем называть автомобиль с координатами $(\bar{a}, \bar{v})$ лидером, а следующий за ним — не-лидером. Заданная функция $\Sigma(a' \to a|v, \bar{a}, \bar{v}, \mathbf{m}_f(t))$ – взвешенная плотность взаимодействий, которая в общем случае имеет следующий вид

$$\Sigma(a' \to a|v, \bar{a}, \bar{v}, \mathbf{m}_f(t)) = \int\limits_{h_{\min}}^{\infty} \sigma(a' \to a|h, v, \bar{a}, \bar{v}) \cdot Q(h, a', v, \bar{a}, \bar{v}) \cdot \mathcal{D}(h|a', v, \mathbf{m}_f(t), \mathcal{K})\, \mathtt{d}h. \qquad (2)$$

В выражении (2) использованы следующие обозначения:

$\mathbf{m}_f(t)$ — некоторый вектор моментов решения $f(a, v, t)$, таких как, например, средние скорость и ускорение, а также их средние квадраты, определяющие разброс этих величин;

$h_{\min}$ — наименьшее возможное расстояние между автомобилями в состоянии покоя, то есть средняя длина автомобиля (размерность m);

$\sigma(a' \to a|h, v, \bar{a}, \bar{v})$ — плотность распределения нового ускорения не-лидера при условии, что имело место взаимодействие между автомобилями с координатами (a', v) и $(\bar{a}, \bar{v})$, и которые находятся на расстоянии h друг от друга (размерность s^2m^{-1});

$Q(h, a', v, \bar{a}, \bar{v})$ — скорость или частота взаимодействий, которая зависит от текущего состояния пары взаимодействующих автомобилей и от расстояния между ними (размерность s^{-1});

$\mathcal{D}(h|a', v, \mathbf{m}_f(t), \mathcal{K})$ — условная плотность распределения расстояния h, зависящая от состояния (a', v) не-лидера, вектора $\mathbf{m}_f(t)$, а также от плотности автомобилей $\mathcal{K}$ (размерность m^{-1}).

Уравнение (1) нужно дополнить некоторым заданным начальным условием $f(a,v,0)=f_0(a,v)$ и краевыми условиями, гарантирующими отсутствие отрицательных скоростей и скоростей, превышающих некоторую разрешенную или возможную максимальную скорость.

Отметим, что в результате введения ускорения в число фазовых координат, в отличие от газовой динамики, в данной модели при парных взаимодействиях скачкообразно меняется не скорость, а ускорение автомобиля. Кроме того, при взаимодействии двух автомобилей лидер не меняет своего ускорения, что выражено в несимметричности функции (2). Поэтому взаимодействующие автомобили можно представлять упорядоченными парами $\pi=(i,j)$. Для определенности полагаем, что первый индекс i — номер не-лидера, а второй индекс j — номер лидера.

2 Интегральное уравнение

Рассмотрим многочастичную (а именно N-частичную) систему, т. е. систему из N автомобилей. Обозначим векторы, состоящие из набора ускорений и скоростей всех автомобилей в заданной системе через $A=(a_1,\ldots,a_N)$ и $V=(v_1,\ldots,v_N)$, соответственно. Рассмотрим также плотность распределения $F\equiv F(Z,t)$ взаимодействий в системе с фазовыми координатами $(Z,t)\equiv(\pi,A,V,t)$. Ранее авторами на основе N-частичного кинетического уравнения было показано, что F является решением следующего интегрального уравнения второго рода (см. подробнее [1]):

$$F(Z,t)=\int\limits_0^t\int F(Z',t')K(Z',t'\to Z,t)\,\mathtt{d}Z'\,\mathtt{d}t'+\delta(t)P_0(A,V)\delta(\pi_0),\tag{3}$$

или в операторном виде $F=\mathbf{K}F+F_0$. Здесь $\delta(\cdot)$ — дельта-функция Дирака, $P_0(\cdot)$ — некоторое заданное начальное распределение ускорений и скоростей, $\mathtt{d}Z=\mathtt{d}A\,\mathtt{d}V\,\mathtt{d}\mu(\pi)$, а интегрирование по мере μ означает суммирование по всем упорядоченным парам $\pi=(i,j)$. При этом, решение F тесно связано с решением $f(a,v,t)$ исходного уравнения (1), а ядро оператора $\mathbf{K}$ описывает динамику системы автомобилей в соответствии со своей мультипликативной структурой [1]:

$$K(Z',t'\to Z,t)=K_1(t'\to t|A',V')K_2(V'\to V|A',t-t')K_3(\pi)K_4(a_i'\to a_i|\pi,V)\cdot\left\{\prod_{m\neq i,m=1}^{N}\delta(a_m'-a_m)\right\}.\tag{4}$$

Таким образом, ядро $K(Z',t'\to Z,t)$ является произведением плотностей перехода. Наличие построенного интегрального уравнения (3), а также связанного с ним марковского процесса динамики системы автомобилей, позволяет нам использовать хорошо развитый аппарат весовых и рандомизированных оценок метода Монте-Карло [3]. В том числе и метод мажорантной частоты, который обеспечивает линейную зависимость времени расчетов от числа тестовых частиц N [4]. Кроме того, это дает возможность исследовать зависимость модели от различных параметров, оценивать параметрические производные (см., например, [1]), а также уменьшать трудоемкость статистических алгоритмов с помощью ценностного моделирования.

2.1 Оценки функционалов

Обычно при решении уравнения (1) требуется оценить линейные функционалы $I_{\mathbf{h}}(T)$ от одночастичной функции распределения f следующего вида $I_{\mathbf{h}}(T)=\int\int\mathbf{h}(a,v)f(a,v,T)\,\mathtt{d}a\,\mathtt{d}v=\int\mathbf{h}(a_1,v_1)P(A,V,T)\,\mathtt{d}A\,\mathtt{d}V$. Здесь T — некоторый фиксированный момент времени, $P(A,V,T)$ — плотность распределения N-частичной системы. По аналогии с [5] можно показать, что

$$I_{\mathbf{h}}(T)=\int\limits_{\mathbb{Z}}\int\limits_0^T\mathbf{H}(A,V+A(T-t'))\mathrm{e}^{\left\{-\int\limits_{t'}^{T}\nu(A,V+A(\tau-t'))\,\mathtt{d}\tau\right\}}F(Z,t')\,\mathtt{d}Z\,\mathtt{d}t'=\int\limits_{\mathbb{Z}}\int\limits_0^T\tilde{\mathbf{H}}(A,V,T-t')F(Z,t')\,\mathtt{d}Z\,\mathtt{d}t',$$

где $\mathbf{H}(A,V)=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\mathbf{h}(a_i,v_i)$. В результате получаем выражение для искомого линейного функционала от решения интегро-дифференциального уравнения (1) в виде функционала от решения интегрального уравнения (3): $I_{\mathbf{h}}(T)=(\tilde{\mathbf{H}},F)$.

Например, для приближения распределения скоростей и ускорений, в качестве функций $\mathbf{h}(\cdot)$ следует выбирать индикаторы некоторого разбиения соответствующих областей (по скоростям или ускорениям). Весовые функции для оценки средней скорости и среднеквадратичного рассеяния скорости автомобилей имеют вид $\mathbf{h}(a,v)=v$ и $\mathbf{h}(a,v)=v^2$, соответственно.

2.2 Моделирование цепи Маркова

Для приближенного вычисления функционала $I_{\mathbf{h}}(T)$ можно использовать оценки по столкновениям ξ или по поглощениям η – случайные величины, являющиеся функционалами от траекторий моделируемой цепи Маркова $\{Z_n,t_n\}$ (см., например, [3,6]): $I_{\mathbf{h}}(T)=\mathbf{E}\xi=\mathbf{E}\eta$, при этом

$$\xi=\sum_{n=0}^{\kappa}\tilde{\mathbf{H}}(A_n,V_n,T-t_n),\ \ \eta=\frac{\tilde{\mathbf{H}}(A_\kappa,V_\kappa,T-t_\kappa)}{q(A_\kappa,V_\kappa,t_\kappa)},\ \text{где } q(A,V,t')=1-\int\limits_0^{T-t'}K_1(\tau|A,V,t')\,\mathtt{d}\tau.$$

Здесь κ — номер взаимодействия, непосредственно предшествующего выходу системы за рассматриваемую временну́ю границу T. Каждый переход при моделировании цепи Маркова, которая связана с интегральным уравнением (3), состоит из четырех элементарных переходов в следующем порядке, который определяется формой ядра интегрального оператора (4):

1. момент времени t следующего взаимодействия в системе выбирается в соответствии с экспоненциальной плотностью (здесь и далее $\Theta(\cdot)$ — единичная ступенчатая функция Хевисайда)

$$K_1(t'\to t|A',V')=\Theta(t-t')\nu(A',V'+A'(t-t'))\exp\left\{-\int\limits_{t'}^{t}\nu(A',V'+A'(\tau-t'))\,\mathtt{d}\tau\right\},$$

$$\text{где функция } \nu(A,V)=\sum_{\pi}\frac{\nu_{(i,j)}}{N-1},\ \text{при этом } \nu_{(i,j)}=\int\Sigma(a_i\to a|v_i,v_j,a_j)\,\mathtt{d}a;$$

2. скорости всех автомобилей в момент времени t перевычисляются в соответствии с дельта-плотностью $K_2(V'\to V|A',t-t')=\delta(V-V'-A'(t-t'))$;

3. упорядоченная пара взаимодействующих автомобилей (i,j) выбирается в соответствии с вероятностями

$$K_3(i,j)=\frac{1}{N-1}\cdot\frac{\nu_{(i,j)}}{\nu(A',V)};$$

4. новое ускорение автомобиля с номером i (т.е. автомобиля, у которого лидером является автомобиль с номером j) изменяется в соответствии с плотностью $K_4(a_i'\to a_i|\pi,V)=\Sigma(a_i'\to a_i|v_i,a_j,v_j)/\nu_{(i,j)}$.

Отметим, что часто более эффективно задействовать вспомогательную цепь Маркова $\{Z_n,t_n\}$, $n=0,1...,\kappa$, с нормированной плотностью перехода $P(Z',t'\to Z,t)$ (см. подробнее [6]), которая имеет те же особенности, что и плотность (4). В этом случае дополнительно вычисляются веса $Q_n=Q_{n-1}\left\{\dfrac{K(Z',t'\to Z,t)}{P(Z',t'\to Z,t)}\right\}$ (при этом $Q_0\equiv 1$) и используются весовые варианты оценок ξ и η:

$$\xi=\sum_{n=0}^{\kappa}Q_n\tilde{\mathbf{H}}(A_n,V_n,T-t_n),\ \ \eta=\frac{Q_\kappa\tilde{\mathbf{H}}(A_\kappa,V_\kappa,T-t_\kappa)}{q(A_\kappa,V_\kappa,t_\kappa)},\ \text{где } q(A,V,t')=1-\int\limits_0^{T-t'}P_1(\tau|A,V,t')\,\mathtt{d}\tau.$$

Здесь P_1 — моделируемая плотность распределения времени следующего взаимодействия в системе.

3 Индивидуальные особенности и случайные параметры

Несмотря на наличие общих правил дорожного движения, множество типов автомобилей, составляющих АТП, обладают индивидуальными техническими характеристиками, а также управляются водителями с

большим разнообразием навыков и манер вождения. Это обстоятельство мотивирует нас перейти от моделей с набором некоторых осредненных параметров к моделям, в которых параметры каждого автомобиля отличаются внутри системы и, более того, имеют вероятностную природу. Кроме того, в адекватной модели должны учитываться как влияние случайных факторов (таких как дорожно-транспортные происшествия, ремонтные работы, погодные условия, качество дорожного покрытия и др.), так и флуктуации, связанные с сезонностью, днями недели и пр.

Постановка задачи. Пусть коэффициенты уравнения (3) зависят от случайного векторного параметра ψ. Рассмотрим задачу оценки математического ожидания функционалов $I_{\mathbf{h}_1}(T,\psi)$ и $I_{\mathbf{h}_1}(T,\psi)I_{\mathbf{h}_2}(T,\psi)$ с соответствующими весовыми функциями: $I_{(1)}(T) = \mathbf{E}_{(\psi)}I_{\mathbf{h}_1}(T,\psi)$ и $I_{(1,2)}(T) = \mathbf{E}_{(\psi)}[I_{\mathbf{h}_1}(T,\psi)I_{\mathbf{h}_2}(T,\psi)]$. Принцип построения новых алгоритмов решения автотранспортных задач предложен авторами в рамках подхода решения кинетических задач [7, 8], в том числе со случайными коэффициентами [9, 10] и параметрами [11].

3.1 Простая частота взаимодействий

В качестве первого примера для введения случайных параметром рассмотрим модель взаимодействия, предложенную в [12] и рассмотренную нами в простейшем случае в [1]. Частота взаимодействия в этой модели задается некоторым распределением: $Q \sim \mathcal{D}_{freq}$. С этой частотой Q водитель автомобиля не-лидера сравнивает расстояние h до лидера со значением некоторой пороговой функции $H(v)$. В случае относительно большого расстояния ($h > H(v)$) не-лидер приобретает ускорение, распределенное с некоторой плотностью $\mathcal{D}_{acc}$, в противном случае ($h \le H(v)$) не-лидер приобретает отрицательное ускорение (тормозит), задаваемое другой плотностью $\mathcal{D}_{br}$. Таким образом, плотность $\sigma(\cdot)$ имеет вид:

$$\sigma(a' \to a|h,v,\bar{a},\bar{v}) \equiv \sigma(a|h,v) = \Theta(h - H(v)) \cdot \mathcal{D}_{acc}(a) + \Theta(H(v) - h) \cdot \mathcal{D}_{br}(a).$$

где $\Theta(\cdot)$ — функция Хевисайда. При этом в простейшем случае, учитывающем различие в автомобилях и водителях, используются дельта-функции: $\mathcal{D}_{acc}(a) = \delta(a - a_i^+)$ и $\mathcal{D}_{br}(a) = \delta(a + a_i^-)$. Пороговые функции также могут отличаться параметрами для учета индивидуальных особенностей водителей, а именно: $H_i(v) = \alpha_i \cdot v + h_{\min}$, при этом α_i также могут принимать случайные значения. При этом бóльшее значение параметра α_i соответствует более спокойному (консервативному) стилю вождения, а меньшее — более агрессивному.

Взвешенная плотность взаимодействий (2) при фиксированной на каждой траектории частоте Q будет иметь рандомизированный вид:

$$\Sigma(a_i' \to a_i|v_i,a_j,v_j) = p \cdot Q \cdot \mathcal{D}_{br}(a_i) + (1-p) \cdot Q \cdot \mathcal{D}_{acc}(a_i), \text{ где } p = \mathbb{P}(h < H(v_i)) = \int\limits_{h_{\min}}^{H(v_i)} \mathcal{D}(h)\, \mathtt{d}h.$$

В качестве функции $\mathcal{D}(\cdot)$ при небольших плотностях автомобилей хорошо подходит экспоненциальное распределение $\mathcal{D}(h) = \frac{1}{\bar{H}-h_{\min}} \exp\left\{-\frac{h-h_{\min}}{\bar{H}-h_{\min}}\right\} \Theta(h - h_{\min})$, здесь $\bar{H}$ — среднее расстояние между автомобилями во взаимодействующей паре, то есть $\bar{H} = 1/\mathcal{K}$. При этом $\nu_{(i,j)} \equiv Q$.

3.2 Пространственно-однородный поток

Рассмотрим пространственно-однородный (без зависимости от расстояний) почти свободный стационарный АТП, в котором скорости всех автомобилей распределены около некоторой средней скорости $\mathrm{V} \gg 0$. В таком потоке, при достаточно больших расстояниях между взаимодействующими автомобилями, измерения показывают сильную зависимость частоты взаимодействия Q от относительной скорости $v_r \equiv \bar{v} - v$ автомобилей в паре $Q(h,a',v,\bar{a},\bar{v}) = Q(|v_r|)$. Приблизим ее рядом Тейлора до первого порядка по v_r:

$$Q(|v_r|) \approx Q_0 + r_0|v_r|. \tag{5}$$

Кроме того, в таком почти свободном потоке каждое взаимодействие вносит только небольшое, корректирующее изменение в ускорение не-лидера. В этом случае функцию $\sigma(\cdot)$ можно представить в виде двух распределений, в зависимости от знака относительной скорости, то есть

$$\sigma(a' \to a|h,v,\bar{a},\bar{v}) \approx \sigma(a|\bar{v}-v) = \Theta(v_r) \cdot \mathcal{D}_{acc}(a) + \Theta(-v_r) \cdot \mathcal{D}_{br}(a).$$

В этом случае из выражения (2) получаем, что плотность взаимодействий имеет следующий вид:

$$\Sigma(a_i' \to a_i|v_i,a_j,v_j) = (Q_0 + r_0|v_i - v_j|)\left[\Theta(v_j - v_i) \cdot \mathcal{D}_{acc}(a_i) + \Theta(v_i - v_j) \cdot \mathcal{D}_{br}(a_i)\right].$$

При этом $\nu_{(i,j)} \equiv Q_0 + r_0|v_i - v_j| \leq Q_0 + r_0 V_{\max}$ и при моделировании целесообразно использовать метод мажорантной частоты [1]. Заметим, что для крайних случаев значений (5) существуют аналоги в газовой динамике: максвелловское взаимодействие ($r_0 = 0$) и взаимодействие типа твердых сфер ($Q_0 = 0$). Более того, для этих случаев известны аналитические решения исходного уравнения (1) в простейшем случае, когда $\mathcal{D}_{acc}(a) = \delta(a - a_0)$ и $\mathcal{D}_{br}(a) = \delta(a + a_0)$ (см., например, [2]). Для бóльшей рандомизации модели можно использовать случайные параметры r_0 и Q_0.

3.3 Взаимодействие на порогах

Далее рассмотрим модель, в которой профиль взаимодействия (закон изменения ускорения водителями) зависит от расстояния между лидером и не-лидером, а также от их относительной скорости. Такой профиль делает транспортный поток более однородным (см. подробнее в [13] и список литературы там же). В данной модели не-лидер с координатами (v, a') взаимодействует с лидером (и таким образом изменяет свое ускорение a'), если расстояние h до лидера совпадает со значением одной из двух пороговых функций $H_1(v) < H_2(v)$. Частота взаимодействий Q в этом случае выражается следующим образом

$$Q(h, a', v, \bar{a}, \bar{v}) = |S_1(\bar{v}, v, a')| \cdot \delta(h - H_1(v)) + |S_2(\bar{v}, v, a')| \cdot \delta(h - H_2(v)),$$

где $S_k(\bar{v}, v, a') = \left[\bar{v} - v - \frac{dH_k}{dv}(v) \cdot a'\right]$, $k = 1, 2$. Для простоты предположим, что первая пороговая функция $H_1(v) = \alpha_1 \cdot v + h_{\min}$ зависит от скорости линейно, так как данные измерений для этой функции имеют большой разброс. Для аппроксимации бóльшего порога $H_2(v)$ с учетом данных, полученных при измерениях мы используем функцию квадратного корня: $H_2(v) = \alpha_2\sqrt{v + \beta} + \gamma$. Заметим, что в общем случае величины α_1, α_2, β, γ описывают индивидуальные особенности не-лидера, поэтому могут зависеть от номера i, а также могут быть случайными с некоторым заданным распределением.

В момент пересечения не-лидером одного из порогов происходит его взаимодействие с лидером, в результате которого ускорение не-лидера изменяется. Таким образом плотность распределения $\sigma(\cdot)$ нового ускорения необходимо определить только на заданных порогах. При этом на первом пороге следует учитывать знак функции $S(\bar{v}, v, a')$, который отвечает за увеличение ($S > 0$) или уменьшение ($S \leq 0$) расстояния между взаимодействующими автомобилями: $\sigma(a' \to a|H_1(v), v, \bar{a}, \bar{v}) = \Theta(S_1) \cdot \mathcal{D}^1_{acc}(a) + \Theta(-S_1) \cdot \mathcal{D}^1_{br}(a)$.

Также, помимо знака функции $S(\cdot)$, на втором пороге следует принимать во внимание знак относительной скорости взаимодействующих автомобилей v_r:

$$\sigma(a' \to a|H_2(v), v, \bar{a}, \bar{v}) = \Theta(v_r)[\Theta(S_2) \cdot \mathcal{D}^{2+}_{acc}(a) + \Theta(-S_2) \cdot \mathcal{D}^{2+}_{br}(a)] + \Theta(-v_r)[\Theta(-S_2) \cdot \mathcal{D}^{2-}_{br}(a) + \Theta(S_2) \cdot \mathcal{D}^{2-}_{acc}(a)].$$

Заметим, что данные измерений расстояний h между автомобилями в потоке часто аппроксимируют гамма распределениями. При этом, в основе выбора плотности распределения $D(h)$ лежит однозначное соответствие между текущим ускорением a' и порогом H_i, на котором происходит скачок $a' \to a$ (см. [13]).

4 Метод двойной рандомизации

Для решения поставленной задачи со случайным векторным параметром ψ будем использовать метод "двойной рандомизации" (см., например, [3]), который строится на основе следующих равенств (здесь для примера использована оценка по поглощениям η):

$$I_{(1)}(T) = \mathbf{E}_{(\psi)} I_{\mathbf{h}}(T, \psi) = \mathbf{E}_{(\psi)} \mathbf{E}_{\omega} \eta(T, \psi, \omega) = \mathbf{E}_{(\psi,\omega)} \eta(T, \psi, \omega);$$

$$I_{(1,2)}(T) = \mathbf{E}_{(\psi)}[I_{\mathbf{h}_1}(T, \psi) I_{\mathbf{h}_2}(T, \psi)] = \mathbf{E}_{(\psi,\omega_1,\omega_2)} \left[\eta_1(T, \psi, \omega_1)\eta_2(T, \psi, \omega_2)\right],$$

где ω_1 и ω_2 - условно-независимые траектории, которые моделируются для одной выборочной реализации ψ, индекс математического ожидания $\mathbf{E}$ означает соответствующее распределение. Последние равенства показывают, что для численной оценки функционалов $I_{(1)}(T)$ и $I_{(1,2)}(T)$ достаточно моделировать только одну траекторию и две условно-независимые траектории для каждого фиксированного набора (ω_1, ω_2), соответственно. При этом отметим, что для выполнения вышеприведенных равенств необходимо предполагать существование используемых в них математических ожиданий.

Заключение

Авторами в рамках кинетической модели автотранспортных потоков предложены новые алгоритмы метода Монте-Карло для оценки вероятностных моментов линейных функционалов от решения уравнений типа Больцмана. Кинетическая модель автотранспортного потока достаточно точно описывает динамику плотного движения автомобилей в силу включения ускорения в число фазовых координат модели. Именно поэтому решение уравнений больцмановского типа, возникающих в этой модели, является актуальной проблемой. С целью учета, с одной стороны, разнообразия типов автомобилей, и, с другой стороны, индивидуальных особенностей отдельных водителей, модель была усовершенствована введением параметров, в том числе случайных. Для оценки соответствующих функционалов от решения возникающего интегрального уравнения предложено использовать метод двойной рандомизации, алгоритмы которого эффективно распараллеливаются.

Авторы признательны организаторам Международной конференции “Актуальные проблемы вычислительной и прикладной математики 2019”, в том числе Новосибирскому государственному университету.

Список литературы

[1] Burmistrov A.V., Korotchenko M.A. Application of Statistical Methods for the Study of Kinetic Model of Traffic Flow with Separated Accelerations // Russian Journal of Numerical Analysis and Mathematical Modelling. 2011. Vol. 26. N. 3. P. 275–293.

[2] Waldeer K.T. The Direct Simulation Monte Carlo Method Applied to a Boltzmann-like Vehicular Traffic Flow Model // Computer Physics Communications. 2003. V. 156. No. 1. P. 1–12.

[3] Михайлов Г.А., Войтишек А.В. Численное статистическое моделирование (Метод Монте-Карло). Москва: Издательский центр “Академия”, 2006.

[4] Иванов М.С., Рогазинский С.В. Метод прямого статистического моделирования в динамике разреженного газа. Новосибирск : ВЦ СО АН СССР, 1988.

[5] Михайлов Г.А., Рогазинский С.В. Весовые методы Монте-Карло для приближенного решения нелинейного уравнения Больцмана // Сиб. мат. журн. 2002. Т. 43, № 3. С. 620–628.

[6] Burmistrov A., Korotchenko M. Weight Monte Carlo Method Applied to Acceleration Oriented Traffic Flow Model // Springer Proceedings in Mathematics and Statistics. 2012. Vol. 23. P. 297–311.

[7] Burmistrov A.V., Korotchenko M.A. Application of Kinetic Models and Simulation of Multi-particle Systems in Some Fields of Science // Математическое моделирование, компьютерный и натурный эксперимент в естественных науках. 2019. № 1. С. 4–13.

[8] Коротченко М.А. Построение статистических оценок с помощью моделирования динамики цен в рамках кинетической модели ценообразования // В сборнике: Математическое и компьютерное моделирование естественно-научных и социальных проблем под ред. д.ф.-м.н., проф. И.В. Бойкова. 2018. С. 174–179.

[9] Бурмистров А.В., Новиков А.В. Определение коэффициентов в стохастической дифференциальной модели формирования цены // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2018. № 3. С. 39–44.

[10] Коротченко М.А., Бурмистров А.В. Алгоритмы метода Монте-Карло для решения уравнения Смолуховского со случайными коэффициентами // В сборнике: Математическое и компьютерное моделирование естественно-научных и социальных проблем под ред. д.ф.-м.н., проф. И.В. Бойкова. 2019. С. 35–40.

[11] Бурмистров А.В. Коротченко М.А. Моделирование автотранспортного потока в кинетической модели со случайными параметрами // Тезисы XIX Всероссийской конференции молодых учёных по математическому моделированию и информационным технологиям. 2018. С. 12–13.

[12] Waldeer K.T. A vehicular traffic flow model based on a stochastic acceleration process // Transport Theory and Statistical Physics. 2004. Vol. 33. No. 1. P. 7–30.

[13] Burmistrov A., Korotchenko M. Monte Carlo Algorithm for Simulation of the Vehicular Traffic Flow Within the Kinetic Model with Velocity Dependent Thresholds // Springer Proceedings in Mathematics and Statistics. 2014. Vol. 114. P. 109–117.

Бурмистров Александр Васильевич — к.ф.-м.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН (ИВМиМГ СО РАН); ст. преп. Новосибирского государственного университета; e-mail: burm@osmf.sscc.ru
Коротченко Мария Андреевна — к.ф.-м.н., науч.сотр. ИВМиМГ СО РАН; e-mail: kmaria@osmf.sscc.ru
Дата поступления — 1 июня 2019 г.

ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ ДЛЯ РЕШЕНИЯ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ, ОПИСЫВАЮЩИХ ПРОЦЕСС ОКИСЛЕНИЯ ИЗОПРОПИЛБЕНЗОЛА

М. К. Вовденко[1], С. А. Габитов[1,2], И. М. Губайдуллин[1,2]

[1] *Уфимский Федеральный исследовательский центр РАН Институт нефтехимии и катализа, 450075, Уфа*
[2] *ФГБОУ ВО Уфимский государственный нефтяной технический университет, 450062, Уфа*

УДК 004.942
DOI: 10.24411/9999-016A-2019-100 13

Окисление изопропилбензола (кумола) кислородом воздуха является одной из промежуточных стадий получения фенола и ацетона в кумольном методе, самым распространенном методе получения данных веществ на сегодняшний день. Реакция протекает по радикально-цепному механизму, при этом наряду с целевыми продуктами получаются нежелательные побочные компоненты. Различные авторы указывают разные направления развития цепей реакции, механизмы инициирования и степень влияния тех или иных факторов на ход реакции.
Ключевые слова: кумол, жесткая задача, метод Гира, метод имитации отжига

Введение

Процесс окисления изопропилбензола (ИПБ) кислородом воздуха является промежуточной стадией технологического процесса получения фенола и ацетона в т.н. кумольным способом. На сегодняшний день более 90% производимого в мире фенола и ацетона получают данным способом [1]. В ходе данной стадии происходит химическое превращение изопропилбензола в гидроперекись изопропилбензола (ГП ИПБ), которая впоследствии распадается на фенол и ацетон на следующей технологической стадии.

1 Модель

Брутто уравнение целевой реакции окисления имеет вид (1).

$$RH + O_2 \rightarrow ROOH \tag{1}$$

где R — кумил-радикал ($C_6H_5CH(CH_3)_2$) [1].

Однако фактически реакция окисления является радикально-цепной, в ходе которой могут проходить одновременно десятки стадий [2, 3]. На основании анализа работ, посвящённых изучению реакционного механизма окисления ИПБ, составим реакционную схемы окисления ИПБ, представленную ниже, в таблице 1 [2–4].

Ключевые вещества в таблице: RH — изопропилбензол, ROH — диметилфенилкарбинол, R'O — ацетофенон.

Реакции (1)-(5) относятся к реакциями инициирования, в ходе которых образуются радикалы, которые в последствии зарождают основную реакционную цепь. Существуют ращличные механизмы инициирования. В данной реакционной схеме представлены как реакции инициирования через распад гидроперекиси на радикалы (3)-(6) [1–3], так и инициирование без участия гидроперекиси [2].

Реакции (6)-(7) относятся к основной реакционной цепи, которая приводит к образованию основного количество гидропероксида. Практически во всех исследованиях реакционного механизма данные стадии не меняются [2, 3].

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00015.

ISBN 978-5-901548-42-4

Таблица 1: Реакционная схема окисления ИПБ

№	Тип реакции	Реакция
1	Инициирование	$RH + O_2 \rightarrow R\bullet + HO_2\bullet$
2	Инициирование	$RH + HO_2\bullet \rightarrow R\bullet + H_2O_2$
3	Инициирование	$ROOH \rightarrow RO\bullet + HO\bullet$
4	Инициирование	$RO\bullet + RH \rightarrow R\bullet + ROH$
5	Инициирование	$\bullet OH + RH \rightarrow R\bullet + H_2O$
6	Основная цепь	$R\bullet + O_2 \rightarrow RO_2\bullet$
7	Основная цепь	$RO_2\bullet + RH \rightarrow R\bullet + ROOH$
8	Побочная	$RO\bullet \rightarrow R'O + CH_3\bullet$
9	Побочная	$2RO_2\bullet \rightarrow 2R'O + 2CH_3\bullet + O_2$
10	Побочная	$CH_3\bullet + O_2 \rightarrow CH_3O_2\bullet$
11	Побочная	$CH_3O_2\bullet + RH \rightarrow CH_3OOH + R\bullet$
12	Побочная	$CH_3OOH \rightarrow H_2CO + OH_2$
13	Побочная	$CH_3O_2\bullet + RO_2\bullet \rightarrow ROH + HCOH + O_2$
14	Побочная	$2HCOH + O_2 \rightarrow 2HCOOH$
15	Побочная	$2R\bullet \rightarrow R\text{-}R$
16	Побочная	$R\bullet + RO_2\bullet \rightarrow ROOR$
17	Побочная	$2RO_2\bullet \rightarrow ROOR + O_2$
18	Побочная	$2RO_2\bullet \rightarrow 2RO\bullet + O_2$
19	Побочная	$RO\bullet + ROOH \rightarrow ROH + RO_2\bullet$
20	Побочная	$RO\bullet + RH \rightarrow \alpha\text{-MS} + H_2O + R\bullet$
21	Побочная	$2ROOH \rightarrow RO_2\bullet + RO\bullet + H_2O$

Теоретически основная реакционная цепь может продолжаться бесконечно. На практике, как правило, основная реакционная цепь обрывается с образованием побочных продуктов, или же приводит к возникновению побочной цепи. В ходе окисления может образовываться большое количество различных типов побочных продуктов, и в различных исследованиях представлены различные типы реакций приводящие к их образованию [1–4]. Для нашей реакционной схемы мы выбрали прежде всего реакции, приводящие к образованию ацетофенона (8), диметилфенилкарбинола (13,19), органических кислот (10-14), перекиси кумола (16,17), альфаметилстирола (20), а также распад гидроперекиси изопропилбензола (21) [1–4].

Для идентификации математической модели, описывающей реакцию, на основе известных данных о поведении многокомпонентной смеси во времени [2, 3, 5] необходимо решить обратную задачу химической кинетики [6], математически представляющуюся задачей глобальной условной оптимизации. Необходимо минимизировать критерий отклонения (2) расчетных данных концентраций от экспериментальных:

$$F = \sum_{i=1}^{N} \sum_{j=1}^{M} |x_{ij}^{\text{р}} - x_{ij}^{\text{э}}|^2 \rightarrow min \tag{2}$$

Для вычисления расчетных концентраций необходимо решить прямую задачу химической кинетики, т.е. на основе известных начальных концентраций и заданных констант скоростей элементарных стадий получить кривые зависимости концентрации от времени. Для этого необходимо решить систему обыкновенных дифференциальных уравнений, которая в простейшем случае строится как раз на основе закона действующих масс. На основе предложенной схемы реакции была составлена следующая система ОДУ, которая оказалась нелинейной (3).

$$\begin{cases}
\frac{dx_1}{dt} = -w_1 - w_2 - w_4 - w_5 - w_7 - w_{11} - w_{20} \\
\frac{dx_2}{dt} = -w_1 - w_6 + w_9 - w_{14} + w_{17} + w_{18} \\
\frac{dx_3}{dt} = w_1 + w_2 + w_4 + w_5 - w_6 + w_7 + w_{11} - 2w_{15} - w_{16} + w_{20} + w_{21} \\
\frac{dx_4}{dt} = w_1 - w_2 \\
\frac{dx_5}{dt} = w_2 \\
\frac{dx_6}{dt} = -w_3 + w_7 - w_{19} - w_{21} \\
\frac{dx_7}{dt} = w_3 - w_4 - w_8 - w_{13} + w_{18} - w_{19} - w_{20} + w_{21} \\
\frac{dx_8}{dt} = w_3 - w_5 \\
\frac{dx_9}{dt} = w_4 + w_{13} + w_{19} \\
\frac{dx_{10}}{dt} = w_5 + w_{20} \\
\frac{dx_{11}}{dt} = w_6 - w_7 - 2w_9 - w_{16} - 2w_{17} - 2w_{18} + w_{19} + w_{21} \\
\frac{dx_{12}}{dt} = w_8 + 2w_9 \\
\frac{dx_{13}}{dt} = w_8 + 2w_9 - w_{10} \\
\frac{dx_{14}}{dt} = w_{10} - w_{11} - w_{13} \\
\frac{dx_{15}}{dt} = w_{11} - w_{12} \\
\frac{dx_{16}}{dt} = w_{12} \\
\frac{dx_{17}}{dt} = w_{12} \\
\frac{dx_{18}}{dt} = w_{13} - 2w_{14} \\
\frac{dx_{19}}{dt} = 2w_{14} \\
\frac{dx_{20}}{dt} = w_{15} \\
\frac{dx_{21}}{dt} = w_{17} \\
\frac{dx_{22}}{dt} = w_{20}
\end{cases} \tag{3}$$

где слагаемые в правой части уравнений выражаются согласно (4).

$$\begin{cases}
w_1 = k_1 x_1 x_2 \\
w_2 = k_2 x_1 x_4 \\
w_3 = k_3 x_6 \\
w_4 = k_4 x_7 x_1 \\
w_5 = k_5 x_8 x_1 \\
w_6 = k_6 x_2 x_3 \\
w_7 = k_7 x_{11} x_1 \\
w_8 = k_8 x_2 \\
w_9 = k_9 x_{11} x_{11} \\
w_{10} = k_{10} x_{13} x_2 \\
w_{11} = k_{11} x_{14} x_1 \\
w_{12} = k_{12} x_{15} \\
w_{13} = k_{13} x_{14} x_7 \\
w_{14} = k_{14} x_{18} x_{18} x_2 \\
w_{15} = k_{15} x_3 x_3 \\
w_{16} = k_{16} x_3 x_{11} \\
w_{17} = k_{16} x_{11} x_{11} \\
w_{18} = k_{16} x_{11} x_{11} \\
w_{19} = k_{16} x_7 x_6 \\
w_{20} = k_{16} x_7 x_1 \\
w_{21} = k_{16} x_6 x_6
\end{cases} \tag{4}$$

2 Численный метод

Как и большинство систем ОДУ, описывающих реальные процессы, приведенная выше система является жесткой, т.е. содержит как быстро, так и медленно изменяющиеся компоненты [7]. Таким образом, метод, решающий систему должен «не потерять» медленно изменяющие составляющие и при этом вести интегрирование с шагом, который не затянет вычисления на очень долгое время, то кроме того чревато накоплением ошибки ЭВМ в решении. С подобной задачей по сравнению с явными методами лучше справляются неявные методы. При этом неявные методы требуют больше вычислительных ресурсов, т.к. на каждом шаге решения ОДУ необходимо вместо вычисления простой формулы решать систему нелинейных уравнений. Примерами неявных методов являются: неявный метод Эйлера, метод Розенброка, метод Гира, метод Адамса–Моултона [8, 9]. В настоящей работе был выбран метод Гира (5), который является многошаговым (от 2 до 5 шагов), т.е. требует для вычислений в очередной точке времени не только значение в предыдущей точке, а значения в двух предыдущих точках [10]. Соответственно, несколько начальных шагов приходится решать более простыми методами (например, для 1-го шага нужен одношаговый метод и т.д.).

$$y_{n+2} - \frac{4}{3}y_{n+1} + \frac{1}{3}y_n = \frac{2}{3}hf(t_{n+2}, y_{n+2}) \tag{5}$$

Выбор метода объясняется рядом причин. Метод Гира обладает лучшей сходимостью по сравнению с ранее указанными, а также не требует расчёта производной, в отличие от метода Розенброка. В построенной математической модели матрица якобиана плохо обусловлена — содержит много нулей, и поэтому методы, не требующие вычисления производной, не сталкиваются с проблемой расхождения решения.

Для решения систем нелинейных уравнений был выбран метод Вегстейна (6), который показал более быструю сходимость по сравнению с методами итераций, Эйткена–Стефенссона и Ньютона [11]. Он является двухшаговым, поэтому первый шаг при решении системы необходимо вычислить методом простых итераций.

$$x_k = x_{k-1} - \frac{f(x_{k-1})(x_{k-1} - x_{k-2})}{f(x_{k-1}) - f(x_{k-2})} \tag{6}$$

Для решения обратной задачи необходимо многократно решить прямую задачу. Ввиду большого количества локальных минимумов, необходимо применять глобальные методы, среди которых для решения был выбран метод имитации отжига [12]. Этот метод основывается на имитации физического процесса, который происходит при кристаллизации вещества, в том числе при отжиге металлов. Предполагается, что атомы уже выстроились в кристаллическую решётку, но ещё допустимы переходы отдельных атомов из одной ячейки в другую. Предполагается, что процесс протекает при постепенно понижающейся температуре. Переход атома из одной ячейки в другую происходит с некоторой вероятностью, причём вероятность уменьшается с понижением температуры. Устойчивая кристаллическая решётка соответствует минимуму энергии атомов, поэтому атом либо переходит в состояние с меньшим уровнем энергии, либо остаётся на месте [6]. Достоинством метода является отсутствие необходимости вычисления производных, т.к. требуется лишь вычисление значения критерия минимизации. Однако, время и точность получения решения зависят от параметров отжига (функций температуры и перехода к новому состоянию).

3 Результаты

Для расчётов математической модели была реализована программа на языке программирования C++, который хорошо зарекомендовал себя для решения задач численными методами [12], несмотря на наличие готовых математических пакетов (Matlab, MathCad, Maple и др.). Перечисленные среды позволяют управлять многими параметрами моделирования, однако не застрахованы от получения некорректного решения и ошибки в ходе расчётов. Стоит отметить, что реализация неявных методов в этих средах как правило базируется на градиентных методах решения систем нелинейных уравнений, которые требуют дифференцирования, что малоэффективно при плохо обусловленном якобиане системы (что часто встречается в задачах химической кинетики).

Расчётами получены следующие смоделированные графики зависимости концентраций ключевых веществ от времени — на рис. 1 представлено изменение концентрации ИПБ во времени, а на рис. 2 — ГП ИПБ во времени.

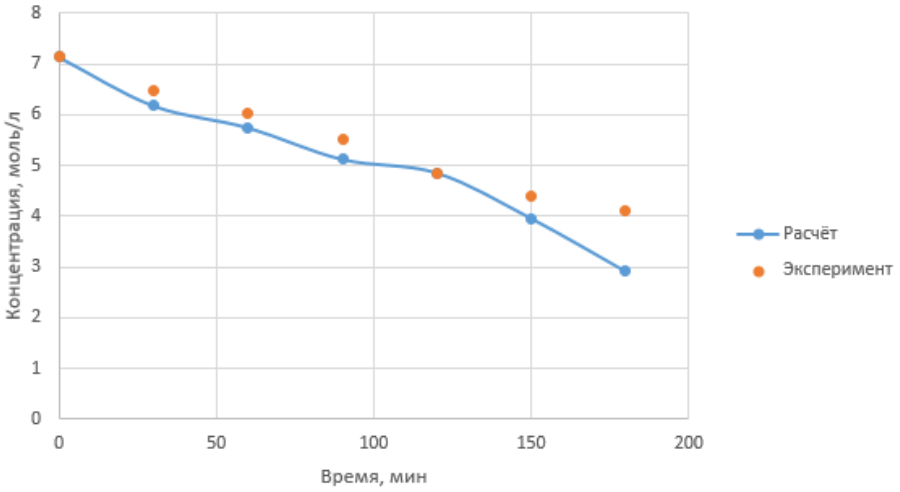


Рис. 1: Изменение концентрации ИПБ во времени

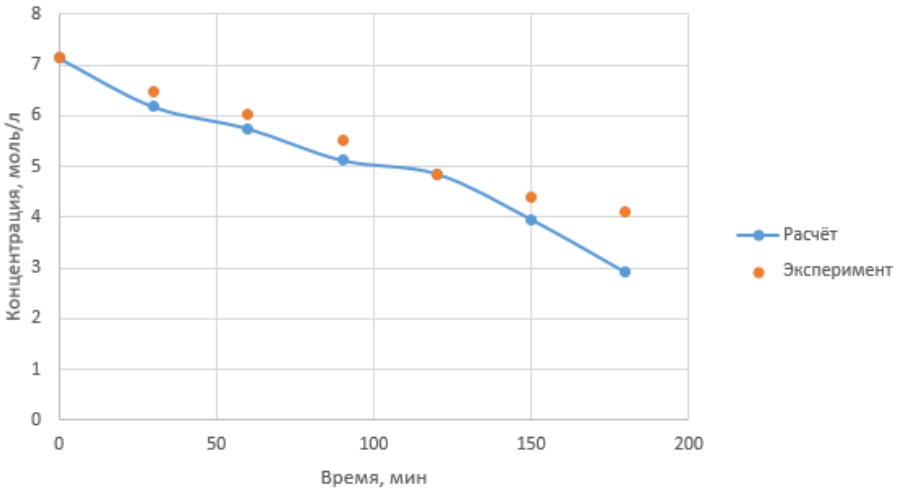


Рис. 2: Изменение концентрации ГП ИПБ во времени

Список литературы

[1] Закошанский В.М. Фенол и ацетон: анализ технологий, кинетики и механизма основных реакций. — СПб.: ХИМИЗДАТ, 2009. — 608 с.

[2] Kazuo Hattori, Yuxi Tanaka, Hiroyuki Suzuki, Tsuneo Ikawa and Hiroshi Kubota. Kinetics of liquid phase oxidation of cumene in bubble column// Journal of chemical Engineering of Japan — 1970 — P.72–78.

[3] Макалец Б.И., Кириченко Г.С., Стрыгин Е.И. и др. Кинетическая модель жидкофазного окисления кумола в гидроперекись// Нефтехимия. — 1978 — Т. 18. — №2. — С. 250–255.

[4] Bhattacharya, A. Kinetic modeling of liquid phase autoxidation of cumene / Bhattacharya, A // Chemical Engineering Journal — 2008. — P. 308–319

[5] Губайдуллин И.М., Сайфуллина Л.В., Еникеев М.Р. «Информационно-аналитическая система обратных задач химической кинетики». Учебное пособие. Изд-е Башкирск. Ун-та. — Уфа, 2003. — 89 с.

[6] Барский А.Б. Параллельные процессы в вычислительных системах: планирование и организация. М.: Радио и связь, 1990. — 255 с.

[7] Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. — М.: Мир, 1999. — 685 с.

[8] Калиткин Н.Н. Численные методы решения жестких систем // Математическое моделирование. — 1995. — Т. 7. — No 5. — С. 8–11.

[9] Розенброк Х., Стори С. Вычислительные методы для инженеров-химиков. — М.: Мир, 1968. — 443 с.

[10] Современные численные методы решения обыкновенных дифференциальных уравнений: Пер. с англ. / под ред. Дж. Холла, Дж. Уатта. — М.: Мир, 1979. — 312 с.

[11] Вержбицкий В.М. Численные методы (математический анализ и обыкновенные дифференциальные уравнения). — М.: Высшая школа, 2001. — 382 с.

[12] Каханер Д., Моулер К., Нэш С. Численные методы и программное обеспечение. — М.: Мир, 1998. — 320 с.

Вовденко Михаил Константинович — аспирант Института нефтехимии и катализа УФИЦ РАН;
e-mail: mikhail_vovdenko@rambler.ru;
Габитов Салих Азатович — магистр Уфимского государственного нефтяного технического университета;
e-mail: salihgabitov@yandex.ru;
Губайдуллин Ирек Марсович — д. ф.-м. н., ст. науч. сотр. Института нефтехимии и катализа УФИЦ РАН, проф. Уфимского государственного нефтяного технического университета;
e-mail: irekmars@mail.ru.
Дата поступления — 1 июня 2019 г.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ БИОМАРКЕРОВ ДЕГЕНЕРАТИВНЫХ ЗАБОЛЕВАНИЙ

О. Ф. Воропаева[1], К. С. Гаврилова[1,2], С. Д. Сенотрусова[1]

[1] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519-6; 51-76
DOI: 10.24411/9999-016A-2019-10014

Работа посвящена численному исследованию функционирования наиболее важных петель положительной и отрицательной обратной и прямой связи сигнального пути белка p53. Рассматриваются четыре известные математические модели, основанные на разных биологических идеализациях процесса. Выполнен сопоставительный анализ математических моделей, даны оценки области применимости моделей. Обнаружены бифуркации Андронова-Хопфа и Неймарка-Сакера в диапазонах фазовых состояний p53 и его ингибиторов, соответствующих нормальной реакции сигнального пути p53 на стресс. В рамках принятых моделей исследованы варианты гипотетических терапевтических противораковых стратегий.

Ключевые слова: численное моделирование, уравнение с запаздыванием, онкомаркер, микроРНК, p53.

Введение

Белок p53 (супрессор опухолей) играет одну из ключевых ролей в управлении ростом, старением и отмиранием клеток. Нарушение функции этого белка приводит к развитию дегенеративных заболеваний, характеризующихся чрезмерным накоплением в организме дефектных клеток или, наоборот, массовой преждевременной клеточной смертью. Установлено, что функции p53 нарушены в клетках большинства видов рака, что позволяет рассматривать p53 в качестве перспективного диагностического маркера и терапевтической мишени. В отсутствие стрессовых воздействий p53 сохраняется на достаточно низком уровне, что предотвращает ненужную в нормальных условиях экспрессию его генов-мишеней. В случае превышения физиологически допустимого уровня повреждений ДНК p53 активируется для избавления от генетически опасных дефектных клеток. Результат активации p53 – остановка клеточного цикла и репарация ДНК при незначительном повреждении или запуск программы апоптоза (генетически запрограммированной смерти клеток) при сильном стрессовом сигнале [1, 2]. В регулировании динамической реакции белка p53 участвует целый ряд положительных и отрицательных циклов обратной и прямой связи, через которые с p53 взаимодействуют, в частности, его ингибиторы — белки Mdm2 и Wip1, а также многочисленные семейства микроРНК (miR). Исследование роли и особенностей функционирования этих циклов в p53-сети — это ключ к более глубокому пониманию процессов дегенерации, старения, механизмов развития онкологических заболеваний и их подавления. Учитывая чрезвычайную сложность структуры сети p53, механизмов взаимодействия и функций ее участников, планирование новых лабораторных экспериментов в этой области биомедицины может быть существенным образом упрощено предварительными оценками, основанными на результатах математического моделирования.

Математические модели, предназначенные для изучения особенностей кинетики, регуляторной функции и/или терапевтического потенциала ключевых участников сети p53 при различных заболеваниях, предложены в целом ряде работ (см. обзоры в [3, 4] и цитируемой там литературе). Многообразие существующих математических моделей сети p53, как правило, отражает разнонаправленность этих исследований. Однако скрупулезный подбор экспериментальных фактов, определяющих границы применимости каждой разработанной модели, является пока нерешенной проблемой.

ISBN 978-5-901548-42-4

Одной из главных целей настоящей работы является сопоставительный анализ четырех известных математических моделей динамики сигнального пути p53, основанный на использовании лабораторных данных [1]. Предложены модификации моделей, расширяющие область их применимости за счет уточнения описания механизма отрицательной обратной связи p53–ингибитор (Mdm2 или Wip1). В работе представлены результаты численного анализа решений моделей в широком диапазоне значений параметров, позволившего выявить, в частности, бифуркации рождения предельных циклов и тороидальных аттракторов как свидетельство весьма тонкой организации ответа биологической системы на стрессовые воздействия.

1 Математические модели

Модель 1. Для описания взаимосвязи p53–ингибитор–микроРНК привлекается нелинейная система дифференциальных уравнений с запаздывающими аргументами [4]. Модель составлена из балансных соотношений, отражающих вклад в совместную динамику системы p53–белок-ингибитор–микроРНК следующих трех ключевых механизмов: генерации, деградации (в рамках модели эти термины подразумевают как конститутивные спонтанные процессы, так и обусловленные влиянием каких-либо не описанных явно молекул), а также особо выделяемых процессов взаимодействия p53–ингибитор–miR (p53-зависимая активация генерации белка-ингибитора p53 и микроРНК, инактивация/деградация p53 под влиянием ингибитора и микроРНК-зависимое подавление генерации ингибитора):

$$\frac{dP}{dt} = \beta_P - \alpha_{PI} F(P(t), I(t), K_f) - \alpha_P P(t),$$

$$\frac{dI}{dt} = \beta_{IP} G(P(t-\tau_P), I(t-\tau_P)) - \alpha_I I(t) - \alpha_{Im} F(I(t-\tau_I), m(t-\tau_I), K_m),$$

$$\frac{dm}{dt} = \beta_m + \beta_{mP} F(P(t-\tau_m), m(t-\tau_m), K_P) - \alpha_m m(t).$$

Для определения функций взаимодействия использовался подход, основанный на аппроксимациях типа Гольдбетера-Кошланда и Хилла:

$$F(P, I, K) = \frac{1}{2}(P + I + K - \sqrt{(P + I + K)^2 - 4PI}),$$

$$G(P, I, K_g) = \frac{P - F(P, I, K_f)}{P + K_g - F(P, I, K_f)}.$$

Здесь P, I, m — уровни p53, белка-ингибитора p53 и микроРНК, соответственно; β_P и β_m — константы скорости генерации белка p53 и микроРНК; α_{PI} — константа скорости деградации p53 под влиянием ингибитора; $\alpha_P, \alpha_I, \alpha_m$ — константы скорости деградации p53, ингибитора и микроРНК; β_{IP} и β_{mP} — константы скорости p53-зависимой генерации ингибитора и микроРНК; α_{Im} — константа скорости микроРНК-зависимого подавления белка-ингибитора; параметры K_f, K_g, K_P, K_m имеют смысл констант диссоциации, которым в рамках принятой модели отводится роль регуляторов, обеспечивающих «тонкую настройку» уровней взаимосвязи образующих комплексы участников сети. Параметры τ_P, τ_I и τ_m определяют время запаздывания реакции белков и микроРНК на сигналы. При анализе адекватности модели 1 в качестве белков-ингибиторов p53 рассматриваются Mdm2 и Wip1.

Модель 2. Для моделирования совместного функционирования двух петель отрицательной обратной связи p53–Mdm2 и p53–Wip1 привлекается предложенная в [2] система нелинейных уравнений с запаздывающими аргументами. Поскольку данная модель рассматривалась в [2], в первую очередь, как инструмент для описания серии лабораторных экспериментов, в уравнения дополнительно включены слагаемые, позволяющие рассматривать воздействие нутлина — вещества, используемого для ингибирования взаимодействия p53 и Mdm2 в терапевтических целях:

$$\frac{dP_i}{dt} = \beta_P - \alpha_{MP_i} \frac{K^n}{K^n + N^n(t-\tau_N)} MP_i - \beta_{SP} \frac{S^{N_s}}{S^{N_s} + T_S^{N_s}} P_i - \alpha_{P_i} P_i,$$

$$\frac{dP_a}{dt} = \beta_{SP} \frac{S^{N_s}}{S^{N_s} + T_S^{N_s}} P_i - \alpha_{MP_a} \frac{K^n}{K^n + N^n(t-\tau_N)} MP_a,$$

$$\frac{dM}{dt} = \beta_{M_i} + \beta_M P_a\,(t-\tau_M) - \alpha_{SM} SM - \alpha_M M,$$

$$\frac{dW}{dt} = \beta_W P_a\,(t-\tau_W) - \alpha_W W,$$

$$\frac{dS}{dt} = \beta_S - \alpha_{W_S} \frac{W^{n_w}}{W^{n_w} + T_W^{n_w}}\, S - \alpha_S S.$$

Здесь P_i, P_a — уровни неактивированного и активного белка p53; M, W — уровни Mdm2 и Wip1; S — уровень активного сигнала о повреждении; τ_N, τ_M, τ_W– параметры запаздывания; N — уровень нутлина.

Модель 3. Последние данные свидетельствуют о бимодальном переключении p53, наблюдаемом при решении «судьбы» клетки в условиях стресса — при малом повреждении в системе возникают периодические колебания уровней белков, а при сильном повреждении наблюдается монотонное увеличение p53, ассоциированное с запуском программы клеточной смерти (апоптоза). Для описания этого эффекта в [5] предложена математическая модель, представляющая собой систему нелинейных дифференциальных уравнений с запаздывающими аргументами. Модель содержит 10 уравнений динамики ключевых белков сети p53 с учетом 28 биохимических реакций, которые считаются наиболее важными (в рамках принятой биологической модели) в механизме, обеспечивающем ответ сигнального пути p53 на стресс. В модели 3 рассматриваются две основные петли отрицательной обратной связи — p53–Mdm2 и p53–Wip1–ATM. Модель учитывает не только конститутивные спонтанные процессы генерации и деградации p53 и двух его ключевых ингибиторов Mdm2 и Wip1, но и активацию белков через механизм фосфорилирования. Уровень этопозида рассматривается как входной параметр, инициирующий стресс. Предполагается, что этопозид как источник стресса может активировать белок ATM, что согласуется с известными экспериментальными наблюдениями. Поскольку известно, что общий уровень белка ATM практически не изменяется в течение 72 часов, но существенно изменяется его активность, в рамках принятой модели центральное место отводится именно активной форме ATM. Активированный белок ATM может фосфорилировать p53 и Mdm2, что способствует изменению активности p53 и Mdm2. Фосфорилированный p53 может активировать Wip1, который, в свою очередь, обеспечивает (в рамках принятой модели) регуляцию механизма фосфорилирования.

Модель 4. В работе [6] предложена математическая модель, которая предназначалась ее авторами для изучения функционирования механизма запуска программы p53-зависимого апоптоза в ответ на сигналы о повреждении ДНК. Модель представляет собой нелинейную систему 14 обыкновенных дифференциальных и четырех алгебраических уравнений. Она разделена на два блока. Такое разделение обусловлено принятой биологической моделью, в рамках которой ключевую роль, как и в моделях 1–3, играет петля отрицательной обратной связи p53–Mdm2 (блок 1). Во втором блоке, где непосредственно решается «судьба» клетки, p53 функционально разделен на три группы. Группа p53-киллеров активирует апоптотические гены, такие как PUMA, p53DINP1 и p53AIP1. PUMA-белок индуцирует выделение из митохондрий цитохрома, который стимулирует белковый комплекс APAF1 для активации апоптосом, непосредственно (в рамках принятой модели) запускающих программу апоптоза. Другая форма p53 — p53-хелперы — ответственна за индукцию белка p21. Последний инициирует остановку клеточного цикла, давая время клетке для восстановления поврежденной ДНК, или участвует в запуске программы необратимого прекращения деления (старения). Третья, вспомогательная, форма p53 индуцирует Wip1, который, согласно принятой модели, сдерживает образование p53-киллеров. Отметим, что конкретная реализация биологической связи между блоками и внутри каждого из них представляет собой предмет исследований как в [6], так и в рамках настоящей работы.

В связи с тем, что уравнения моделей 1–4 представляют собой нелинейные системы дифференциальных уравнений, часть из которых содержит функции с запаздывающими аргументами, возникает проблема нарушения гладкости решения в точке $t = 0$ и последующего переноса разрывов производных в точки, кратные величине запаздывания. Численное решение представленных систем уравнений основано на применении идеи метода шагов, который позволяет рассматривать возникающие задачи с начальными условиями как задачи Коши для нелинейных систем ОДУ первого порядка, решаемые последовательно на интервалах времени, равных величине запаздывания. На каждом из этих интервалов функции с запаздывающими аргументами заданы в узлах расчетной сетки — они определяются из начальных условий или из численного решения, полученного на предыдущем этапе реализации метода шагов. Такой подход позволяет привлечь к решению задач широко распространенные численные методы решения задачи Коши для системы ОДУ — метод предиктор-корректор второго порядка, методы Адамса–Бэшфорта–Моултона или Гира (фактический порядок точности, как правило, не превосходил 3). Из соображений простоты численной реализации алгоритма решения систем уравнений шаг по времени h выбирался таким образом, чтобы величины запаздывания

были кратными h. Расчеты показали, что данное упрощение не оказывает существенного влияния на характер и точность решения, если величина шага сетки является достаточно малой величиной. Окончательный выбор численного метода осуществлялся на основе анализа результатов серий методических расчетов. При решении каждой нелинейной системы дифференциальных уравнений использовалась идея метода Зейделя.

2 Результаты численного анализа

Прежде всего, представляет интерес оценка области применимости моделей 1–4, адекватность каждой из которых в [2, 4–6] обосновывалась сопоставлением с существенно разными лабораторными экспериментами. В настоящей работе в качестве единого критерия для оценки всех моделей используются лабораторные данные [1], представленные в виде совокупности характерных фазовых состояний системы p53–Wip1, наблюдаемых in vitro при облучении раковых клеток с p53 дикого типа. Данные [1] показывают, что значительная часть клеток характеризуется относительно низким уровнем p53 и его отрицательного регулятора Wip1. Это следует рассматривать, по-видимому, как нормальную реакцию системы на стрессовое воздействие, приводящее к остановке клеточного цикла и репарации поврежденной ДНК. Из анализа экспериментальных данных [1] следует также, что клеткам с относительно низким сигналом p53 могут соответствовать весьма высокие уровни сигнала Wip1. Такие состояния системы p53–Wip1 характеризуют, в частности, ситуацию затухания p53-зависимого апоптоза и риска рака. Точно так же наблюдаются состояния с высоким уровнем p53 при низком уровне Wip1, которые связываются с запуском программы клеточной смерти (p53-зависимого апоптоза). Известно, что именно два последних состояния позволяют рассматривать p53, его ингибиторы и p53-зависимые микроРНК как биомаркеры наиболее опасных дегенеративных заболеваний, включая рак, ишемическую болезнь сердца, болезни Альцгеймера и Паркинсона. Результаты численных экспериментов показали, что совокупность рассчитанных неподвижных точек решений модели 1 (стационарных значений уровней p53 и Wip1) располагается в фазовой плоскости в согласии с данными [1]. При этом модель 1 воспроизводит также наблюдаемые in vitro [2] колебательные режимы функционирования системы (характерные, как правило, для низкого уровня повреждения ДНК), которые регулируются петлей отрицательной обратной связи p53–Wip1. Важно, что все характерные значения численных решений модели 1 располагаются на фазовой плоскости (p53, Wip1) в том же диапазоне величин, что и экспериментальные данные. Подобная степень согласия с экспериментальными данными получена и для модели 3. В то же время, численные эксперименты показали, что в рамках моделей 2 и 4 отрицательная обратная связь p53 и его ингибиторов выражена слабо — изменение p53 вызывает сколько-нибудь значимые изменения уровня Wip1 (и наоборот) лишь в относительно узком диапазоне значений уровней p53 и Wip1. Это существенно ограничивает область применимости моделей 2 и 4, фактически не позволяя их использовать как модели функционирования системы онкомаркеров p53–Wip1 и/или как инструмент анализа эффективности противоопухолевых терапевтических стратегий.

Модель 1 нацелена, в первую очередь, на исследование динамики p53-зависимых микроРНК на основе минимального подхода к моделированию. Ранее авторами разработана иерархия минимальных математических моделей функционирования сети p53–ингибитор–miR с положительным прямым и обратным вариантами связи p53–miR и выполнен комплекс численных экспериментов, иллюстрирующих адекватность принятого подхода. Модель 1 — единственная модель этой иерархии, учитывающая не только прямую положительную связь p53–микроРНК, но и позволяющая «различать» разные микроРНК по уровню связи с p53 (речь идет о конкретных p53-зависимых микроРНК, для которых обратное воздействие микроРНК на p53 является несущественным либо, напротив, значимость этого воздействия чрезвычайно высока для функционирования системы в целом). Показано, в частности, что численные решения модели 1 качественно и количественно согласуются с экспериментальными данными [4] о динамике p53, а также адекватно воспроизводят изменение уровня микроРНК в ответ на активацию p53. Отдельно анализировалась адекватность математической модели взаимосвязи белка-ингибитора и микроРНК. С этой целью использовались полученные в культивируемых клетках карциномы экспериментальные данные для мРНК Mdm2 и miR-143. Численные решения модели 1 удовлетворительно согласуются с результатами лабораторных исследований, в которых было показано, что экспрессия Mdm2 отрицательно коррелирует с экспрессией miR-143. На этой основе в рамках модели 1 и ее более простых модификаций даны обобщенные оценки диагностического потенциала p53-зависимых микроРНК при онкологических и нейродегенеративных заболеваниях. Рассмотрены возможные стратегии восстановления нормального уровня p53 и связанных с ним микроРНК в целях профилактики угрозы рака. Изучены варианты противораковой терапии, связанные с одновременной гиперактивацией двух регуляторов апоптоза – p53 и микроРНК. В рамках модели 1 показана потенциально

высокая эффективность противораковой терапии, мишенью которой является белок-ингибитор p53 как основное звено петли положительной обратной связи p53–микроРНК. Выполненный комплекс исследований дает основания для того, чтобы рассматривать модель 1 как базовую при разработке более полных математических моделей сигнального пути p53.

Анализ численного решения модели 2 показал, что, в соответствии с известными представлениями, при отсутствии повреждения p53 находится в неактивном состоянии. Процесс активации p53 сопровождается стремительным ростом уровня активного p53 в ответ на сигнал и последующей регуляцией этой активности белками-ингибиторами Mdm2 и Wip1, при этом следует отметить определенную цикличность данного затухающего процесса. Результаты расчетов демонстрируют доминирующую роль Mdm2 как отрицательного регулятора p53, что, как представляется, соответствует принятой биологической модели [2]. Предложена модификация модели 2, корректирующая описание отрицательной обратной связи p53–ингибитор для согласования с наблюдениями [1]. Анализ адекватности модели 2 и ее модификации основан на сопоставлении полученных численных решений с экспериментальными данными о динамике p53 в условиях стресса, порожденного ингибированием взаимодействия p53–Mdm2 малыми молекулами нутлина. Лабораторные измерения показывают, что реакция p53 и системы p53–Mdm2 на введение нутлина является достаточно сложной и зависит от уровня нутлина, продолжительности и режима его введения в клетку. В лабораторном эксперименте постоянное введение нутлина (при постепенном наращивании его дозы на первом этапе эксперимента) достаточно быстро приводит к росту и стабилизации p53 на уровнях, значительно превышающих «безнутлиновый» в те же самые моменты времени. Аналогичный результат демонстрируют и приближенные решения модели 2 и ее модификации. При этом особенность принятой математической модели 2 состоит в том, что динамика p53 зависит от распределения во времени концентрации нутлина N, но для остальных компонент решения эта зависимость выражена крайне слабо. Выполнен численный анализ решений исходной модели и ее модификации в широком диапазоне параметров, который дает представление о качественных свойствах решений. Обнаружены бифуркации Андронова-Хопфа и Неймарка-Сакера. Для модифицированной модели обнаружено существование двух решений – с неподвижной предельной точкой и с предельным циклом – при одинаковых значениях параметров модели.

В численных экспериментах с привлечением моделей 3 и 4 главное внимание уделялось анализу области применимости, о котором шла речь в начале данного раздела. Дополнительно с использованием модели 3 проведено численное моделирование динамики сети p53 под влиянием стрессового сигнала и терапевтического воздействия малыми молекулами этопозида. Показано, что результаты настоящих расчетов согласуются с известными представлениями (см., в частности, [5]) о бимодальном переключении пути p53 в зависимости от уровня повреждения: при малом повреждении возникают периодические колебания уровней белков, а при сильном повреждении наблюдается монотонное увеличение p53, ассоциированное с массовой гибелью клеток. Выполнено численное моделирование механизма запуска p53-зависимого апоптоза на основе математической модели 4. Рассмотрены модификации модели 4, корректирующие описание отрицательной обратной связи p53-Wip1. Выполнено численное исследование функционирования сигнального пути p53 как ключевого звена в программе клеточного ответа на кратковременное повреждение ДНК в зависимости от уровня, количества и длительности сигналов о повреждении. Показано, в частности, что при нескольких последовательных повреждениях ДНК может наблюдаться усиление процессов гибели клеток, связанное исключительно с взаимным влиянием разных, даже относительно слабых, сигналов о повреждении.

Список литературы

[1] Batchelor E., Mock C.S., Bhan I., Loewer A., Lahav G. Recurrent initiation: A mechanism for triggering p53 pulses in response to DNA damage // Molecular Cell. 2008. V. 30, iss. 3. P. 277–289.

[2] Purvis J.E., Karhohs K.W., Mock C., Batchelor E., Loewer A., Lahav G. p53 dynamics control cell fate // Science. 2012. V. 336, iss. 6087. P. 1440–1444.

[3] Воропаева О.Ф., Сенотрусова С.Д., Шокин Ю.И. Дерегуляция p53-зависимых микроРНК: результаты математического моделирования // Математическая биология и биоинформатика. 2017. Т. 12, iss. 1. С. 151–175.

[4] Сенотрусова С.Д., Воропаева О.Ф. Математическое моделирование функционирования положительной связи в системе онкомаркеров p53–микроРНК // Сиб. журн. вычисл. математики. 2019. Т. 22, iss. 3.

[5] Sun T. , Cui J. A plausible model for bimodal p53 switch in DNA damage response // FEBS Letters. 2014. V. 588. P. 815–821.

[6] Zhang T., Brazhnik P., Tyson J.J. Exploring Mechanisms of the DNA-Damage Response: p53 Pulses and their Possible Relevance to Apoptosis // Cell Cycle. 2007. V. 6, iss. 1. P. 85–94.

Воропаева Ольга Фалалеевна — д.ф.-м.н., вед. науч. сотр. Института вычислительных технологий;
e-mail: vorop@ict.nsc.ru;
Сенотрусова Софья Дмитриевна — мл. науч. сотр. Института вычислительных технологий;
e-mail: senotrusova.s@mail.ru;
Гаврилова Ксения Сергеевна — студентка Новосибирского государственного университета;
e-mail: ksu483@yandex.ru.
Дата поступления — 30 апреля 2019 г.

ОБ ОДНОМ АЛГОРИТМЕ СТАТИСТИЧЕСКОЙ РЕГУЛЯРИЗАЦИИ

Ю. В. Гласко

Научно-исследовательский вычислительный центр МГУ им. М.В. Ломоносова, 119992, Москва

УДК 519.63; 519.24; 550.83; 553.9
DOI: 10.24411/9999-016A-2019-10015

Рассмотрена обратная задача концентрации источника гравитационной либо магнитной аномалии и алгоритм ее решения. Рассмотрен случай положительной магнитной аномалии на фоне отрицательной гравитационной аномалии. Предложена методика определения геометрических характеристик области где сконцентрирован источник. Обратная задача концентрации распределения плотности источника решается на основе метода Монте–Карло в рамках статистической регуляризации. Приведены результаты нескольких модельных примеров. Оценены их погрешности по невязке и сглаживающему функционалу относительно точных данных и данных с погрешностью.

Ключевые слова: обратная задача, статистическая регуляризация, метод Монте–Карло.

Введение

Применение вычислительной математики относительно месторождений углеводородов позволяет разработать новые эффективные и экономичные методики. При этом возникают проблемы регуляризации интерпретации ограниченных выборок данных гравитационной и магнитной разведки [8]. В этой связи, актуальна оптимизация необходимой для регуляризации априорной информации.

Решение рассматриваемой в статье обратной задачи относительно источника аномалии посредством многократного решения граничной обратной задачи является альтернативой актуальной проблемы продолжения поля с дневной поверхности [2], [6] и др.

При определении геометрических характеристик объектов-источников аномалии мы используем гравитационное поле от коллектора (в основном отрицательное) и всплески магнитного поля от окружающих коллектор эпигенетических элементов. Априорная информация сведена в web-ориентированную информационную систему. Для определения по потенциальному полю распределения плотностей на контурах охватывающих источники следует использовать регуляризирующий алгоритм для некорректной задачи относительно системы линейных алгебраических уравнений [12].

Рассматриваемый здесь 3D алгоритм решения обратной задачи для плотностей, является применением идеи 2D balayage метода А. Пуанкаре касательно потенциалов, которая [16] обобщает метод Л. Шварца [4]. Развитие подхода А. Пуанкаре в работах В.Н. Страхова [19] для плотностей проходило параллельно с разработкой метода установления С.К. Годунова [5]. В работах болгарского геофизика Д. Зидарова [25] развивается метод сметания-концентрации относительно потенциальных полей. В работе вводятся классы "материнских"тел. Отметим, что в нашем случае (газ, нефть) мы имеем отрицательную эффективную плотность [8]. Заметим, что геофизическая классификация месторождений является естественной априорной информацией.

Рассматривается численный подход основанный на методике А.А. Самарского [17]: объект – математическая модель – сеточная аппроксимация – вычислительный алгоритм – программа – вычислительный эксперимент – анализ результатов.

Здесь мы рассматриваем комбинаторную 3D задачу интерпретации. Анализ результатов вычислительных экспериментов включает построение альбома потенциальных полей для различных геометрических характеристик изучаемого объекта. Относительно него строится альбом выметенных плотностей. На следующем этапе, наряду с математической моделью мы рассматриваем статистическую модель. Для аппроксимации

ISBN 978-5-901548-42-4

модели (в том числе и сеточной) используем вариант итерационно-проекционного метода. При этом вычислительный эксперимент включает серию статистических испытаний в рамках статистической регуляризация.

Развитие методики решения обратных задач интерпретации в настоящее время связано с расширением технических возможностей вычислительной техники и развитием информационных технологий. В связи с указанными возможностями оптимизируется время вычислений для численных реализаций статистических моделей и соответствующей регуляризации.

1 Об обратной задаче интерпретации геометрических характеристик месторождения углеводородов

Рассматриваемая комбинаторная обратная задача интерпретации включает два этапа. Первый этап состоит в определении геометрических характеристик источника $f(X)$, $X \in \Omega \subset V$ (формы, глубины залегания). Второй этап состоит в определении плотности источника $f(X)$. На втором этапе статистические испытания проводятся на заданном интервал значений искомой плотности [10].

Рассмотрим первый этап указанной задачи для гравитационной аномалии от месторождений УВ. Создаваемая коллектором (Ω) гравитационная аномалия $u(s) = \Delta g(s)$, $s \in \Gamma \equiv \partial V$ является отрицательной [13]. Отрицательны соответствующие эффективные плотность и масса δ_{ef}, m_{ef} [8]. Для расчета важно к какому множеству структур месторождения УВ ($C_q H_s$) $\Phi = \{\Phi_p\}$, $p = 1, ..., P$ принадлежит коллектор. Здесь мы можем использовать web-ориентированную ИС "Deposit"относительно 235 месторождений созданную нами на основе совместных с В.М. Мегерей, В.И. Старостенко и др. публикаций [9–12, 14]. В вычислительном эксперименте мы не будем детализировать плотностные характеристики самого коллектора, а ограничимся модельным случаем предполагая область заполненной исключительно УВ. В свою очередь собственно коллектор может быть аппроксимирован односязной или многосвязной областью $\Omega = \bigcup\limits_{k=1}^{K} \Omega_k$. В рамках классической антиклинальной модели [9] Ω является односвязной выпуклой областью без дырок. Геосолитонная концепция [12] рассматривает Ω как объединение областей указанной структуры. При этом предполагаем, что $\bigcap\limits_{k=1}^{K} \Omega_k = \emptyset$ и $\{s\} \in \bigcap\limits_{k=1}^{K} \partial\Omega_k = \emptyset$. Указанная область может быть аппроксимирована множеством шаров: $\Omega \simeq \bigcup\limits_{l=1}^{L} \Upsilon_l$, в каждом из которых его плотность распределена однородно. В данном случае, мы вычисляем гравитационное поле с учетом эффективной плотности δ_{ef} для каждого Υ_l (точки являющейся центром шара).

Нас интересуют $(x, y, z) \in \Gamma$. При этом Γ может быть границей нижнего полупространства, либо границей выпуклой односвязной области содержащей Ω.

При этом потенциал гравитационного поля имеет вид:

$$g(x, y, z) = \int\limits_{\Omega}\int\int \frac{GM_{\text{ef}} d\xi d\eta d\zeta}{((x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2)^{\frac{1}{2}}}.$$

В случае без шаров для потенциала притяжения V (соответствует g) и его производной V_z (соответствует Δg) используются эффективные плотности [8].

Более экономичным представляется подход основанный на выделении локальной составляющей поля $U_{lok}(s)$ из $U(s) = U_{lok}(s) + U_{fon}(s) + \delta U(s)$, где $U_{lok}(s)$ — гравитационное поле Ω, $U_{fon}(s)$ — гравитационное поле фона (остальных элементов структуры), $\delta U(s)$ — погрешность измерений. Таким образом используется гравитационная аномалия Буге. В данном случае, в рассматриваемой обратной задаче на втором этапе мы будем определять плотность коллектора $\delta(X)$, $X \in \Omega$. В рассматриваемой нами модели в качестве $\delta(X)$ рассматривается плотность $C_q H_s$. Указанные выше методики позволяют решить первый этап поставленной обратной задачи интерпретации используя невязку относительно экспериментально измеренных данных.

Обратная задача определения источника $f(X)$, $X \in \Omega_{me} \subset V_{me}$ магнитной аномалии $u(s) = \Delta\mu(s)$, $s \in \Gamma_{me} \equiv \partial V_{me}$ формулируется и для терригенного слоя геомагнитной модели [9] антиклинального нефтегазового месторождения. Как следует из исследований на месторождениях [14], эпигинетические элементы терригенного слоя создают магнитное поле $\Delta\mu(\bar{X}), \bar{X} \in V_{me} \setminus \Omega_{me}$ сосредоточенное над центром коллектора. Указанные магнитные минералы создают в наблюдаемом поле локальные аномалии не превышающие 10–20 нТл.

В данном случае, мы будем рассматривать обратную задачу интерпретации относительно источника $f(X)$ для заданного граничного условия $\Delta\mu(s), s \in \Gamma_{me} \equiv \partial V_{me}$. Нас интересует двухуровневое (N_1 уровневое, где $N_1 >= 2$) расположение эпигенетических элементов. Таким образом, над Ω на фоне отрицательного эффективного гравитационного поля $\Delta g(s),\ s \in \Gamma$ на OXY мы наблюдаем положительное магнитное поле $\Delta\mu(s),\ s \in \Gamma_{me} = \Gamma$ (случай V_{me} — нижнее полупространство).

Если V_{me} включающее терригенный слой отлично от нижнего полупространства, но является параллелепипедом либо кубом, то мы фиксируем магнитную аномалию и на его боковых гранях ($\Gamma_{me} \neq \Gamma$) [9]. При этом мы знаем интервал глубин терригенного слоя где расположен $\Omega_m e$. Для обратной задачи касательно эпигенетических элементов мы используем как вариант проекционного метода относительно Ω_{me} и V_{me}, так и 3D-вариант метода с использованием особых точек полного нормированного градиента В.М. Березкина и интропродолжения поля вдоль оси OZ [14].

2 Обратная задача интерпретации плотности относительно месторождения углеводородов

Здесь мы будем многократно решать обратную граничную задачу [1] относительно плотности $u(X) = \delta(X)$ $X \in V \supset \Omega$ с целью определения плотности $u_0(X)$ в области Ω, геометрические характеристики которой известны (либо частично известны). Это второй этап рассматриваемой комбинаторной задачи.

В качестве модели мы рассмотрим стационарный процесс с источником в рамках следующей краевой задачи для уравнений Пуассона и Лапласа:

$$D\Delta u(X) = -f(X),\ X \in \Omega \subset V, \tag{1}$$

$$D\Delta u(X) = 0,\ X \in V \setminus \Omega, \tag{2}$$

$$\left.\frac{\partial u(X)}{\partial n}\right|_{\Gamma} = 0, \tag{3}$$

$$u(X) = u_0(X) = cf(X),\ X \in \Omega \tag{4}$$

Здесь D и c обеспечивают совпадение размерностей. В общем случае, $f(X) = \sum\limits_{i=1}^{P} f_i(X),\ X \in \Omega \subset V$, $\Omega = \bigcup\limits_{i=1}^{P} \Omega^i$. Предполагается, что вся масса из источника сосредотачивается на Γ.

Граничная задача [1] состоит в определении:

$$u(X)|_{\Gamma} = u_{\Gamma}(X) \tag{5}$$

Модель для потенциала гравитационного поля $u(X) = V$ и плотности $\delta(X)$ имеет вид:

$$D\Delta u(X) = -F(X),\ X \in \Omega \subset V,$$

$$D\Delta u(X) = 0,\ X \in V \setminus \Omega,$$

$$\left.\frac{\partial \delta(X)}{\partial n}\right|_{\Gamma} = 0,$$

$$\delta(X) = \delta_0(X) = cf(X),\ X \in \Omega$$

$$u(X)|_{\Gamma} = u_{\Gamma}(X)$$

где в 3D для единичной сферы $F(X) = 4\pi G\delta(X)$ гравитационная постоянная $G = 6.67 \times 10^{-8}$ см^3/ (гс^2) [24].

Модель процесса на основе параболического уравнения для плотности имеет вид:

$$D\Delta u(X,t) = u_t(X,t) - f(X,t),\ X \in V, t \in [0, +\infty)$$

$$\left.\frac{\partial u(X,t)}{\partial n}\right|_{\Gamma} = 0,$$

$$u(X,0) = u_0(X) = cf(X,0),\ X \in \Omega$$

Граничная задача состоит в определении:

$$u(X,T) = u_\Gamma(X) = cf(X,T),\ X \in \Gamma$$

С момента времени $t = T$ можно рассматривать стационарный процесс [5].

В геофизической практике возможна интерпретация относительно гравитационного (либо магнитного) поля либо относительно физического распределения плотности. Мы рассмотрим второй случай. Предполагаем заданными $f(X)$ и Ω. Отметим, что геометрические характеристики Ω мы определили на первом этапе комбинаторной обратной задачи интерпретации. Обратная задача (5) на основе математической модели процесса (1–4) может быть записана в операторном виде:

$$A_1 u_\Gamma(X) = f(X),\ X \in V$$

Будем решать ее методом Монте–Карло. В качестве статистических испытаний мы рассмотрим цикл относительно структурной схемы процесса на основе balayage метода А. Пуанкаре для плотностей [16]. Эта структурная схема процесса состоит в заполнении области $V \supset \Omega$ системой шаров $\{\Sigma_i\}, i = 1, ..., M$ и перераспределении $u(X)$ из шаров на их границы $\{\partial\Sigma_i\}, i = 1, ..., M$ до тех пор пока вся $u(X)$ окажется на Γ. Предполагается, что объединение частей поверхностей обращенных к Γ граничных шаров $\{\bigcup_{j=1}^{J} \partial\Sigma_j^\Gamma\}$ достаточно точно аппроксимирует Γ. Предположим, что каждые 2 из пересекающихся шаров $\Sigma_i,\ \Sigma_j$ являются вариантом составной области рассматриваемой в альтернирующем методе Шварца. При этом в каждое $\Sigma_i \bigcap \Sigma_j$ можно поместить луночку — то есть выполнен критерий Шварца [4]. Указанную структурную схему обозначим оператором B: $Bu(X),\ X \in V$. В рамках статистических испытаний нужно провести достаточное количество циклов balayage метода, что бы получить распределение плотности на границе $u(X)|_\Gamma$ приближающееся к достоверному в силу закона больших чисел. В качестве такового следует взять его математическое ожидание $Mu_\Gamma(X)$. Отметим, что условия (3–4) являются аналитическим описанием указанных через balayage процесс условий относительно границы Γ. На гильбертовом пространстве с помощью проекционного оператора можно проводить balayage процесс в V сразу для всей области Ω.

Численная реализация [25] оператора B на 3D сетке с шагом h осуществляется посредством B^h: $B^h u(X)$, $X \in V^h$. Оператор реализуется с помощью n итераций для трех вложенных циклов по координатам x, y, z относительно плотности $u(X),\ X \in V^h$. При этом шары $\{\Sigma_i\}, i = 1, ..., M$ аппроксимируются семиточечными схемами типа "крест"$\{\Sigma_i^h\}, i = 1, ..., M$. Итерационный цикл заканчивается когда все сеточные плотности окажутся с достаточной точностью на сеточном аналоге Γ^h границы Γ [27].

Обратная задача определения плотности источника (4) основана на математической модели (1–2), (5). При этом геометрические характеристики области Ω известны полностью либо частично (определены на первом этапе комбинаторной задачи). Соответствующую обратную задачу можно записать в операторном виде:

$$Af(X) = u_\Gamma(X),\ X \in V \tag{6}$$

В рассматриваемом случае обратная задача (6) решается на основе статистической модели для обратной граничной задачи (1–5). Эта статистическая модель определена структурной схемой — оператором B: $Bu(X),\ X \in V$. Поскольку в математической модели не определено условие (4), то в рамках метода Монте–Карло мы будем это условие разыгрывать. В то же время заданы граничные условия. Эти значения определены посредством экспериментов в лабораторных и естественных условиях. Здесь мы не будем проводить сопоставление между гравитационным (или магнитным) полем и физическим распределением плотностей на границе Γ. В рассматриваемом случае они определяются посредством программного решения ($\bar{u}_\Gamma(X)$) "генеральной"обратной граничной задачи. В том числе, относительно $u_\Gamma(X)$ можно использовать достаточное количество статистических испытаний относительно balayage метода ($Bu(X)$). Таким образом, нам следует определить физическое распределение плотностей $f(X)$ зная геометрию области Ω и граничное распределение физической плотности $\bar{u}_\Gamma(X)$). При этом мы используем статистическую модель $Bu(X),\ X \in V$. В случае магнитного поля мы имеем дело с плотностями эпигенетических образований. Отметим, что плотность зависит от множества параметров структуры месторождения Φ: $\Theta = \{\Theta_1, \Theta_2, ..., \Theta_L\}$. Следовательно, $f(X), X \in \Omega$ соответствует $\delta(X, \Theta)$. При этом для нефти большое значение имеют температура T и глубина H: $\Theta_1 = T$, $\Theta_2 = H$ относительно коллектора Ω. Для газа — глубина и давление. Мы знаем экспериментально установленные интервалы значений плотности нефти $[\delta_{bttm}^{oil}; \delta_{top}^{oil}]$ и газа $[\delta_{bttm}^{gas}; \delta_{top}^{gas}]$ [10]. Для нефти имеем интервал [0.4 г/ см^3; 1 г/ см^3]. Будем разыгрывать случайные значения плотности в точках $X \in \Omega$ на этом

интервале. В данном статистическом эксперименте рассматриваем равномерное распределение относительно $\delta(X,\Theta)$. При численной генерации псевдослучайных чисел мы использовали две функции являющиеся датчиками линейных конгруэнтных последовательностей случайных чисел.

Относительно математической модели для обратной задачи касательно источника нам следует проводить итерационный процесс для $f(X) \in F \subset U$:

$$\rho(Af(X), \bar{u}_\Gamma(X)) \longrightarrow min$$

В рамках статистической модели для обеспечения большой вероятности единственности решения и устойчивости результата мы будем проводить описанный статистический процесс концентрации плотностей $\delta(X,\Theta)$, $X \in \Omega$ с достаточным (относительно закона больших чисел) количеством экспериментов K:

$$\rho(B\delta(X,\Theta), \bar{u}_\Gamma(X)) \longrightarrow min$$

В начале каждого из K экспериментов $\delta(X,\Theta) = 0$, для $X \in V \setminus \Omega$. В случае погрешности $\tilde{u}_\Gamma(X)$ мы используем как невязку, так и сглаживающий функционал А.Н. Тихонова M^α с классическим стабилизатором.

3 Численные эксперименты

Рассмотрим куб $V = \{\xi, \eta, \zeta\}$, $\xi \in [0,1]$, $\eta \in [0,1]$, $\zeta \in [1/3, 4/3]$. Рассматривается сеточный аналог куба V^h. Предполагаем, что первый этап комбинаторной задачи состоящий в определении геометрии Ω (Ω^h) решен. Используем теперь результаты решения "генеральной"обратной граничной задачи относительно различных распределений плотности. Соответствующие результаты получены посредством статистического процесса $Bu(X)$ ($B^h u(X)$). Результаты сведены в альбом выметенных на границу плотностей.

Рассмотрим случай относительно гравитационного поля. Область Ω для листов альбома представлена различными одной и двумя связными выпуклыми областями. Сеточная аппроксимация Ω^h дается 1, 2, 4, 8, 12, 16 точками на сетке. Шаг сеток $h = 1/3$ и $h = 1/6$. Расстояние h измеряется в км. Так же представлен соответствующий сеточный альбом. Расположение объекта определяется всплесками граничной плотности δ^x_Γ, δ^y_Γ, δ^z_Γ на соответствующих гранях куба V (V^h).

Для магнитного (либо гравитационного) поля в рамках эпигенетической трактовки варианты области Ω^h представляют одноуровневые и двухуровневые скопления магнитных эпигенетических образований. Здесь плотность полагается равной 1г/см^3. Полученные альбомы можно использовать и в этом случае.

На втором этапе количество экспериментов метода Монте–Карло на $(0.4, 1)$ со структурной схемой B составляет $K = 900$ и $K = 3000$. Первый вариант K более экономичен. В случае использования интервала $(3 \times min(\delta^x_\Gamma, \delta^y_\Gamma, \delta^z_\Gamma); 1)$ мы отсеиваем большую часть области $V \setminus \Omega$ до статистических испытаний. При расчетах мы использовали как $||Af(X) - \bar{u}_\Gamma(X)||_U$, так и $||Af(X) - \bar{u}_\Gamma(X)||^2_U$. Второй вариант оказался предпочтителен. Для $\bar{\delta}_\Gamma$ и Ω заданной 1-й точкой (точечная масса) расчет займет 5 сек, погрешность по квадрату невязки не более 1% при $\epsilon_{MC} = 10^{-4}$, $K = 900$. Частота процессора составляет 1 GHz. При погрешности в $\tilde{u}$ 5% погрешность по квадрату невязки 5.1%. Для модели, где Ω заданно параллелепипедом (в случае магнитного поля двухуровневым скоплениями 12 магнитных тел) при $\tilde{\delta}_\Gamma$ заданной с погрешностью из $[6\%; 10\%]$ погрешность концентрации по квадрату невязки 4.7% при $\epsilon_{MC} = 10^{-3}$.

Заключение

В работе рассмотрены алгоритмы определения геометрических и плотностных характеристик относительно месторождений углеводородов. При определении геометрических характеристик мы проводим комплексную интерпретацию выборок относительно гравитационного и магнитного полей. Предложен алгоритм относительно задачи расчета распределения плотностей на контуре охватывающем источники аномалий. Процедура расчета распределения плотностей в источнике аномалии основана на статистических испытаниях имеющих целью минимизацию невязки либо сглаживающего функционала А.Н. Тихонова. Указанная процедура вместе с методом Монте–Карло использует статистическую регуляризацию и balayage метода А. Пуанкаре. Представлены результаты относительно нескольких вычислительных экспериментов для источников имеющих различную форму. Рассмотренные алгоритмы является экономичными и эффективными.

Автор выражает искреннюю благодарность профессору В.Г. Филатову, академику НАНУ В.И. Старостенко, академику РАН М.М. Лаврентьеву, профессору А.Г. Яголе, профессору И.Э. Степановой — за полезные рекомендации и внимание к работе, доктору физико-математических наук М.А. Шишленину, профессору А.В. Баеву, профессору Н.Н. Нефедову, профессору Ю.П. Пытьеву — за полезные замечания во время обсуждения результатов.

Список литературы

[1] Алифанов О.М. Идентификация процессов теплообмена летательных аппаратов. Москва: Машиностроение, 1979.

[2] Березкин В.М., Жбанков Ю.В., Филатов В.Г., Трайнин П.Н., Булычев Е.В., Метиев Р. Методические рекомендации по технологии площадной обработки и интерпретации гравимагнитных данных. Москва: Нефтегеофизика, 1992.

[3] Вапник В.Н. Восстановление зависимостей по эмпирическим данным. Москва: Наука, 1979.

[4] Годунов С.К. Уравнения математической физики. Москва: Наука, 1979.

[5] Годунов С.К., Рябенький В.С. Разностные схемы. Москва: Наука, 1977.

[6] Kabanikhin S. I., Gasimov Y. S., Nurseitov D. B., Shishlenin M. A., Sholpanbaev B. B., Kasenov S. E. Regularization of the continuation problem for elliptic equations // Journal of Inverse and Ill-Posed Problems. 2013. V. 21. N 6. P. 871–884.

[7] Кочетков Е.С., Смерчинская С.О., Соколов В.В. Теория вероятностей и математическая статистика. Москва: Форум, Инфа-М, 2014.

[8] Лаврентьев М.М., Романов В.Г., Шишатский С.П., Некорректные задачи математической физики и анализа. Москва: Наука, 1980.

[9] Лаврентьев М.М., Старостенко В.М. Филатов В.Г., Гласко Ю.В. и др. Применение регуляризации в гравимагниторазведке при поисках месторождений углеводородов. Москва: РГГРУ-МГРИ, 2010.

[10] Мегеря В.М., Старостенко В.И., Филатов В.Г., Зиновкин С.В. и др. Определение плотностей горных пород и геологических объектов в нефтегазоразведке. Москва: МГРИ-РГГРУ, 2011.

[11] Мегеря В.М., Старостенко В.И., Филатов В.Г., Лобанов А.Н. и др. Месторождения углеводородов неструктурного типа и их проявления в гравитационных полях. Москва: РГГРУ-МГРИ, 2011.

[12] Мегеря В.М., Филатов В.Г., Старостенко В.И., Корчагин И.Н., Лобанов А.Н., Гласко Ю.В. и др. Возможности и перспективы применения несейсмических методов для поисков скоплений углеводородов и геосолитонная концепция их образования // Геофизический журнал. Киев. 2012. Т. 34, N 3. С. 4–21.

[13] Мудрецова Е.И. (ред.) Гравиразведка. Справочник геофизика. Москва: Недра, 1981.

[14] Никитин А.А., Петров А.В., Мегеря В.М., Старостенко В.Н. и др. Оптимальная фильтрация и интропродолжение геополей с учетом вторичного магнитоминералообразования в нефтегазоразведке. Москва: НТ Пресс, 2011.

[15] Пытьев Ю.П., Шишмарев И.А. Курс теории вероятностей и математической статистики для физиков. Москва: МГУ, 1983.

[16] Анри Пуанкаре. Избранные труды. Том 3. Москва: Наука, 1974.

[17] Самарский А.А. Теория разностных схем. Москва: Наука, 1989.

[18] Самарский А.А., Вабищевич П.Н. Численные методы решения задач конвекции-диффузии. Москва: Эдиториал УРСС, 1999.

[19] Страхов В.Н. К теории плоских задач гравиметрии и магнитометрии —"аналитический мир порождаемый выметанием Пуанкаре // Журнал Известия АН СССР, серия Физика Земли. Москва. 1978. N 2. С. 47–73.

[20] Свешников А.Г., Боголюбов В.В., Кравцов В.В. Лекции по математической физике. Москва: МГУ, Наука, 2004.

[21] Соболь И.М. Метод Монте–Карло. Москва: Наука, 1985.

[22] Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. Москва: Наука, 1979.

[23] Yagola A.G., Dorofeev K.Yu. Sourcewise representation and a posteriori error estimates for ill-posed problems // Ramm A.G., Shivakumar P.N., Strauss A.V. (eds.) Fields Institute Communications: Operator Theory and Its Applications. Providence, RI: American Mathematical Society. 2000. V. 25. P. 543–550.

[24] Ягола А.Г., Степанова И.Э., Титаренко В.Н. Обратные задачи и методы их решения. Приложения к геофизике. Москва: БИНОМ, 2014.

[25] Зидаров Д.П. О решении некоторых обратных задач потенциальных полей и их применение к вопросам геофизики. София: БАН, 1980.

[26] Гласко Ю.В. Задача концентрации масс // Физика Земли. 2015. Т. 51. N 2. С. 37–43.

[27] Glasko Yuri V. Interpretation Algorithms for Hydrocarbon Deposits // Nurgaliev D., Matveeva N. (eds.) Practical and Theoretical Aspects of Geological Interpretation of Gravitational, Magnetic and Electric Fields. Switzerland: Springer. 2019. P. 113–125.

Гласко Юрий Владленович — к.ф.-м.н., ведущий математик
Научно-исследовательского вычислительного центра
Московского государственного университета им. М. В. Ломоносова;
e-mail: glaskoyv@mail.ru.
Дата поступления — 31 мая 2019 г.

КОМБИНАТОРНАЯ СТРУКТУРА 5-МЕРНОЙ МОДЕЛИ КОЛЬЦЕВОЙ ГЕННОЙ СЕТИ

В. П. Голубятников[1,2], В. С. Градов[2]

[1] *Институт математики им. С.Л.Соболева СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 514.763.81
DOI: 10.24411/9999-016A-2019-10016

Установлены условия существования двух периодических траекторий (циклов) у пятимерной блочно-линейной динамической системы, моделирующей функционирование кольцевой генной сети. Фазовый портрет системы разбивается на блоки, в каждом из которых система линейна. Это позволяет сконструировать комбинаторную схему фазового портрета и локализовать положения указанных циклов.

Ключевые слова: модели генных сетей, блочно-линейные динамические системы, дискретизация фазовых портретов, валентность блока, циклы.

Введение

Генными сетями называют наборы генов, контролирующих кинетику биохимических процессов, происходящих в организме. Моделирование функционирования этих сетей позволяет прогнозировать морфогенез различных тканей организмов, см. [1, 4, 5, 10] и цитированную там литературу. Мы изучаем один класс генных сетей, в котором вещества взаимодействуют “по кольцу”, и скорость синтеза каждого из них является монотонно убывающей функцией концентрации вещества, предшествующего ему относительно этого кольцевого порядка. Такие ингибирующие взаимодействия называют отрицательными обратными связями. Основная наша задача состоит в описании периодических траекторий динамических систем указанного типа, поскольку периодические режимы функционирования генных сетей играют важную роль в задачах биоинформатики.

1 Модель

В качестве моделей генных сетей, регулируемых только отрицательными обратными связями, рассматриваются блочно-линейные динамические системы размерности 5, симметричные относительно циклической перестановки координат:

$$\dot{X}_1 = L(X_5) - X_1;\ \ \dot{X}_2 = L(X_1) - X_2;\ \ \dot{X}_3 = L(X_2) - X_3;\ \ \dot{X}_4 = L(X_3) - X_4;\ \ \dot{X}_5 = L(X_4) - X_5, \qquad (1)$$

где переменные $X_j \geq 0$ задают концентрации веществ, участвующих в реакциях генной сети, вычитаемые описывают процессы их естественного разложения, а процессы синтеза этих веществ моделируются ступенчатой функцией L, которая определяется следующим образом:

$$L(x) = a \quad \text{для} \quad x \in (0,1); \quad L(x) = 0 \quad \text{для} \quad 1 < x; \quad a > 1.$$

Как было установлено в [8], в случае $a < 1$ система (1) циклов не имеет. В работах [2, 3, 5] рассматривались аналогичные системы других размерностей с кусочно-линейными правыми частями, в работах [4, 6, 7] с помощью различных подходов изучалось поведение траекторий подобных динамических систем с гладкими правыми частями.

Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-01-00057) и СО РАН, грант № 0314-2018-0011.

ISBN 978-5-901548-42-4

2 Дискретизация фазового портрета

Рассмотрим куб $Q = [0,a] \times [0,a] \times [0,a] \times [0,a] \times [0,a] \subset \mathbb{R}^5_+$ и точку $E = (1,1,1,1,1)$ в его внутренности. В работах [1, 7, 8] было установлено, что область Q положительно инвариантна для системы (1), и что с течением времени все траектории этой системы попадают в эту область и в дальнейшем не выходят из нее.

Аналогичные утверждения имеют место и для подобных (1) динамических систем других размерностей, их доказательства состоят в проверке знаков производных $\dot{x_j}$ на парах противоположных граней $x_j = 0$ и $x_j = a$ куба Q, см., например, [2, 9].

Разобьем Q координатными плоскостями $x_j = 1$ на 2^5 более мелких параллелепипедов, которые будем называть блоками и нумеровать бинарными мультииндексами $\epsilon = \{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4\varepsilon_5\}$, где $\varepsilon_j = 0$, если $x_j < 1$, и $\varepsilon_j = 1$, если $x_j > 1$ — как у вершин 5-мерного куба. Здесь и далее $j = 1,2,3,4,5$.

Из результатов работ [1, 2] следует, что для каждых двух соседних блоков ϵ_1 и ϵ_2, пересекающихся по 4-мерной грани, траектории всех точек внутренности этой грани переходят либо из ϵ_1 в ϵ_2, либо из ϵ_2 в ϵ_1. Тем самым граф с вершинами $\{\varepsilon_1\varepsilon_2\ldots\varepsilon_5\}$ можно ориентировать согласно направлению переходов траекторий из блока в блок. Такой ориентированный граф G конструируется и для любой динамической системы вида $\dot{x}_m = f_m(x_{m-1}) - k_m x_m,\ m = 1,2,\ldots,N$, аналогичной системе (1). Здесь все функции f_m монотонны (возрастающие либо убывающие) и либо гладкие, либо ступенчатые; убывающие функции моделируют отрицательные обратные связи, а возрастающие — положительные, см. [1, 5, 9]. Следует подчеркнуть, что если в таких системах функции f_m не являются монотонными, то поведение траекторий может оказаться хаотическим, см. [10].

Как было показано в [2, 8], если $k_m = 1$ при всех m, то в каждом блоке ϵ траектории таких блочно-линейных систем прямолинейны и описываются гомотетиями с центрами в вершинах параллелепипеда Q. В случаях, когда функции f_m различны, динамическая система не имеет симметрий относительно циклической перестановки координат, и область Q кубом не является.

Назовем валентностью блока ϵ количество его граней, через которые траектории системы могут из него выходить, то есть валентность блока — это количество ребер, выходящих из соответствующей вершины ориентированного графа G, см. [1, 9].

Среди 32 блоков разбиения области Q есть блоки валентностей 1, 3 и 5. Блоки с мультииндексами $\{00000\}$ и $\{11111\}$ имеют валентность $V = 5$, в них траектории системы (1) попадают только из области $\mathbb{R}^5_+ \setminus Q$.

В работах [7, 8] были установлены условия существования цикла $\mathcal{C}_1$ систем вида (1), проходящего через блоки валентности 1, а также изучались вопросы единственности и устойчивости этого цикла. Теперь мы рассматриваем блоки валентности $V = 3$, для которых переходы из блока в блок вдоль траекторий системы (1) описываются следующими диаграммами:

$$\begin{array}{ccccccccc}
\{11110\} & \longrightarrow & \{11100\} & \longrightarrow & \{11101\} & \longrightarrow & \{11001\} & \longrightarrow & \{11011\} \\
\uparrow & & & & & & & & \downarrow \\
\{01110\} & \longleftarrow & \{01111\} & \longleftarrow & \{00111\} & \longleftarrow & \{10111\} & \longleftarrow & \{10011\}
\end{array} \tag{2}$$

$$\begin{array}{ccccccccc}
\{11100\} & \longrightarrow & \{11000\} & \longrightarrow & \{11001\} & \longrightarrow & \{10001\} & \longrightarrow & \{10011\} \\
\uparrow & & & & & & & & \downarrow \\
\{01100\} & \longleftarrow & \{01110\} & \longleftarrow & \{00110\} & \longleftarrow & \{00111\} & \longleftarrow & \{00011\}
\end{array} \tag{3}$$

$$\begin{array}{ccccccccc}
\{11000\} & \longrightarrow & \{10000\} & \longrightarrow & \{10001\} & \longrightarrow & \{00001\} & \longrightarrow & \{00011\} \\
\uparrow & & & & & & & & \downarrow \\
\{01000\} & \longleftarrow & \{01100\} & \longleftarrow & \{00100\} & \longleftarrow & \{00110\} & \longleftarrow & \{00010\}
\end{array} \tag{4}$$

2.1 Вспомогательные алгебраические утверждения

Пусть $a > 1$. Рассмотрим последовательность многочленов:

$$P_1(x) = x + \left(1 - \frac{a^2}{a-1}\right), \qquad P_2(x) = x^2 + \left(1 - \frac{a^2}{a-1}\right)x + 1,$$

$$P_n(x) = x^n + \left(1 - \frac{a^2}{a-1}\right)x^{n-1} + x^{n-2} + \ldots + x + 1, \quad n > 1.$$

При $n \geq 2$ выполняются рекуррентные соотношения (из которых вытекают следующие леммы):

$$P_n(x) = xP_{n-1}(x) + 1. \tag{5}$$

Лемма 1. 1) *Если $n \geq 2$ нечетное, то $P_n(x) = P_1(x)$ при $x = \frac{a}{a-1}$ и при $x = a$. Если n четное, то $P_n(x) = P_1(x)$ при $x = -1$, $x = \frac{a}{a-1}$ и при $x = a$.*

2) *Если $2 \leq k < n$ и $n - k$ нечетное, то $P_n(x) = P_k(x)$ при $x = 0$, $x = \frac{a}{a-1}$ и при $x = a$. Если $n - k$ четное, то $P_n(x) = P_k(x)$ при $x = -1$, $x = 0$, $x = \frac{a}{a-1}$ и при $x = a$.*

Лемма 2. 1) *Если $a > 1$, то многочлен $P_2(x)$ имеет единственный корень на интервале $(0, 1)$;*
2) *Если $a > 1$, $a \neq 2$, то многочлен $P_3(x)$ имеет единственный корень на интервале $(0, 1)$;*
3) *Если $n \geq 4$ и $a^2 > (n + 1)(a - 1)$, то многочлен $P_n(x)$ имеет единственный корень на интервале $(0, 1)$.*

Из предыдущих лемм следует, что для $n \geq 2$ выполняются неравенства

$$P_n(0) = 1 > 0, \quad P_n(a) = -\frac{1}{a-1} < 0, \quad P_n\left(\frac{a}{a-1}\right) = -(a-1) < 0.$$

Значит $P_n(x)$ имеет два положительных корня, причем один из корней больше единицы, так как $a > 1$ и $\frac{a}{a-1} > 1$ при $a > 1$. Если $P_n(1) = n + 1 - \frac{a^2}{a-1} < 0$, то на интервале $(0, 1)$ этот многочлен имеет единственный корень. Если $P_n(1) > 0$, то оба его положительных корня лежат в интервалах $(1, a)$ и (a, x_2), где $x_2 > a$ является корнем многочлена $P_2(x)$. Так как $P_n(1) > 0$, $P_n(a) < 0$, $P_n\left(\frac{a}{a-1}\right) < 0$ и $P_n(x_2) > 0$, то оба интервала $(1, a)$ и (a, x_2) содержат корни многочлена $P_n(x)$.

Лемма 3. *Пусть $n \geq 1$ и $x^{(n)}, x^{(n+1)} \in (0, 1)$ — корни многочленов $P_n(x)$ и $P_{n+1}(x)$ соответственно, тогда $x^{(n+1)} > x^{(n)}$.*

3 Поиски циклов системы (1)

1. Рассмотрим диаграмму (2) и введем обозначения:

$$F_0^{[h]} = \{11110\} \cap \{11100\},\ F_1^{[h]} = \{11100\} \cap \{11101\},\ F_2^{[h]} = \{11101\} \cap \{11001\}.$$

Пусть $X = (X_1, X_2, X_3, X_4, X_5)$. Если $X \in F_0^{[h]}$, то координаты этой точки удовлетворяют условиям:

$$X_1 > 1, \quad X_2 > 1, \quad X_3 > 1, \quad X_4 = 1, \quad X_5 < 1. \tag{6}$$

Если $X \in F_1^{[h]}$, то выполняются условия:

$$X_1 > 1, \quad X_2 > 1, \quad X_3 > 1, \quad X_4 < 1, \quad X_5 = 1. \tag{7}$$

Если $X \in F_2^{[h]}$, то выполняются условия: $X_1 > 1, \quad X_2 > 1, \quad X_3 = 1, \quad X_4 < 1, \quad X_5 > 1.$

Теорема 1. *Если $a > \dfrac{5 + \sqrt{5}}{2}$, то система (1) имеет единственный цикл, симметричный относительно циклической перестановки координат и проходящий через блоки диаграммы (2).*

Доказательство. Траектория точки $X^{(0)} \in F_0^{[h]} \subset \{11100\}$ описывается в этом блоке уравнениями

$$X_1 = a - e^{-kt}\left(a - X_1^{(0)}\right),\ X_2 = e^{-kt}X_2^{(0)},\ X_3 = e^{-kt}X_3^{(0)}, X_4 = e^{-kt}X_4^{(0)},\ X_5 = a - e^{-kt}\left(a - X_5^{(0)}\right).$$

Найдем время t_1, за которое точка $X^{(0)}$ перейдет вдоль своей траектории на грань $F_1^{[h]}$, и координаты точки $X^{(1)} \in F_1^{[h]}$, в которую она попадет:

$$\begin{gathered} X_1^{(1)} = a - \frac{\left(a - X_1^{(0)}\right)(a-1)}{a - X_5^{(0)}},\ X_2^{(1)} = \frac{X_2^{(0)}(a-1)}{a - X_5^{(0)}}, \\ X_3^{(1)} = \frac{X_3^{(0)}(a-1)}{a - X_5^{(0)}},\ X_4^{(1)} = \frac{a-1}{a - X_5^{(0)}},\ X_5^{(1)} = 1,\ e^{-t_1} = \frac{a-1}{a - X_5^{(0)}}. \end{gathered} \tag{8}$$

Аналогично найдем время t_2 и координаты $X^{(2)} \in F_2^{[h]}$, в которую попадает точка $X^{(1)}$:

$$X_1^{(2)} = \frac{a\left(a - X_5^{(0)}\right)}{X_3^{(0)}\left(a - X_5^{(0)}\right)} - \frac{\left(a - X_1^{(0)}\right)}{X_3^{(0)}}, \quad X_2^{(2)} = \frac{X_2^{(0)}}{X_3^{(0)}}, \quad X_3^{(2)} = 1,$$

$$X_4^{(2)} = \frac{1}{X_3^{(0)}}, \quad X_5^{(2)} = a - \frac{a - X_5^{(0)}}{X_3^{(0)}}, \quad e^{-t_2} = \frac{a - X_5^{(0)}}{X_3^{(0)}(a-1)}.$$

Для существования цикла, симметричного относительно циклической перестановки координат, необходимы и достаточны условия: $X_j^{(2)} = X_{j+1}^{(0)}$ при $j = 1, 2, 3, 4$, $X_5^{(2)} = X_1^{(0)}$, то есть

$$X_2^{(0)} = \frac{a\left(a - X_5^{(0)}\right)}{X_3^{(0)}(a-1)} - \frac{\left(a - X_1^{(0)}\right)}{X_3^{(0)}}, \; X_3^{(0)} = \frac{X_2^{(0)}}{X_3^{(0)}}, \; X_4^{(0)} = 1, \; X_5^{(0)} = \frac{1}{X_3^{(0)}}, \; X_1^{(0)} = a - \frac{a - X_5^{(0)}}{X_3^{(0)}}. \tag{9}$$

Из уравнений (9) получаем координаты точки $X^{(0)}$, для которых надо проверить условия (6), а для соотношений (8) нужно будет проверить условия (7). Если все указанные условия выполнены, то через точку $X^{(0)}$ проходит симметричная относительно циклической перестановки координат периодическая траектория системы (1). Введем обозначение $y := X_5^{(0)}$. Из уравнений (9) следует, что $(y-1)\cdot P_4(y) = 0$. Согласно лемме 2, это уравнение имеет единственный корень $y_1 \in (0,1)$, если $a > 1$ и $a^2 > 5(a-1)$, то есть при

$$a \in \left(1, \frac{5-\sqrt{5}}{2}\right) \cup \left(\frac{5+\sqrt{5}}{2}, \infty\right) \tag{10}$$

Можно проверить, что условия (6) выполняются только в правом интервале (10). Проверим теперь условия (7). Очевидно, что $X_4^{(1)} < 1$ при $y_1 < 1$. Неравенство $X_3^{(1)} > 1$ следует из $X_1^{(0)} > 1$. Значит неравенство $X_2^{(1)} > 1$ эквивалентно неравенству $X_3^{(1)} > y_1$, и неравенство $X_1^{(1)} > 1$ эквивалентно $X_5^{(0)} < X_1^{(0)}$.

Таким образом, при $a > \dfrac{5+\sqrt{5}}{2}$ через блоки диаграммы (2) проходит цикл $\mathcal{C}_3^{[h]}$ системы (1), симметричный относительно циклической перестановки координат. □

Здесь нижний индекс 3 означает, что цикл $\mathcal{C}_3^{[h]}$ проходит по блокам валентности 3.

В теореме 1, так же, как и в последующих рассуждениях, дробно-линейные функции, описывающие сдвиги вдоль траекторий системы (1) с грани F_j на грань F_{j+1}, задают проективные преобразования, которые переводят прямолинейные отрезки в прямолинейные отрезки, что позволяет строить в Q инвариантные многогранные поверхности, содержащие циклы этой системы, см. [2, 3].

2. Рассмотрим диаграмму (3) и введем обозначения:

$$F_0^{[m]} = \{11100\} \cap \{11000\}, \; F_1^{[m]} = \{11000\} \cap \{11001\}, \; F_2^{[m]} = \{11001\} \cap \{10001\}.$$

Если $X \in F_0^{[m]}$, то координаты этой точки удовлетворяют условиям:

$$X_1 > 1, \quad X_2 > 1, \quad X_3 = 1, \quad X_4 < 1, \quad X_5 < 1. \tag{11}$$

Если $X \in F_1^{[m]}$, то выполнены условия: $X_1 > 1$, $X_2 > 1$, $X_3 < 1$, $X_4 < 1$, $X_5 = 1$.
Если $X \in F_2^{[m]}$, то выполнены условия: $X_1 > 1$, $X_2 = 1$, $X_3 < 1$, $X_4 < 1$, $X_5 > 1$.

Теорема 2. *Если $a > 1$, то система* (1) *не имеет цикла, симметричного относительно циклической перестановки координат и проходящего через блоки диаграммы* (3).

Доказательство. Для точки $X^{(0)} \in F_0^{[m]} \subset \{11000\}$ найдем время t_1, за которое эта точка перейдет вдоль своей траектории на грань $F_1^{[m]}$ в точку $X^{(1)} \in F_1^{[m]}$, с координатами:

$$X_1^{(1)} = a - \frac{(a - X_1^{(0)})(a-1)}{a - X_5^{(0)}}, \; X_2^{(1)} = \frac{X_2^{(0)}(a-1)}{a - X_5^{(0)}}, \; X_3^{(1)} = \frac{X_3^{(0)}(a-1)}{a - X_5^{(0)}},$$

$$X_4^{(1)} = a - \frac{(a - X_4^{(0)})(a-1)}{a - X_5^{(0)}}, \; X_5^{(1)} = 1, \; e^{-t_1} = \frac{a-1}{a - X_5^{(0)}}.$$

Аналогично найдем t_2 и координаты точки $X^{(2)} \in F_2^{[m]}$, в которую переходит вдоль траектории точка $X^{(1)}$:

$$X_1^{(2)} = \frac{a(a-X_5^{(0)})}{X_2^{(0)}(a-1)} - \frac{(a-X_1^{(0)})}{X_2^{(0)}}, \quad X_2^{(2)} = 1, \quad X_3^{(2)} = \frac{1}{X_2^{(0)}},$$
$$X_4^{(2)} = a - \frac{a-X_4^{(0)}}{X_2^{(0)}}, \quad X_5^{(2)} = a - \frac{a-X_5^{(0)}}{X_2^{(0)}}, \quad e^{-t_2} = \frac{a-X_5^{(0)}}{X_2^{(0)}(a-1)}.$$

Как и в теореме 1, для существования цикла, симметричного относительно циклической перестановки координат, необходимо и достаточно чтобы $X_j^{(2)} = X_{j+1}^{(0)}$, $j = 1,2,3,4$, $X_5^{(2)} = X_1^{(0)}$. Отсюда следует, что:

$$X_2^{(0)} = \frac{a(a-X_5^{(0)})}{X_2^{(0)}(a-1)} - \frac{(a-X_1^{(0)})}{X_2^{(0)}},\ X_3^{(0)} = 1,\ X_4^{(0)} = \frac{1}{X_2^{(0)}},\ X_5^{(0)} = a - \frac{a-X_4^{(0)}}{X_2^{(0)}},\ X_1^{(0)} = a - \frac{a-X_5^{(0)}}{X_2^{(0)}}. \quad (12)$$

Выразив $X_1^{(0)}, X_2^{(0)}, X_5^{(0)}$ через $y := X_4^{(0)}$, получим уравнение $P_4(y) = 0$, которое при условии (10) имеет единственный корень $y_2 \in (0,1)$. Из (12) следует, что $X_2^{(0)} > 1$. Неравенства $X_1^{(0)} > 1$ и $X_5^{(0)} < 1$ сводятся к неравенствам:

$$a - (a-y_2)y_2^2 > 1, \qquad a - (a-y_2)y_2 < 1. \quad (13)$$

Если $a \geq 2$, то второе неравенство выполнено при $y_2 \in [1, a-1)$, что противоречит условию $y_2 < 1$. Если же $1 < a < 2$, то это неравенство выполнено при $y_2 \in (a-1, 1)$, но тогда

$$P_4(a-1) = -a^3 + 3a^2 - 2a + 1 = -(a-1)^3 + a > 0.$$

Значит из $a < 2$ следует,что $X_5^{(0)} < 1$. Пусть теперь $a \in \left(1, \frac{5-\sqrt{5}}{2}\right)$. Первое из неравенств (13) сводится к неравенству $y_2^2 - (a-1)(y_2+1) < 0$, которое выполнено при $y_2 \in (0,b)$, где $2b = a-1+\sqrt{(a-1)^2+4(a-1)} < 2$. Проверим, может ли в этом интервале содержаться корень уравнения $P_4(y) = 0$. Допустим, что $P_4(b) < 0$. Поскольку $P_3(b) = -(a-1)$, из соотношения (5) следует, что

$$P_4(b) = -\frac{a-1+\sqrt{(a-1)^2+4(a-1)}}{2}(a-1) + 1 < 0.$$

Значит если $a \in \left(1, \frac{5-\sqrt{5}}{2}\right)$, то $\sqrt{(a-1)^2+4(a-1)} > \dfrac{-a^2+2a+1}{a-1} > 0$, и тогда $f(a) = a^3 - 2a^2 + a - 1 > 0$. Покажем, что при $a \in \left(1, \frac{5-\sqrt{5}}{2}\right)$ это неравенство не выполнено. Из формулы $f'(a) \equiv 3a^2 - 4a + 1$ видно, что на интервале $(1, \frac{3}{2})$ функция $f(a)$ монотонно растет, и значит $P_4(b) > 0$ при $a \in \left(1, \frac{5-\sqrt{5}}{2}\right)$.

Поэтому на интервале $(0,b)$ корней у многочлена $P_4(y)$ нет, и условия (11) не выполняются при всех $a > 1$. Следовательно, система (1) не имеет циклов, симметричных относительно циклической перестановки координат и проходящих по блокам диаграммы (3). □

3. Рассмотрим теперь диаграмму (4) и введем обозначения:

$$F_0^{[b]} = \{10000\} \cap \{10001\},\ F_1^{[b]} = \{10001\} \cap \{00001\},\ F_2^{[b]} = \{00001\} \cap \{00011\}$$

Координаты точки $X \in F_0^{[b]}$ удовлетворяют условиям:

$$X_1 > 1, \quad X_2 < 1, \quad X_3 < 1, \quad X_4 < 1, \quad X_5 = 1. \quad (14)$$

Если $X \in F_1^{[b]}$, то выполняются условия:

$$X_1 = 1, \quad X_2 < 1, \quad X_3 < 1, \quad X_4 < 1, \quad X_5 > 1. \quad (15)$$

Если $X \in F_2^{[b]}$, то выполняются условия: $X_1 < 1, \quad X_2 < 1, \quad X_3 < 1, \quad X_4 = 1, \quad X_5 > 1.$

Теорема 3. *Если* $1 < a < \dfrac{5-\sqrt{5}}{2}$, *то система* (1) *имеет единственный цикл, симметричный относительно циклической перестановки координат и проходящий через блоки диаграммы* (4).

Доказательство. Траектория точки $X^{(0)} \in F_0^{[b]} \subset \{10001\}$ попадает на грань $F_1^{[b]}$ в точку $X^{(1)}$ с координатами

$$\begin{gathered} X_1^{(1)} = 1, \quad X_2^{(1)} = \frac{X_2^{(0)}}{X_1^{(0)}}, \quad X_3^{(1)} = a - \frac{a - X_3^{(0)}}{X_1^{(0)}}, \\ X_4^{(1)} = a - \frac{a - X_4^{(0)}}{X_1^{(0)}}, \quad X_5^{(1)} = a - \frac{a-1}{X_1^{(0)}}, \quad e^{-t_1} = \frac{1}{X_1^{(0)}}. \end{gathered} \tag{16}$$

Аналогично найдем время t_2 и координаты точки $X^{(2)} \in F_2^{[b]}$.

$$X_1^{(2)} = \frac{X_1^{(0)}(a-1)}{a - X_4^{(0)}}, \quad X_2^{(2)} = a - \frac{X_1^{(0)}(a-1)}{a - X_4^{(0)}} \left(a - \frac{X_2^{(0)}}{X_1^{(0)}} \right), \quad X_4^{(2)} = 1,$$

$$X_3^{(2)} = a - \frac{(a-1)\left(a - X_3^{(0)}\right)}{a - X_4^{(0)}}, \quad X_5^{(2)} = a - \frac{(a-1)^2}{a - X_4^{(0)}}, \quad e^{-t_2} = \frac{(a-1)X_1^{(0)}}{a - X_4^{(0)}}.$$

Для существования цикла, симметричного относительно циклической перестановки координат, необходимо и достаточно, чтобы $X_j^{(2)} = X_{j+1}^{(0)}$, $j = 1, 2, 3, 4$, $X_5^{(2)} = X_1^{(0)}$, и поэтому

$$\begin{gathered} X_2^{(0)} = \frac{X_1^{(0)}(A-1)}{A - X_4^{(0)}}, \quad X_3^{(0)} = a - \frac{X_1^{(0)}(a-1)}{a - X_4^{(0)}} \left(a - \frac{X_2^{(0)}}{X_1^{(0)}} \right), \\ X_4^{(0)} = a - \frac{(a-1)\left(a - X_3^{(0)}\right)}{a - X_4^{(0)}}, \quad X_5^{(0)} = 1, \quad X_1^{(0)} = a - \frac{(a-1)^2}{a - X_4^{(0)}}. \end{gathered} \tag{17}$$

Исключая из этих соотношений $X_4^{(0)}$, $X_3^{(0)}$ и $X_2^{(0)}$, получим уравнение

$$w^5 - \frac{a^2}{a-1} w^4 + \frac{a^2}{a-1} w^3 - 1 = 0; \qquad w := \frac{a - X_1^{(0)}}{a-1}.$$

Как и в предыдущих теоремах, отсюда следует уравнение $P_4(w) = 0$, которое имеет единственный корень $w_1 \in (0,1)$ при условии (10). Очевидно, что $X_1^{(0)} > 1$. Проверим оставшиеся неравенства (14). Поскольку $X_2^{(0)} = e^{-t_2} < 1$, из соотношений (17) видно, что: $\dfrac{a - X_4^{(0)}}{a-1} = \dfrac{a-1}{a - X_1^{(0)}} = \dfrac{1}{w_1} > 1$, и тогда $X_4^{(0)} < 1$, а также $a - X_3^{(0)} = \dfrac{\left(a - X_4^{(0)}\right)^2}{a-1} = \dfrac{(a-1)^3}{\left(a - X_1^{(0)}\right)^2} = \dfrac{a-1}{w_1^2}$. Поэтому $w_1^2 = \dfrac{a-1}{a - X_3^{(0)}} < 1$ и $X_3^{(0)} < 1$, и значит из условия (10) вытекают условия (14).

Проверим условия (15). Из системы (16) следует, что $X_2^{(1)} < 1$, так как $X_1^{(0)} > 1$, $X_2^{(0)} < 1$, и что $X_5^{(1)} > 1$, поскольку $\dfrac{a - X_5^{(1)}}{A-1} < 1$. Учитывая соотношения на $X_3^{(0)}$, $X_4^{(0)}$ из (17), перепишем условия (16) в виде

$$\frac{a - X_3^{(1)}}{a-1} = \frac{1}{w_1^2 X_1^{(0)}}, \qquad \frac{a - X_4^{(1)}}{a-1} = \frac{1}{w_1 X_1^{(0)}}.$$

Найдем значения a, при которых оба эти выражения больше единицы. Неравенство $w_1 X_1^{(0)} < 1$ эквивалентно

$$w_1^2 - \frac{a}{a-1} w_1 + \frac{1}{a-1} > 0. \tag{18}$$

Если $a > 2$, то это неравенство выполнено на интервале $\left(0, \frac{1}{a-1}\right)$. Из Леммы 3 следует, что если корень многочлена $P_3(w)$ меньше единицы, то он лежит в интервале $\left(\frac{1}{a-1}, 1\right)$, и поэтому там содержится и корень многочлена $P_4(w)$. Значит $P_4(w)$ не имеет корней в интервале $\left(0, \frac{1}{a-1}\right)$ при $a > 2$.

Если $1 < a < 2$, то есть a принадлежит левому интервалу из условия (10), то неравенство (18) выполнено при $w_1 \in (0,1)$. Следовательно, при $a \in \left(1, \frac{5-\sqrt{5}}{2}\right)$ выполняются условия (15), и значит в условиях теоремы 3 у системы (1) существует цикл $\mathcal{C}_3^{[b]}$, симметричный относительно циклической перестановки координат и проходящий по имеющим валентность 3 блокам диаграммы (4). □

Заключение

Из приведенных рассуждений следует, что все перечисленные циклы $\mathcal{C}_1$, $\mathcal{C}_3^{[b]}$, $\mathcal{C}_3^{[h]}$ содержатся в инвариантных кусочно-линейных поверхностях, симметричных относительно циклической перестановки переменных и имеющих одну вершину — точку E с координатами $(1,1,1,1,1)$. При выполнении условий (10) динамическая система (1) имеет по крайней мере два цикла, которые проходят по блокам валентностей 1 и 3, соответственно.

Авторы выражают искреннюю признательность Н.Б. Аюповой и В.В. Иванову за полезные обсуждения.

Список литературы

[1] Акиньшин А. А., Голубятников В. П. Циклы в симметричных динамических системах // Сибирский журнал чистой и прикладной математики. 2012. Т. 12, №2. С. 3–12.

[2] Аюпова Н. Б., Голубятников В. П. О единственности цикла в несимметричной трехмерной модели молекулярного репрессилятора // Сибирский журнал индустриальной математики. 2014. Т. 17, № 1. С. 3–7.

[3] Аюпова Н.Б., Голубятников В.П. О двух классах нелинейных динамических систем. Четырехмерный случай // Сибирский математический журнал. 2015. Т. 56, N 2. С. 282 – 289.

[4] Elowitz M. B., Leibler S. A synthetic oscillatory network of transcriptional regulators // Nature. 2000. V. 403. P. 335–338.

[5] Glass L., Pasternack J. S. Stable oscillations in mathematical models of biological control systems // Journal of Mathematical Biology. 1978. V.6. P. 207–223.

[6] Глызин С. Д., Колесов А.Ю., Розов Н.Х. Явления буферности в кольцевых генных сетях // Теоретическая и математическая физика. 2016. Т. 187, N. 3. С. 560–579.

[7] Голубятников В. П., Голубятников И. В., Лихошвай В. А. О существовании и устойчивости циклов в пятимерных моделях генных сетей // Сибирский журнал вычислительной математики 2010. Т. 13, № 4, С. 403–411.

[8] Голубятников В. П., Иванов В. В. Циклы в нечетномерных моделях кольцевых генных сетей // Сибирский журнал индустриальной математики. 2018. Т. 21, N 4. С. 28–38.

[9] Казанцев М. В. О некоторых свойствах графов доменов динамических систем // Сибирский журнал индустриальной математики. 2015. Т. 18, № 4. С. 42–49.

[10] Likhoshvai V. A., Fadeev S. I., Kogai V. V., Khlebodarova T. M. Alternative splicing can lead to chaos // Journal of Bioinformatics and Computational Biology. 2015. V l3, N 1. P. 1540003-1 — 1540003-25.

Голубятников Владимир Петрович — д.ф.-м.н., гл. науч.сотр. Института математики им. С.Л.Соболева СО РАН;
e-mail: vladimir.golubyatnikov1@fulbrightmail.org;
Градов Вячеслав Сергеевич — магистрант ММФ Новосибирского государственного университета;
e-mail: gnets2008@outlook.com.
Дата поступления — 30 апреля 2019 г.

О ЦИКЛАХ В БЛОЧНО-ЛИНЕЙНЫХ МОДЕЛЯХ КОЛЬЦЕВЫХ ГЕННЫХ СЕТЕЙ

В. П. Голубятников[1,2], Л. С. Минушкина[2]

[1] *Институт математики им. С. Л. Соболева СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 514.763.81
DOI: 10.24411/9999-016A-2019-10017

Установлены достаточные условия существования цикла у блочно-линейного несимметричного варианта динамической системы типа Еловица-Лейблера, моделирующей функционирование кольцевой генной сети. Для такого цикла описана его инвариантная окрестность в фазовом портрете этой системы.

Ключевые слова: Блочно-линейные динамические системы, фазовые портреты, инвариантные области, кольцевая генная сеть, отображение Пуанкаре, монотонность, циклы.

Введение

В ряде недавних публикаций [4, 5, 9, 11] исследовались нелинейные динамические системы биохимической кинетики, моделирующие функционирование кольцевых генных сетей, общая структура которых была описана в [3,10,12] вместе с их биологической интерпретацией. В работах [5,10,11] рассматривался случай, когда указанные динамические системы симметричны относительно циклических перестановок пар переменных, соответствующих парам мРНК и кодируемых ими белков. В настоящей публикации, так же, как и в [2,4,9], такие симметрии в генных сетях и в их моделях не предполагаются.

Основные результаты перечисленных статей заключались в описании условий существования циклов у этих моделей генных сетей.

1 Модель

Мы изучаем аналогичные модели генных сетей, функционирование которых описывается кусочно-линейными системами вида

$$\frac{dm_1}{dt} = L_1(p_3) - k_1 m_1; \qquad \frac{dp_1}{dt} = \Gamma_1(m_1) - l_1 p_1; \qquad \frac{dm_2}{dt} = L_2(p_1) - k_2 m_2;$$

$$\frac{dp_2}{dt} = \Gamma_2(m_2) - l_2 p_2; \qquad \frac{dm_3}{dt} = L_3(p_2) - k_3 m_3; \qquad \frac{dp_3}{dt} = \Gamma_3(m_3) - l_3 p_3. \tag{1}$$

Здесь все коэффициенты k_j, l_j постоянны и положительны, функции L_j моделируют отрицательные обратные связи и представляют собой убывающие ступеньки:

$$L_j(x_{j-1}) = a_j \cdot k_j > 0 \quad \text{при } 0 \le x_{j-1} < 1; \;\; L_j(x_{j-1}) = 0 \quad \text{при } 1 \le x_{j-1},$$

а функции Γ_j описывают положительные обратные связи — возрастающие ступеньки:

$$\Gamma_j(x_{j-1}) = 0 \quad \text{при } 0 \le x_{j-1} < 1; \;\; \Gamma_j(x_{j-1}) = b_j \cdot l_j > 0 \quad \text{при } 1 \le x_{j-1}, \quad \text{где } j = 1, 2, 3.$$

Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-01-00057) и СО РАН, грант № 0314-2018-0011.

ISBN 978-5-901548-42-4

В симметричном гладком случае $k_j = 1,\ L_j(w) = \alpha \cdot (1 + w^m)^{-1},\ \Gamma_j(u) = \mu u,\ l_j = \mu$, где все параметры положительны, система (1) была предложена в [10] для моделирования генной сети, связывающей три белка с концентрациями $p_1(t)$, $p_2(t)$, $p_3(t)$, репрессирующих друг друга по кольцевой схеме $1 \to 2 \to 3 \to 1$, и соответствующие мРНК с концентрациями $m_1(t)$, $m_2(t)$, $m_3(t)$, см. также [3,12]. В бóльших размерностях аналоги такой динамической системы изучались в [2,4,9].

Система (1) представляет собой блочно-линейный несимметричный вариант системы, рассмотренной в [10]. Положительные параметры a_j и b_j пропорциональны скоростям реакций синтеза веществ с концентрациями m_j, p_j, соответственно. Эти неотрицательные переменные имеют ту же самую биологическую интерпретацию, что и в [3,10].

Вычитаемые во всех этих уравнениях моделируют процессы естественного разложения компонент генной сети, коэффициенты k_j, l_j характеризуют скорости процессов разложения, монотонно убывающее слагаемое $L_1(p_3)$ в первом уравнении системы (1) описывает зависимость скорости синтеза мРНК с концентрацией m_1 от концентрации p_3 белка, "предшествующего" этой мРНК в кольцевой генной сети; таким же образом интерпретируются и остальные уравнения этой системы. В дальнейшем полагаем, что $j = 1,2,3$ и что $j - 1 = 3$ при $j = 1$.

Отметим, что если в правых частях уравнений системы вида (1) присутствуют немонотонные функции, например, унимодальные функции вида $f(x) = a \cdot (1 + x^n)^{-1},\ f(x) = x \cdot e^{-nx}$ и т.п., то поведение траекторий такой системы может оказаться хаотическим, см. [13].

2 Результаты

Целью настоящей работы является рассмотрение несимметричного случая — все функции L_j вообще говоря различны, различны также и все функции Γ_j, и все коэффициенты k_j, l_j. Основным нашим результатом является

Теорема 1. *Если $a_j > 1$ и $b_j > 1$ при всех j, то динамическая система* (1) *имеет цикл.*

3 Вспомогательные построения

Рассмотрим параллелепипед $Q = [0, a_1] \times [0, b_1] \times [0, a_2] \times [0, b_2] \times [0, a_3] \times [0, b_3] \subset \mathbb{R}^6_+$. Так же, как и в [1,6,7], можно показать, что если у системы (1) есть цикл, то он должен содержаться во внутренности области Q, и что что эта область положительно инвариантна для этой системы.

При описании геометрии фазовых портретов блочно-линейных систем вида (1) важную роль играет точка E с координатами $(1,1,1,1,1,1)$, поскольку при $m_j = 1$, $p_j = 1$ правые части уравнений этой системы имеют разрывы первого рода. Во всех дальнейших наших построениях мы будем следить за тем, чтобы рассматриваемые нами траектории не проходили бы через эту точку, а также через другие пересечения гиперплоскостей $m_j = 1$ и $p_j = 1$.

Так же, как и в [1,6,7], где рассматривались подобные "биллиардообразные" динамические системы других типов, разобьем параллелепипед Q гиперплоскостями, параллельными координатным и проходящими через точку E, на более мелкие параллелепипеды, которые мы будем называть блоками и обозначать бинарными мультииндексами:

$$\{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4\varepsilon_5\varepsilon_6\}. \tag{2}$$

Всюду будем полагать, что $\varepsilon_{2j-1} = 0$, если для всех точек рассматриваемого блока выполняется неравенство $m_j \le 1$, и $\varepsilon_{2j} = 0$, если для всех точек рассматриваемого блока выполняется неравенство $p_j \le 1$, и что $\varepsilon_{2j-1} = 1$, если для всех его внутренних точек выполняется противоположное неравенство $m_j > 1$, и $\varepsilon_{2j} = 1$, если в для всех таких точек $p_j > 1$, см. [2,4,9], где было установлено, что для любой пары B_1, B_2 смежных блоков разбиения (3) траектории всех точек их общей грани $F = B_1 \cap B_2$ проходят через эту грань только в одну сторону, либо из блока B_1 в B_2, либо из блока B_2 в B_1. В тех же работах было введено определение: *Валентностью блока B* назовем количество смежных с ним блоков (2), в которые могут переходить траектории точек блока B.

Для доказательства указанного утверждения достаточно проверить знак производной dm_j/dt (либо dp_j/dt) на грани F, то-есть в точках гиперплоскости $m_j = 1$ (либо, соответственно, $p_j = 1$), разделяющей блоки B_1 и B_2.

Аналогичными рассуждениями проверяется, см., например, [1, 2] что все траектории динамической системы (1), начинающиеся в блоке {110011}, проходят через имеющие валентность один блоки разбиения (2) в соответствии с циклической диаграммой

$$\begin{array}{ccccccc}
\{110011\} & \longrightarrow & \{010011\} & \longrightarrow & \{000011\} & \longrightarrow & \{001011\} \\
\uparrow & & & & & & \downarrow \\
\{110010\} & & & & & & \{001111\} \\
\uparrow & & & & & & \downarrow \\
\{110000\} & & & & & & \{001101\} \\
\uparrow & & & & & & \downarrow \\
\{110100\} & \longleftarrow & \{111100\} & \longleftarrow & \{101100\} & \longleftarrow & \{001100\}
\end{array} \qquad (3)$$

Лемма 1. *Из каждого блока разбиения* (2)*, перечисленного в диаграмме* (3)*, траектории системы* (1) *могут перейти только в тот соседний блок, куда указывает соответствующая стрелка диаграммы* (3)*.*

Пусть W — внутренность объединения всех этих двенадцати блоков; область W является положительно инвариантной для системы (1), в ней мы и будем искать цикл $\mathcal{C}$ этой системы.

Из доказательства теоремы мы увидим, что траектории всех внутренних точек области W переходят из блока в блок только через внутренние точки разделяющих их пятимерных граней и не проходят через их грани, имеющие меньшую размерность. Следовательно, область W является инвариантной окрестностью искомого цикла $\mathcal{C}$. Отметим, что в диаграмме (3) перечислены все блоки валентности 1 системы (1).

Для гладких монотонно убывающих функций L_j и $\Gamma_j(u) = \mu_j u$, и аналогичных (1) динамических систем других размерностей подобные рассуждения были проведены в [2, 4, 9].

4 Описание траекторий точек области W

Общие грани каждой пары соседних блоков, перечисленных в диаграмме (3), будем обозначать так:

$F_0 = \{110011\} \cap \{010011\}$, где $m_1 = 1$; $F_1 = \{010011\} \cap \{000011\}$, где $p_1 = 1$;
$F_2 = \{000011\} \cap \{001011\}$, где $m_2 = 1$; $F_3 = \{001011\} \cap \{001111\}$, где $p_2 = 1$;
$F_4 = \{001111\} \cap \{001101\}$, где $m_3 = 1$; $F_5 = \{001101\} \cap \{001100\}$, где $p_3 = 1$;
$F_6 = \{001100\} \cap \{101100\}$, где $m_1 = 1$; $F_7 = \{101100\} \cap \{111100\}$, где $p_1 = 1$;
$F_8 = \{111100\} \cap \{110100\}$, где $m_2 = 1$; $F_9 = \{110100\} \cap \{110000\}$, где $p_2 = 1$;
$F_{10} = \{110000\} \cap \{110010\}$, где $m_3 = 1$; $F_{11} = \{110010\} \cap \{110011\}$, где $p_3 = 1$.

Схема дальнейших рассуждений следует идеям доказательств основных результатов работ [6–8], в которых рассматривались случаи несколько более простых динамических систем, моделирующих функционирование генных сетей только с отрицательными обратными связями.

Для точек всех указанных выше граней их сдвиги вдоль траекторий динамической системы (1) внутри каждого из перечисленных в диаграмме (3) блоков обозначим следующим образом:

$$\varphi_0 : F_0 \to F_1, \quad \varphi_1 : F_1 \to F_2, \quad \varphi_2 : F_2 \to F_3, \quad \varphi_3 : F_3 \to F_4, \quad \varphi_4 : F_4 \to F_5, \quad \varphi_5 : F_5 \to F_6,$$

$$\varphi_6 : F_6 \to F_7, \quad \varphi_7 : F_7 \to F_8, \quad \varphi_8 : F_8 \to F_9, \quad \varphi_9 : F_9 \to F_{10} \quad \varphi_{10} : F_{10} \to F_{11}, \quad \varphi_{11} : F_{11} \to F_0.$$

Пусть $\Phi : F_0 \to F_0$ — композиция всех этих сдвигов. Нетрудно проверить, что $\Phi(E) = E$.

В каждом из блоков разбиения (2) решения системы (1) находятся явно, например в “первом” блоке {010011} диаграммы (3)) имеем

$\dot m_1 = -k_1 m_1$; $\dot p_1 = -l_1 p_1$; $\dot m_2 = -k_2 m_2$; $\dot p_2 = -l_2 p_2$; $\dot m_3 = -k_3(m_3 - a_3)$; $\dot p_3 = -l_3(p_3 - b_3)$.

Рассматривая эти уравнения в замыкании блока {010011}, выберем на грани F_0 внутреннюю точку $X^{(0)}$ с координатами $m_1^{(0)} = 1$, $p_1^{(0)} > 1$, $m_2^{(0)} < 1$, $p_2^{(0)} < 1$, $m_3^{(0)} > 1$, $p_3^{(0)} > 1$ в качестве начальной точки траектории $\mathcal{T}$. В этом блоке решение системы (1) имеет вид

$$m_1(t) = m_1^{(0)} e^{-k_1 t}, \ p_1(t) = p_1^{(0)} e^{-l_1 t}, \ m_2(t) = m_2^{(0)} e^{-k_2 t}, \ p_2(t) = p_2^{(0)} e^{-l_2 t},$$

$$m_3(t) = a_3 + (m_3^{(0)} - a_3)e^{-k_3 t},\ \ p_3(t) = b_3 + (p_3^{(0)} - b_3)e^{-l_3 t}. \tag{4}$$

Согласно Лемме 1, при $t = t_1 > 0$, определяемом из условия $e^{-l_1 t_1} = (p_1^{(0)})^{-1}$, траектория $\mathcal{T}$ выходит из блока $\{010011\}$ через точку $X^{(1)} \in F_1$. Из формул (4) следует, что на другие грани блока $\{010011\}$ эта траектория не попадает, и поэтому точка $X^{(1)}$ имеет координаты

$$m_1^{(1)} = m_1^{(0)} \cdot (p_1^0)^{-k_1/l_1};\ \ p_1^{(1)} = 1;\ \ m_2^{(1)} = m_2^{(0)} \cdot (p_1^0)^{-k_2/l_1};\ \ p_2^{(1)} = p_2^{(0)} \cdot (p_1^0)^{-l_2/l_1};$$

$$m_3^{(1)} = a_3 + (m_3^{(0)} - a_3) \cdot (p_1^0)^{-k_3/l_1};\ \ \ p_3^{(1)} = b_3 + (p_3^{(0)} - b_3) \cdot (p_1^0)^{-k_3/l_1}, \tag{5}$$

При $t = t_2 > t_1$, определяемом условием перехода траектории $\mathcal{T}$ в следующий блок $\{001011\}$ диаграммы (3): $m_2(t_2) = m_2^{(1)} e^{-k_2(t_2 - t_1)} = 1$, точка $X^{(1)}$ попадает на грань F_2, где $m_2 = 1$, в точку, у которой остальные координаты имеют вид:

$$m_1^{(2)} = m_1^{(1)} \cdot \left(\frac{a_2 - 1}{a_2 - p_1^{(0)}}\right)^{k_1/k_2};\quad p_1^{(2)} = p_1^{(1)} \cdot \left(\frac{a_2 - 1}{a_2 - p_1^{(0)}}\right)^{l_1/k_2};\quad p_2^{(2)} = p_2^{(1)} \cdot \left(\frac{a_2 - 1}{a_2 - p_1^{(0)}}\right)^{l_2/k_2};$$

$$m_3^{(2)} = a_3 + (m_3^{(1)} - a_3) \cdot \left(\frac{a_2 - 1}{a_2 - p_1^{(0)}}\right)^{k_3/k_2};\qquad p_3^{(2)} = b_3 + (p_3^{(1)} - b_3) \cdot \left(\frac{a_2 - 1}{a_2 - p_1^{(0)}}\right)^{l_3/k_2}. \tag{6}$$

Из формул перехода (4), (5) и их аналогов для φ_k, $k \geq 2$, видно, что в каждом блоке диаграммы (3) каждая координата точки на любой траектории системы (1) зависит от времени монотонно.

Матрицы Якоби преобразований φ_0 и φ_1, вычисленные в точке E, лежащей на границе всех граней F_k, имеют вид:

$$J(\varphi_0) = \begin{pmatrix} -\dfrac{k_2}{l_1} & 1 & 0 & 0 & 0 \\ -\dfrac{l_2}{l_1} & 0 & 1 & 0 & 0 \\ \dfrac{k_3(a_3-1)}{l_1} & 0 & 0 & 1 & 0 \\ \dfrac{l_3(b_3-1)}{k_1} & 0 & 0 & 0 & 1 \\ -\dfrac{k_1}{l_1} & 0 & 0 & 0 & 0 \end{pmatrix};\qquad J(\varphi_1) = \begin{pmatrix} \dfrac{l_2}{k_2(a_2-1)} & 1 & 0 & 0 & 0 \\ -\dfrac{k_3(a_3-1)}{k_2(a_2-1)} & 0 & 1 & 0 & 0 \\ -\dfrac{l_3(b_3-1)}{k_2(a_2-1)} & 0 & 0 & 1 & 0 \\ \dfrac{k_1}{k_2(a_2-1)} & 0 & 0 & 0 & 1 \\ \dfrac{l_1}{k_2(a_2-1)} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Подобным же образом с помощью соответствующих перестановок индексов элементов первого столбца описываются переходы траекторий системы (1) и их матрицы Якоби для преобразований φ_5 и φ_9, φ_4 и φ_8, соответственно. У переходов φ_2 и φ_3 матрицы Якоби имеют вид:

$$J(\varphi_2) = \begin{pmatrix} -\dfrac{k_3(a_3-1)}{l_2(b_2-1)} & 1 & 0 & 0 & 0 \\ -\dfrac{l_3(b_3-1)}{l_2(b_2-1)} & 0 & 1 & 0 & 0 \\ \dfrac{k_1}{l_2(b_2-1)} & 0 & 0 & 1 & 0 \\ \dfrac{l_1}{l_2(b_2-1)} & 0 & 0 & 0 & 1 \\ -\dfrac{k_2(a_2-1)}{l_2(b_2-1)} & 0 & 0 & 0 & 0 \end{pmatrix};\qquad J(\varphi_3) = \begin{pmatrix} \dfrac{l_3(b_3-1)}{k_3} & 1 & 0 & 0 & 0 \\ -\dfrac{k_1}{k_3} & 0 & 1 & 0 & 0 \\ -\dfrac{l_1}{k_3} & 0 & 0 & 1 & 0 \\ \dfrac{k_2(a_2-1)}{k_3} & 0 & 0 & 0 & 1 \\ \dfrac{l_2(b_2-1)}{k_3} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Так же, как и для матриц $J(\varphi_0)$, и $J(\varphi_1)$, порядок переменных во всех таких матрицах и в соответствующих им гранях определяется следующим образом: сначала записывается переменная, идущая после той, которая на этой грани равна единице, а далее по циклу $1 \to 2 \to 3 \to 4 \to 5 \to 6 \to 1$. Матрицы $J(\varphi_s)$ и $J(\varphi_{s+4})$ отличаются друг от друга только циклической перестановкой индексов элементов первого столбца.

5 Доказательство Теоремы

Вычислим произведения $J(\varphi_1\varphi_0)$ и $J(\varphi_3\varphi_2)$ попарных композиций:

$$J(\varphi_1\varphi_0)=\begin{pmatrix} -\dfrac{l_2a_2}{l_1(a_2-1)} & \dfrac{l_2}{k_2(a_2-1)} & 1 & 0 & 0\\ \dfrac{k_3a_2(a_3-1)}{l_1(a_2-1)} & -\dfrac{k_3(a_3-1)}{k_2(a_2-1)} & 0 & 1 & 0\\ \dfrac{l_3a_2(b_3-1)}{l_1(a_2-1)} & -\dfrac{l_3(b_3-1)}{k_2(a_2-1)} & 0 & 0 & 1\\ -\dfrac{k_1a_2}{l_1(a_2-1)} & \dfrac{k_1}{k_2(a_2-1)} & 0 & 0 & 0\\ -\dfrac{1}{a_2-1} & \dfrac{l_1}{k_2(a_2-1)} & 0 & 0 & 0 \end{pmatrix};\ J(\varphi_3\varphi_2)=\begin{pmatrix} -\dfrac{l_3a_3(b_3-1)}{l_2(b_2-1)} & \dfrac{l_3(b_3-1)}{k_3} & 1 & 0 & 0\\ \dfrac{k_1a_3}{l_2(b_2-1)} & -\dfrac{k_1}{k_3} & 0 & 1 & 0\\ \dfrac{l_1a_3}{l_2(b_2-1)} & -\dfrac{l_1}{k_3} & 0 & 0 & 1\\ -\dfrac{k_2a_3(a_2-1)}{l_2(b_2-1)} & \dfrac{k_2(a_2-1)}{k_3} & 0 & 0 & 0\\ -(a_3-1) & \dfrac{l_2(b_2-1)}{k_3} & 0 & 0 & 0 \end{pmatrix}.$$

Остальные такие попарные произведения $J(\varphi_{4j-3}\varphi_{4j-4})$, а также $J(\varphi_{4j-1}\varphi_{4j-2})$ отличаются от $J(\varphi_1\varphi_0)$, (соответственно от $J(\varphi_3\varphi_2)$) только циклической перестановкой индексов элементов первых двух столбцов.

Рассмотрим вычисленную в точке E матрицу Якоби $J(\Phi)$ отображения Φ:

$$J(\Phi)=J(\varphi_{11}\varphi_{10})\cdot J(\varphi_9\varphi_8)\cdot J(\varphi_7\varphi_6)\cdot J(\varphi_5\varphi_4)\cdot J(\varphi_3\varphi_2)\cdot J(\varphi_1\varphi_0).$$

Непосредственное перемножение всех этих шести матриц приводит к необозримым результатам. Однако, нетрудно вычислить знаки всех элементов матрицы $J(\Phi)$, все эти элементы либо строго положительны, либо строго отрицательны. Кроме того, имеют место неравенства: $J(\Phi)_{11}>1$; $J(\Phi)_{22}>1$.

Зададим в F_0 частичный порядок:
для точек $P=(1,p_1,m_2,p_2,m_3,p_3)\in F_0$ и $\widetilde{P}=(1,\widetilde{p}_1,\widetilde{m}_2,\widetilde{p}_2,\widetilde{m}_3,\widetilde{2}_3)\in F_0$ будем говорить, что P меньше $\widetilde{P}$ (или $P\prec\widetilde{P}$), если выполнены неравенства

$$\widetilde{p}_1\ge p_1;\quad \widetilde{m}_2\le m_2;\quad \widetilde{p}_2\le p_2;\quad \widetilde{m}_3\ge m_3;\quad \widetilde{p}_3\ge p_3. \tag{7}$$

Пусть $\Phi(P):=(1,y_1,x_2,y_2,x_3,y_3)$ и $\Phi(\widetilde{P}):=(1,\widetilde{y}_1,\widetilde{x}_2,\widetilde{y}_2,\widetilde{x}_3,\widetilde{y}_3)$.

Лемма 2. *Отображение Φ монотонно, то есть, если выполнены неравенства (7), то выполняются и неравенства*

$$\widetilde{y}_1\ge y_1;\quad \widetilde{x}_2\le x_2;\quad \widetilde{y}_2\le y_2;\quad \widetilde{x}_3\ge x_3;\quad \widetilde{y}_3\ge y_3. \tag{8}$$

При этом если хотя бы одно из неравенств (7) строгое, то все неравенства (8) строгие.

Доказательство Леммы 2 следует из неравенств вида (7), (8) на гранях $F_1,\ldots,F_{11}$, из формул (4) – (6) и других аналогичных формул переходов $\varphi_k: F_k\to F_{k+1}$.

Лемма 3. *Если точка $P_*=(1,p_1,m_2,1,1,1)\in F_0$ достаточно близка к точке E и отлична от нее, то $\Phi(P_*)\prec P_*$.*

Поскольку $\Phi(E)=E$, при достаточно малых положительных Δp_1 и Δm_2 для $p_1^{(0)}=1+\Delta p_1$, $m_2^{(0)}-\Delta m_2$, $p_2^{(0)}=m_3^{(0)}=p_3^{(0)}=1$ из Леммы 2 и соотношений (8) следует, что $p_1^{(6)}>1+\Delta p_1=p_1^{(0)}$, $m_2^{(6)}<1-\Delta m_2=m_2^{(0)}$, $p_2^{(6)}<1$, $m_3^{(6)}>1$, $p_3^{(6)}>1$, что и составляет утверждение леммы.

Из Лемм 2 и 3 следует, что параллелепипед $\Pi=[p_1^{(0)},b_1]\times[0,m_2^{(0)}]\times[0,1]\times[1,a_3]\times[1,b_3]\subset F_0$ инвариантен относительно монотонного отображения Φ, точнее $\Phi(\Pi)$ содержится во внутренности Π. Пусть $P_1\in\Pi$ — имеющая координаты $(1,b_1,0,0,a_3,b_3)$ точка, противоположная E; она максимальна на грани F_0 относительно порядка $\prec$. Последовательность итераций P_*, $\Phi(P_*)$, $\Phi(\Phi(P_*))=\Phi^{(2)}(P_*)$, ... монотонно удаляется от точки E, а последовательность P_1, $\Phi(P_1)$, $\Phi^{(2)}(P_1)$, ... монотонно приближается к E, причем для всех целых m,n выполняется соотношение $\Phi^{(m)}(P_*)\prec\Phi^{(n)}(P_1)$.

Отсюда следует существование по крайней мере одной точки $P_0\in\Pi$ такой, что $\Phi(P_0)=P_0$, значит траектория этой точки является нетривиальным циклом $\mathcal{C}$ системы (1), и теорема доказана.

Заключение

Из доказательства теоремы следует, что указанный цикл $\mathcal{C}$ проходит по области W из блока в блок в соответствии со стрелками диаграммы (3), кусочно гладок, и его угловые точки расположены на гиперплоскостях, разделяющих соседние блоки области W.

Для трехмерных динамических систем блочного типа, у которых в правых частях уравнений нет функций вида Г, в работах [6, 8] получены аналогичные результаты, а также установлены условия единственности и устойчивости цикла, обнаруженного в объединении блоков валентности один, образующих диаграмму, подобную (3); эти результаты по-видимому можно распространить и на систему (1), и ее многомерные аналоги.

Авторы искренне благодарны В.В. Иванову за полезные обсуждения и советы.

Список литературы

[1] Акиньшин А.А., Голубятников В.П., Голубятников И.В. О некоторых многомерных моделях функционирования генных сетей // СибЖИМ. 2013. Т. 16, iss 1. С. 3–9.

[2] Аюпова Н.Б., Голубятников В.П., Казанцев М.В. О существовании цикла в одной несимметричной модели молекулярного репрессилятора // СибЖВМ. 2017. Т. 20, iss 2. С. 121–130.

[3] Banks H.T., Mahaffy J.M. Stability of cyclic gene models for systems involving repression // Journ. Theor. Biol. 1978. V. 74. P. 323–334

[4] Bukharina T. A., Golubyatnikov V. P., Kazantsev M. V., Kirillova N. E., Furman D. P. Mathematical and numerical models of two asymmetric gene networks // Sib. Electr. Math. Reports. 2018. V. 16, P. 1271–1283.

[5] Глызин С. Д., Колесов А. Ю., Розов Н. Х. Существование и устойчивость релаксационного цикла в математической модели репрессилятора // Матем. заметки. 2017. Т. 101, № 1. С. 58–67.

[6] Голубятников В.П., Иванов В.В., Минушкина Л.С. О существовании цикла в одной несимметричной модели кольцевой генной сети // Сибирский журнал чистой и прикладной математики. 2018. Т. 18, iss 3. С. 26–32.

[7] Голубятников В.П., Иванов В.В. Циклы в нечетномерных моделях кольцевых генных сетей // СибЖИМ. 2018. Т. 21, iss 4. С. 28–38.

[8] Голубятников В.П., Иванов В.В. Единственность и устойчивость цикла в трехмерных блочно-линенйных моделях кольцевых генных сетей // Сибирский журнал чистой и прикладной математики. 2018. Т. 18, iss 4. С. 19–28.

[9] Голубятников В.П., Кириллова Н.Е. О циклах в моделях функционирования кольцевых генных сетей // Сибирский журнал чистой и прикладной математики. 2018. Т. 18, iss 1. С. 54–63.

[10] Elowitz M.B., Leibler S. A synthetic oscillatory network of transcriptional regulators // Nature. 2000. V. 403. P. 335–338.

[11] Колесов А. Ю., Розов Н. Х., Саловничий В. А. Периодические решения типа бегущих волн в кольцевых генных сетях // Известия РАН, серия математическая. 2016. Т. 80, iss 3. С. 67–94.

[12] Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. Задачи функционирования генных сетей // Сибирский журнал индустриальной математики. 2003. Т. 6, iss 2. С. 75–84.

[13] Likhoshvai V. A., Fadeev S. I., Kogai V. V., Khlebodarova T. M. On the Chaos in Gene Networks // Journal of Bioinformatics and Computational Biology. 2013. V. 11, iss 1. P. 1340009-1–1340009-25.

Голубятников Владимир Петрович — д.ф.-м.н., гл. науч.сотр. Института математики им. С. Л. Соболева СО РАН;
e-mail: vladimir.golubyatnikov1@fulbrightmail.org;
Минушкина Лилия Сергеевна — студентка Новосибирского государственного университета;
e-mail: l.minushkina@g.nsu.ru

Дата поступления — 30 апреля 2019 г.

АЛГОРИТМЫ ПОСТРОЕНИЯ КВАЗИСТРУКТУРИРОВАННЫХ СЕТОЧНЫХ МОДЕЛЕЙ ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА НА ОСНОВЕ КОНЦЕПЦИИ БСМ

В. С. Горшунов[1,2], Л. А. Голубева[1], В. П. Ильин[1]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10018

Одним из важных этапов при решении задач математической является этап построение сеточной модели. В настоящей работе описывается подход, предполагающий использование нескольких алгоритмов при генерации одной сеточной модели.

Ключевые слова: сеточные генераторы, сеточная структура данных, САПР, БСМ, Netgen, Cubdat

Введение

При создании программного комплекса (ПК) для решения задач математической физики важно учитывать, что такой программный комплекс должен содержать средства построения расчётной области, генерации расчётных сеток, решения физических уравнения. Например, в ПК на основе концепции Базовой Системы Моделирования (БСМ) [1,2] решаются уравнения Пуассона в плоских или осесимметричных областях практически произвольной формы с граничными условиями Дирихле и Неймана на квазиструктурированных несогласованных сетках. Метод решения таких задач основывается на декомпозиции исходной расчетной области на подобласти, сопрягаемые без наложения. В каждой из подобластей может быть построена сетка своим методом, например, Cubdat-ом или NETGEN-ом [2]. В этом случае должна быть решена проблема сопряжения сеток на границах подобластей.

Вблизи границы исходная прямоугольная сетка подвергается локальной модификации, состоящей в сдвиге приграничных узлов в точки пересечения координатных линий с границей, а в подобластях с сильной неоднородностью решения имеется возможность локального увеличения числа узлов. Тем самым обеспечивается адаптивность сетки.

В статье приводится описание структур данных для работы с квазиструктурированной сеткой пакета прикладных программ и рассматриваются примеры численных расчетов.

1 Сеточная структура данных и сеточные генераторы

Сеточная структура данных (ССД) или же сетки могут представлять из себя набор сеточных вершин, сеточных рёбер, сеточных граней и сеточных элементов. Однако, способ построения и составляющие ССД могут различаться в зависимости от методов решения поставленной задачи. Например, при использовании метода конечных элементов (МКЭ) необходимо, чтобы ССД представляла из себя набор конечных элементов (тетраэдры, криволинейные параллелепипеды и т.д.). Так же существуют так называемые алгебраические сетки — сетки, построенные при помощи алгебраических методов. Обычно, такие способы применяются, когда от сетки требуется лишь наличие вершин, в этом случае ССД задаётся как упорядоченный набор вершин.

ISBN 978-5-901548-42-4

Одними из наиболее известных сеточных генераторов являются Netgen и Gmsh [3]. Так же в работах [1,2,4,5,8] используется сеточный генератор Cubdat. Назначение Cubdat — генерация структурированных сеток в областях, состоящих из параллелепипедов. Данные сеточные генераторы прекрасно справляются со своим назначением, однако их применение налагает на пользователя определённые ограничения — в этих сеточных генераторах отсутствуют построители геометрических моделей, снабжённые удобным пользовательским интерфейсом. Т.е. для построения сетки этим сеточным генераторам требуется на вход геометрическая структура данных в определённом формате, например, STEP или geo, которые представляют собой текстовые файлы, написанные на специальном языке описания геометрии, и которую невозможно задать при помощи внутренних средств этих программ. Каждый их этих генераторов имеет свою уникальную ССД и свой формат выходных данных. Несмотря на это, их можно обобщить в единую форму, например, на базе ИИПП «Гербарий» [6].

2 Модуль генерации сеток

В программном комплексе для решения задач математической физики на основе концепции БСМ [9] реализован модуль построения и отображения ССД. Программный комплекс представляет из себя набор модулей, реализованных на базе ИИПП «Гербарий». Цель такого программного комплекса — обеспечение целостности вычислительного процесса, возможности оперативно вносить коррективы на каждом из этапов вычислительного процесса, предоставление удобно пользовательского интерфейса. В настоящий момент сконструировано три модуля, соответствующих поставленным целям: модуль «Вороной», отвечающий за задание геометрической структуры данных (ГСД) и функциональной структуры данных (ФСД), модуль «Делоне», отвечающий за построение ССД и модуль «Чебышёв», отвечающий за решение уравнений. Рассмотрим подробно модуль «Делоне», в соответствии с заявленным во введении.

Как уже было описано выше, назначение данного модуля — генерация и отображения сеток. Предполагается, что исходная расчётная область была декомпозирована на подобласти [7]. Особенность этого модуля в том, что он предоставляет возможность применять для построения сетки в разных подобластях разные алгоритмы и сеточные генераторы, как сторонние, так и собственной разработки. Т.е. модуль позволяет строить квазиструктурированную сетку — сетку, структурированную в каждой из подобластей, в целом такая сетка может быть неструктурированной. Структурированной сеткой будем назвать такую сетку, в которой можно задать однозначную нумерацию на множестве узлов, иначе говоря, упорядочить это множество. Далее рассмотрим принцип работы, возможности и составляющие модуля «Делоне».

Построение конечно-элементной сетки разбито на несколько этапов:

1. Задание координат вершин.
2. Указание типа конечных элементов (КЭ) для каждой из подобластей.
3. Составление из полученных вершин КЭ заданной формы.
4. Визуализация построенной сетки.

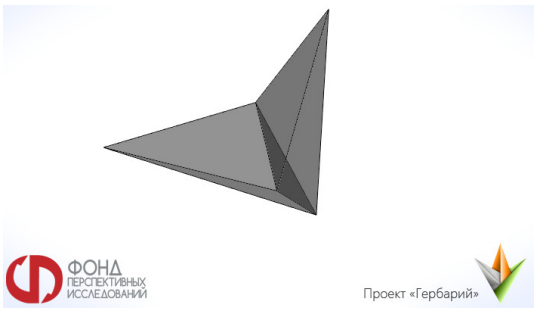


Рис. 1: Два тетраэдра

На рис. 1 представлен простой пример визуализированной КЭ сетки, состоящей из 5 вершин и 2 тетраэдров.

Следующий пример (рис. 2) взят из работы [7]. Данная сетка сегнерирована в Netgen. А в модуле «Делоне» реализована возможность считывания этой сетки во внутреннюю ССД Гербария из внешнего источника. Сетка доступна для дальнейшей работы, например, решения уравнений на ней. Исходные сеточные данные Netgen находятся в файле формата *.vol.

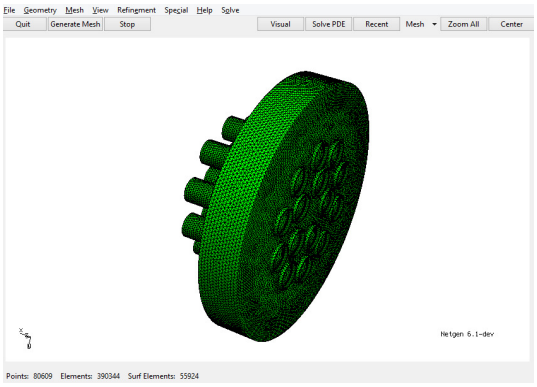


Рис. 2: Сетка «Netgen» для катодов в «Netgen»



Рис. 3: Сетка «Netgen» для катодов в «Гербарий»

Так же реализована возможность генерации сетки с помощью алгоритмов Cubdat посредством интегрирования кода Cubdat в модуль «Делоне».

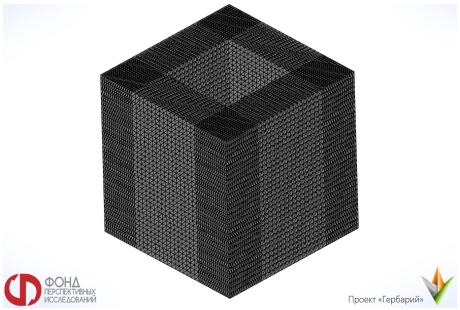


Рис. 4: Сетка «Cubdat» для куба с несквозным отверстием

Пример квазиструктурированной сетки из 2 подобластей — параллелепипеда со сквозным отверстием и лунки на месте сквозного отверстия изображён на рис. 5. Сетка для параллелепипеда построена при помощи алгоритмов Cubdat, сетка для лунки построена с использованием собственных алгоритмов.

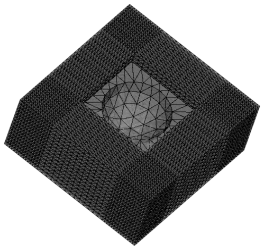


Рис. 5: Сетка для лунки

Заключение

Использование нескольких алгоритмов при построении квазиструктурированной сетки предполагает разбиение исходной расчётной области на подобласти, что значит для эффективного применения такого подхода

необходимо иметь построитель геометрии и сеточный генератор объединённые на базе одной платформы. Такой способ предложен в работе [9] и реализован как часть программного комплекса. В настоящей статье кратко описана структура этого ПК и подробно разобрана его часть, ответственная за генерацию сеточных моделей. Показано, что отличия в способах задания ССД в различных алгоритмах и сеточных генераторах не препятствует сведению ССД каждого из них в другую структуру.

Список литературы

[1] Ильин В.П. Фундаментальные вопросы математического моделирования.//Вестник Российской Академии Наук, т. 86, № 4, 2016, 26–36.

[2] Беляев Д.О., Козырев А.Н., Свешников В.М. Моделирование интенсивных пучков заряженных частиц на квазиструктурированных сетках при помощи пакета прикладных программ ЭРА — 2В. //Актуальные проблемы вычислительной математики и математического моделирования (г. Новосибирск, 12–15 июня 2012 г.): Труды всероссийской конференции. — 2012. — С. 13.

[3] https://ngsolve.org/docu/latest/

[4] Голубева Л.А., Ильин В.П., Козырев А.Н. О программных технологиях в геометрических аспектах математического моделирования.–Вестник НГУ, серия Информационные технологии, т. 10, N 2, 2012, 25–33.

[5] Ильин В.П., Гладких В.С. Базовая система моделирования (БСМ):концепция, архитектура и методология "Современные проблемы математического моделирования , обработки изображений и параллельных вычислений 2017"(СПММОИиПВ-2017),Ростов-на-Дону издательство ДГТУ, 2017, 151–158 .

[6] http://tflex.ru/about/publications/detail/index.php?ID=3846

[7] Бутюгин Д.С., Гурьева Я.Л., Ильин В.П., Перевозкин Д.В., Петухов А.В., Скопин И.Н. Функциональность и технологии алгебраических решателей в библиотеке Krylov.– Вестник ЮУрГУ. Серия “Вычислительная математика и информатика”, т. 2, N 3, 2013, 92–105.

[8] Gurieva Ya.L., Il’in V.P. Program package for 3D boundary-value elliptic Problem.–Bull. NCC, series: “Num. Anal.,” Iss. 11, 2002, 35–52.

[9] Голубева Л.А., Горшунов В.С., Ильин В.П., Эрдыниев Э.Б. Программный комплекс для решения 3-мерных задач математической физики на основе концепции БСМ. //Труды Международной конференции "Вычислительная математика и математическая геофизика"посвященная 90-летию со дня рождения академика А. С. Алексеева. 2018. С. 126–132.

Голубева Людмила Андреевна — к.ф.-м.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: golubeva@labchem.sscc.ru;
Горшунов Василий Сергеевич — инженер Института вычислительной математики и математической геофизики СО РАН;
аспирант Новосибирского государственного университета
e-mail: basil.gorshunov@gmail.com;
Ильин Валерий Павлович — д.ф.-м.н., глав. науч.сотр. Института вычислительной математики и математической геофизики СО РАН; Новосибирский государственный университет;
e-mail: ilin@sscc.ru.
Дата поступления — 1 июня 2019 г.

ПОИСК ЦИКЛОВ В ОДНОЙ n-МЕРНОЙ МОДЕЛИ КОЛЬЦЕВОЙ ГЕННОЙ СЕТИ

В. С. Градов

Новосибирский государственный университет, 630090, Новосибирск

УДК 514.745.82
DOI: 10.24411/9999-016A-2019-10019

Установлены необходимые и достаточные условия существования периодических траектории у динамической системы, моделирующей функционирование кольцевой генной сети, регулируемой отрицательной обратной связью в случае $n > 5$. Фазовый портрет системы разбивается на блоки, в каждом из которых система линейна. Дискретизация фазового портрета позволяет локализовать положение периодических траекторий.

Ключевые слова: модели генных сетей, блочно-линейные динамические системы, дискретизация фазовых портретов, валентность блока, циклы.

Введение

Генные сети контролируют биохимические процессы, происходящие в живых организмах, см. [1, 3, 4, 8]. В данной работе изучаются генные сети, в которых каждое вещество взаимодействует с предыдущим посредством отрицательной обратной связи, то есть когда скорость концентрации одного вещества монотонно убывает с увеличением концентрации предыдущего вещества в сети. Главной задачей является выявление периодических режимов функционирования таких сетей, поскольку такие режимы играют важную роль в молекулярной биологии.

1 Модель

В качестве математических моделей изучаемых генных сетей рассматриваются n - мерные блочно-линейные динамические системы симметричные относительно циклической перестановки координат:

$$\dot{X}_1 = L(X_n) - X_1; \quad \dot{X}_2 = L(X_1) - X_2; \quad \dots \quad \dot{X}_n = L(X_{n-1}) - X_n, \tag{1}$$

где $X_j \geq 0$ являются концентрациями веществ генной сети, вычитаемые описывают процессы их разложения, а ступенчатые функции $L(w)$ описывают синтез компонент сети и определяются следующим образом:

$$L(w) = A > 1 \quad \text{для} \quad w \in (0, 1); \quad L(w) = 0 \quad \text{для} \quad 1 < w;$$

В работах [2, 4] рассматривались аналогичные системы других размерностей.

2 Дискретизация фазового портрета

Рассмотрим куб $Q = [0, A]^n \subset \mathbb{R}^n_+$ и точку $E = (1, \dots, 1)$ в его внутренности. В работах [1,5] было установлено, что область Q положительно инвариантна для системы (1), то есть с течением времени все траектории этой системы попадают в куб Q и дальше не выходят из него.

Доказательства состоит в проверке направления векторного поля на гранях Q, т.е. нужно проверить неравенства: $\dot{X}_j \mid_{X_j=0} \geqslant 0$, $\dot{X}_j \mid_{X_j=A} \leqslant 0$ см., например, [2,6].

Работа выполнена при частичной финансовой поддержке РФФИ (проект 18-01-00057)

ISBN 978-5-901548-42-4

Разобьем Q плоскостями $x_j = 1$ на 2^n параллелепипедов. Будем называть их блоками, и пронумеруем бинарными мультииндексами $\epsilon = \{\varepsilon_1 \dots \varepsilon_n\}$, где $\varepsilon_j = 0$, если $X_j < 1$, и $\varepsilon_j = 1$, если $X_j > 1$ — как у вершин n-мерного булева куба. Здесь $j = \overline{1,n}$.

В работах [1, 2] было установлено, что для каждой пары соседних блоков ϵ_1 и ϵ_2, пересекающихся по $(n-1)$-мерной грани, траектории всех точек внутренности этой грани переходят либо из ϵ_1 в ϵ_2, либо из ϵ_2 и ϵ_1. Это позволяет построить граф на булевом кубе и ориентировать его дуги, согласно переходам из блока в блок.

Будем называть валентностью блока ϵ количество его граней, через которые траектории системы выходят из этого блока, то есть валентность блока — это количество дуг, исходящих из соответствующей вершины ориентированного графа G.

В случае четного n среди 2^n блоков разбиения области Q есть блоки валентностей 0, 2, ..., n, а в случае нечетного n есть блоки валентности 1, 3, ..., n. Блоки с мультииндексами $\{0 \dots 0\}$ и $\{1 \dots 1\}$ имеют валентность $V = n$, то есть траектории системы (1) могут только выходить из этих блоков (см. [1, 6]).

В работе [5] для систем вида (1) были установлены необходимые и достаточные условия существования единственного устойчивого цикла $\mathcal{C}_1$, проходящего через блоки валентности 1. Для таких систем рассмотрим теперь блоки валентности $V = n - 2$:

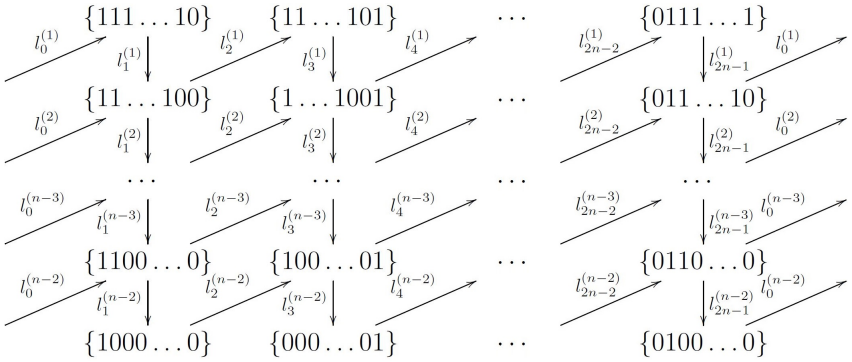


Рис. 1: Диаграмма переходов блоков валентности $V = n - 2$

Индекс каждого блока s-ой строки, $s = 1, 2, \dots, n-1$, содержит s нулей и $n - s$ единиц. В каждой строке мультииндекс следующего блока отличается от предыдущего смещением всех элементов мультииндекса на один шаг влево.

Для фиксированного $k = 1, \dots, n-2$ последовательность блоков со стрелками $l_j^{(k)}$ назовем Диаграммой уровня k, где $j = 0, \dots, 2n-1$.

2.1 Вспомогательные алгебраические утверждения

Пусть $A > 1$. Рассмотрим последовательности многочленов:

$$\begin{aligned} &P_1(x) = x + \left(1 - \frac{A^2}{A-1}\right), \quad P_2(x) = x^2 + \left(1 - \frac{A^2}{A-1}\right)x + 1, \\ &P_n(x) = x^n + \left(1 - \frac{A^2}{A-1}\right)x^{n-1} + x^{n-2} + \ldots + x + 1, \quad n > 1. \end{aligned} \tag{2}$$

Так же рассмотрим ещё одну последовательность многочленов:

$$\begin{aligned} &Q_1(x) = x - (A-1), \quad Q_2(x) = x^2 - (A-1)(x+1), \\ &Q_n(x) = x^n - (A-1)(x^{n-1} + x^{n-2} + \ldots + x + 1), \quad n > 1. \end{aligned} \tag{3}$$

При $n \geq 2$ верны рекуррентные соотношения (из которых вытекают следующие леммы):

$$P_n(x) = xP_{n-1}(x) + 1, \quad Q_n(x) = xQ_{n-1}(x) - (A-1). \tag{4}$$

Лемма 1. *1) Если $n \geqslant 2$ нечетное, то $P_n(x) = P_1(x)$, при $x = \frac{A}{A-1}$, A. Если n четное, то при $x = -1$, $\frac{A}{A-1}$, A.*

2) Если $n-k$, где $2 \leqslant k < n$, нечетное, то $P_n(x) = P_k(x)$ при $x = 0,\ \frac{A}{A-1},\ A$. Если $n-k$ четное, то при $x = -1,\ 0,\ \frac{A}{A-1},\ A$.

Лемма 2. *1) Многочлен $P_2(x)$ имеет единственный корень на интервале $(0,1)$ тогда и только тогда, когда $A > 1$.*
2) Многочлен $P_3(x)$ имеет единственный корень на интервале $(0,1)$ тогда и только тогда, когда $A > 1$, $A \neq 2$;
3) Многочлен $P_n(x)$ при $n \geqslant 4$ имеет единственный корень на интервале $(0,1)$ тогда и только тогда, когда $A^2 > (n+1)(A-1)$.

Из предыдущих лемм следует, что для $n \geq 2$ выполняются неравенства

$$P_n(0) = 1 > 0, \quad P_n(A) = -\frac{1}{A-1} < 0, \quad P_n\left(\frac{A}{A-1}\right) = -(A-1) < 0.$$

Следовательно $P_n(x)$ имеет два положительных корня, причем один из корней больше единицы, так как $A > 1$ и $\frac{A}{A-1} > 1$ при $a > 1$. Если $P_n(1) = n + 1 - \frac{A^2}{A-1} < 0$, то на интервале $(0,1)$ этот многочлен имеет единственный корень. Если $P_n(1) > 0$, то оба его положительных корня лежат в интервалах $(1, A)$ и (A, x_2), где $x_2 > A$ — корень многочлена $P_2(x)$. Так как $P_n(1) > 0$, $P_n(A) < 0$, $P_n\left(\frac{A}{A-1}\right) < 0$ и $P_n(x_2) > 0$, то оба интервала $(1, A)$ и (A, x_2) содержат корни многочлена $P_n(x)$.

Лемма 3. *Пусть $n \geq 1$ и $x^{(n)}, x^{(n+1)} \in (0,1)$ — корни многочленов $P_n(x)$ и $P_{n+1}(x)$ соответственно, тогда $x^{(n+1)} > x^{(n)}$.*

Лемма 4. *Пусть $n \geqslant 2$. Если $y^{(n)}, y^{(n+1)}$ — корни $Q_n(x)$ и $Q_{n+1}(x)$ соответственно, то $y^{(n+1)} > y^{(n)}$.*

Лемма 5. *Пусть $n > 4$, $z \in (0,1)$ — корень $P_n(x)$. Тогда 1) $Q_{n-1}(z) < 0$ для любого $A > 1$,*
2) $Q_{n-2}(z) > 0$, если $2 < 2A < n + 1 - \sqrt{(n+1)^2 - 4(n+1)}$,
3) $Q_{n-2}(z) < 0$, если $2A > n + 1 + \sqrt{(n+1)^2 - 4(n+1)}$.

3 Описание циклов системы (1)

Рассмотрим Диаграмму на рисунке 1 и введем обозначения:

$$F_0^{[k]} = B_0^{[k]} \cap B_1^{[k]},\ F_1^{[k]} = B_1^{[k]} \cap B_2^{[k]},\ F_2^{[k]} = B_2^{[k]} \cap B_3^{[k]},$$
$$B_0^{[k]} = \{\underbrace{1\ldots1}_{n-k}\underbrace{0\ldots0}_{k}\},\ B_1^{[k]} = \{\underbrace{1\ldots1}_{n-k-1}\underbrace{0\ldots0}_{k+1}\},$$
$$B_2^{[k]} = \{\underbrace{1\ldots1}_{n-k-1}\underbrace{0\ldots0}_{k}1\},\ B_3^{[k]} = \{\underbrace{1\ldots1}_{n-k-2}\underbrace{0\ldots0}_{k+1}1\}.$$

Пусть $X = (X_1, X_2, \ldots, X_n)$. Тогда если $X \in F_0^{[k]}$, то:

$$X_1 > 1, \ldots, X_{n-k-1} > 1, X_{n-k} = 1, X_{n-k+1} < 1, \ldots, X_n < 1. \tag{5}$$

Если $X \in F_1^{[k]}$, то:

$$X_1 > 1, \ldots, X_{n-k-1} > 1, X_{n-k} < 1, \ldots, X_{n-1} < 1, X_n = 1. \tag{6}$$

Если $X \in F_2^{[k]}$, то:

$$X_1 > 1, \ldots, X_{n-k-2} > 1, X_{n-k-1} = 1, X_{n-k} < 1, \ldots, X_{n-1} < 1, X_n > 1. \tag{7}$$

Теорема 1. *Пусть $n \geqslant 5$. Тогда*
1) Система (1) имеет единственный цикл $\mathcal{C}_{n-2}^{[1]}$, симметричный относительно циклической перестановки координат σ_n: $X_1 \to X_2 \to \ldots \to X_n \to X_1$ в блоках Диаграммы уровня $k = 1$ тогда и только тогда, когда $A > \frac{n+\sqrt{n^2-4n}}{2}$.
2) Система (1) имеет единственный цикл $\mathcal{C}_{n-2}^{[n-2]}$, симметричный относительно σ_n в блоках Диаграммы уровня $k = n-2$ тогда и только тогда, когда $1 < A < \frac{n-\sqrt{n^2-4n}}{2}$

3) Для любого $A > 1$ система (1) не имеет циклов, симметричных относительно σ_n в блоках Диаграмм уровней $k = 2, \ldots, n-3$.

Доказательство. Рассмотрим уровень k. В дальнейшем будем использовать следующие индексы: $i = 1, \ldots, n-k-2$, $j = 1, \ldots, n-k-1$, $m = n-k+1, \ldots, n$, $h = 2, \ldots, n-k-2$. Пусть траектория системы берет начало в блоке $B_1^{(k)}$ на грани $F_0^{(k)}$ в точке $X^{(0)}$.

Решение системы в этом блоке выглядит следующим образом:

$$X_1 = A - e^{-t}(A - X_1^{(0)}), \quad X_{j+1} = X_{j+1}^{(0)} e^{-t}, \quad X_m = A - e^{-t}(A - X_m^{(0)}).$$

Найдем время t_1, за которое точка $X^{(0)} \in F_0^{(k)}$ переходит в $X^{(1)} \in F_1^{(k)}$ и выпишем координаты точки $X^{(1)}$:

$$\begin{gathered} X_1^{(1)} = A - \frac{A-1}{A - X_n^{(0)}}(A - X_1^{(0)}),\ X_{j+1}^{(1)} = X_{j+1}^{(0)} \frac{A-1}{A - X_n^{(0)}}, \\ X_m^{(1)} = A - \frac{A-1}{A - X_n^{(0)}}(A - X_m^{(0)}),\ e^{-t_1} = \frac{A-1}{A - X_n^{(0)}}. \end{gathered} \tag{8}$$

Аналогично найдем время t_2, за которое точка $X^{(1)} \in F_1^{(k)}$ переходит в точку $X^{(2)} \in F_2^{(k)}$ и выпишем координаты точки $X^{(2)}$:

$$\begin{gathered} X_1^{(2)} = \frac{A - X_n^{(0)}}{(A-1)X_{n-k-1}^{(0)}} \left(A - \frac{(A-1)\left(A - X_1^{(0)}\right)}{A - X_n^{(0)}} \right),\ X_h^{(2)} = \frac{X_h^{(0)}}{X_{n-k-1}^{(0)}}, \\ X_{n-k}^{(2)} = \frac{1}{X_{n-k-1}^{(0)}},\ X_m^{(2)} = A - \frac{A - X_m^{(0)}}{X_{n-k-1}^{(0)}},\ e^{-t_2} = \frac{A - X_n^{(0)}}{(A-1)X_{n-k-1}^{(0)}}. \end{gathered} \tag{9}$$

Для существования цикла, симметричного относительно циклической перестановки координат, нужно потребовать, чтобы $X_i^{(2)} = X_{i+1}^{(0)}$, $i = \overline{1, n-1}$, $X_n^{(2)} = X_1^{(0)}$. Тогда соотношения выше запишутся в виде:

$$\begin{gathered} X_2^{(0)} = \frac{A - X_n^{(0)}}{(A-1)X_{n-k-1}^{(0)}} \left(A - \frac{(A-1)\left(A - X_1^{(0)}\right)}{A - X_n^{(0)}} \right),\ X_{h+1}^{(0)} = \frac{X_h^{(0)}}{X_{n-k-1}^{(0)}}, \\ X_{n-k+1}^{(0)} = \frac{1}{X_{n-k-1}^{(0)}},\ X_{m+1}^{(0)} = A - \frac{A - X_m^{(0)}}{X_{n-k-1}^{(0)}}. \end{gathered}$$

Пусть $X_{n-k+1}^{(0)} = y$. Тогда можем переписать полученные соотношения в терминах y:

$$X_2^{(0)} = \frac{(A-y)y^k}{A-1}(A - (A-1)y), \tag{10}$$

$$X_{h+1}^{(0)} = X_2^{(0)} y^{h-1},\ X_{n-k+1}^{(0)} = y = \frac{1}{X_{n-k-1}^{(0)}}, \tag{11}$$

$$X_{m+1}^{(0)} = A - (A-y)y^{m-n+k},\ X_{n+1}^{(0)} = X_1^{(0)}. \tag{12}$$

Проверим выполнение условий (5). Положим $h = n-k-2$ и из (11) получим, что $X_2^{(0)} = y^{2+k-n}$. Подставим в (10) и получим следующее уравнение относительно y:

$$y^n - \frac{A^2}{A-1} y^{n-1} + \frac{A^2}{A-1} y^{n-2} - 1 = 0.$$

Разделив обе части на $(y-1)$, получим уравнение $P_{n-1}(y) = 0$, которое по лемме 2 имеет единственный корень $z \in (0,1)$ тогда и только тогда, когда $A > 1$ и $A^2 > n(A-1)$. Эти неравенства выполняются при:

$$A \in \left(1, \frac{n - \sqrt{n^2 - 4n}}{2}\right) \cup \left(\frac{n + \sqrt{n^2 - 4n}}{2}, \infty\right). \tag{13}$$

Далее полагаем, что выполняется условие (13). Очевидно, что $X_2^{(0)} = z^{2+k-n} > 1$, откуда следует, что $X_{h+1}^{(0)} > 1$. Проверка условий на координаты $X_{m+1}^{(0)}$ заключается в проверке неравенств:

$$A - (A - z)z^l < 1,\ A - (A - z)z^k > 1,\ l = \overline{1, k-1},$$

которые сводятся к неравенствам $Q_l(z) > 0$, $Q_k(z) < 0$.

1) Рассмотрим случай $k = 1$. Пользуясь леммой 2 нетрудно показать, что при $A > \frac{n+\sqrt{n^2-4n}}{2}$ выполнено неравенство $Q_1(z) < 0$. Так же пользуясь (8) можно показать, что выполнене неравенств (6) сводится к проверке неравенства $Q_1(z) < 0$. Следовательно, при $A > \frac{n+\sqrt{n^2-4n}}{2}$ существует единственный цикл $\mathcal{C}_{n-2}^{[1]}$, симметричный относительно σ_n в диаграмме уровня $k = 1$. Таким образом доказана достаточность.

Пусть у системы (1) существует единственный цикл, симметричный относительно σ_n, и проходящий по блокам уровня $k = 1$. Тогда выполнены условия (5) – (7), из которых следует, что $A > \frac{n+\sqrt{n^2-4n}}{2}$. Так же выполнено условие $X_i^{(2)} = X_{i+1}^{(0)}$, $i = \overline{1, n-1}$, $X_n^{(2)} = X_1^{(0)}$, что влечет собой (10)–(12). А значит $P_{n-1}(X)$ имеет единственный корень на интервале $(0, 1)$. Согласно лемме 3 можно заключить, что все условия выполняются одновременно при $A > \frac{n+\sqrt{n^2-4n}}{2}$. Таким образом доказана необходимость.

2) Рассмотрим случай $k = n - 2$. Из лемм 4, 5 следует, что при $1 < A < \frac{n-\sqrt{n^2-4n}}{2}$ выполняются неравенства $Q_l(z) > 0$, $Q_{n-2}(z) < 0$, $l = \overline{1, n-3}$. Так же пользуясь (8) несложно увидеть, что для любого $A > 1$ выполняются неравенства (6). Следовательно, при $1 < A < \frac{n-\sqrt{n^2-4n}}{2}$ существует единственный цикл $\mathcal{C}_{n-2}^{[n-2]}$, симметричный относительно σ_n в диаграмме уровня $k = n-2$. Таким образом доказана достаточность. Необходимость доказывается аналогично предыдущему пункту.

3) Рассмотрим случай $k = 2, \ldots, n-3$. Если $A \in \left(1, \frac{n-\sqrt{n^2-4n}}{2}\right)$, то по леммам 4, 5 для любого $l = 1, \ldots, n-3$ выполнено $Q_l(z) > 0$. Значит не возможно выполнение неравенства $Q_k(z) < 0$.

Если $A > \frac{n+\sqrt{n^2-4n}}{2} > 2$, то не будет выполнено неравенство $Q_1(z) = z - (A - 1) > 0$, так как $z \in (0, 1)$.

Следовательно у системы (1) не существует циклов, симметричных относительно циклической перестановки координат в диаграммах уровней $k = 2, \ldots, n-3$. □

Заключение

В работе [5] было показано, что для существования единственного устойчивого цикла $\mathcal{C}_1$ системы (1) в блоках валентности 1 необходимо и достаточно, чтобы $A > 1$. Результаты данной работы показывают, что перечисленные циклы $\mathcal{C}_{n-2}^{[1]}$, $\mathcal{C}_{n-2}^{[n-2]}$ содержатся в инвариантных кусочно-линейных двумерных поверхностях, симметричных относительно σ_n и имеющих одну общую вершину — точку E. В случае нечетного n при $A > 1$ динамическая система (1) имеет по крайней мере два цикла, которые проходят по блокам валентностей 1 и $n - 2$, соответственно.

Авторы выражают искреннюю признательность Н.Б. Аюповой за полезные обсуждения.

Список литературы

[1] Акиньшин А. А., Голубятников В. П. Циклы в симметричных динамических системах // Сибирский журнал чистой и прикладной математики. 2012. Т. 12, №2. С. 3–12.

[2] Аюпова Н. Б., Голубятников В. П. О единственности цикла в несимметричной трехмерной модели молекулярного репрессилятора // Сибирский журнал индустриальной математики. 2014. Т. 17, № 1. С. 3–7.

[3] Elowitz M. B., Leibler S. A synthetic oscillatory network of transcriptional regulators // Nature. 2000. V. 403. P. 335–338.

[4] Glass L., Pasternack J. S. Stable oscillations in mathematical models of biological control systems // Journal of Mathematical Biology. 1978. V.6. P. 207–223.

[5] Голубятников В. П., Иванов В. В. Циклы в нечетномерных моделях кольцевых генных сетей // Сибирский журнал индустриальной математики. 2018. Т. 21, N 4. С. 28–38.

[6] Казанцев М. В. О некоторых свойствах графов доменов динамических систем // Сибирский журнал индустриальной математики. 2015. Т. 18, № 4. С. 42–49.

[7] Колесов А.Ю., Розов Н.Х., Садовничий В.А. Периодические решения типа бегущих волн в кольцевых генных сетях // Известия РАН, серия математическая. 2016. Т. 80, N. 3. С. 67–94.

[8] Likhoshvai V. A., Fadeev S. I., Kogai V. V., Khlebodarova T. M. On the chaos in gene networks // Journal of Bioinformatics and Computational Biology. 2013. V 11, N 1. P. 1340009-1 — 1340009-25.

Градов Вячеслав Сергеевич — магистрант ММФ Новосибирского государственного университета;
e-mail: gnets2008@outlook.com.
Дата поступления — 30 апреля 2019 г.

ФРАКТАЛЬНЫЕ ХАРАКТЕРИСТИКИ ГЕНОМОВ КАК ИНСТРУМЕНТ ВНУТРИ- И МЕЖВИДОВОЙ КЛАССИФИКАЦИИ ОРГАНИЗМОВ

В. Д. Гусев[1], Л. А. Мирошниченко[1], Ю.П. Джиоев[2]

[1] *Институт математики им. С.Л.Соболева СО РАН, 630090, Новосибирск*

[2] *НИИ Биомедицинских технологий Иркутского государственного медицинского университета, 664003, Иркутск*

УДК 575.22
DOI: 10.24411/9999-016A-2019-10020

Тандемно повторяющиеся цепочки символов (периодичности) широко представлены в геномах различных организмов. Их часто используют в качестве молекулярных маркеров в задачах диагностики, генотипирования и др. Для достаточно длинных геномов спектры периодичностей становятся слишком объёмными. Предложено использовать вместо них гораздо более компактные и легко трактуемые фрактальные спектры (авторский термин), компонентами которых являются тандемно повторяющиеся цепочки палиндромов и комплементарных палиндромов, обладающие свойствами самоподобия. Показано, что фрактальные спектры позволяют решать задачи дифференциации родственных объектов, в частности, бактерий возбудителей псевдотуберкулёза и чумы, близких по структуре генома, но принципиально отличающихся по степени патогенности для человека.

Ключевые слова: символьные последовательности (тексты), бактериальные геномы, периодичности, локальные фракталы, фрактальные спектры.

Введение

Тандемно повторяющиеся фрагменты ДНК (периодичности) составляют значимую долю геномов различных организмов. О механизмах их возникновения и участии в различных геномных перестройках можно прочитать в [1]. Повторяющуюся единицу обычно называют мономером (M) или периодом. В зависимости от длины мономера $|M|$ условно различают три класса тандемных структур: сателлиты ($|M| > 100$), минисателлиты ($7 \leq |M| \leq 100$) и микросателлиты ($1 \leq |M| \leq 6$), характеризующиеся различной эволюционной историей. Тандемные повторы играют важную роль в регуляции основных генетических процессов – транскрипции и репликации. Но особый интерес представляет возможность их использования в качестве молекулярных маркеров в популяционно-генетических исследованиях: ранняя диагностика различных патологий, генотипирование, ДНК-дактилоскопия, определение предрасположенности к определённым видам спорта и т.п.

Аннотации геномов различных организмов содержат лишь ограниченную информацию о наиболее значимых тандемных повторах, поскольку полные спектры периодичностей для геномов с длинами порядка 10^6 н.п. и выше уже достаточно объёмны. Характеризация таких спектров и выявление наиболее информативных их компонентов требует отдельного рассмотрения. Значительный интерес в этом плане представляют периодичности, в которых повторяющаяся единица (мономер) является палиндромом или комплементарным палиндромом. Они обладают признаками фрактальности (самоподобия) в том плане, что повторение палиндрома (обычного или комплементарного) приводит к появлению нового палиндрома того же типа вдвое большей длины. Возникающие по указанной схеме структуры (локальные фракталы) [2] часто являются функционально значимыми.

Фрактальные спектры геномов гораздо компактнее, чем спектры периодичностей, и обладают значительным классификационным потенциалом [3]. Целью данной работы является иллюстрация возможностей

Работа выполнена при поддержке программы Фундаментальных Научных исследований РАН, проект № 0314-2019-0015.

ISBN 978-5-901548-42-4

использования фрактальных спектров для дифференциации близких геномов, характеризующих родственные организмы. Таковыми в рамках данной работы выбраны два вида бактерий: Yersinia pseudotuberculosis и Yersinia pestis. Представители первого вида являются возбудителями псевдотуберкулёза, относимого к категории не слишком опасных инфекционных заболеваний, тогда как бактерии второго вида являются возбудителями чумы – смертельно опасного заболевания, унёсшего жизни миллионов людей [4]. Несмотря на принципиальные отличия по уровню патогенности (для человека), эти бактерии относятся к родственным видам (бактерию Y. Pseudotuberculosis принято считать прародителем Y.pestis). В [5] показано, что их хромосомы схожи по своему строению более чем на 90%, а значимые отличия наблюдаются в основном по составу и строению сопутствующих плазмид. Тем не менее, поиск значимых (трактуемых) отличий продолжается и на геномном уровне, поскольку «выявление даже небольших структурных и функциональных различий может приблизить нас к пониманию причин высокой патогенности возбудителя чумы» [6].

1 Детализация задачи

В качестве исходных данных используются две подборки полных геномов, характеризующих виды Yersinia pseudotuberculosis и Yersinia pestis, соответственно. Они формируются из числа последовательностей, представленных в базе GenBank (по 14 геномов в каждой подборке). Типичные длины геномов составляют порядка 4.5 миллионов символов.

Основная роль в задаче дифференциации геномов двух видов отводится выбору информативных признаков (маркеров). В данном случае это всевозможные цепочки нуклеотидов с длиной равной или превышающей пороговое значение r, имеющие структуру локального фрактала и присутствующие только в одной из исходных («обучающих») подборок. Наибольший вес придаётся маркерам, представленным во всех геномах своего класса. Очевидно, что с увеличением числа геномов в обучающих подборках невозможно гарантировать наличие такого рода «идеальных» маркеров. Реально приходится ограничиваться цепочками, присутствующими в большинстве геномов одной подборки и отсутствующими (за редкими исключениями) во второй. Окончательное решение в этом случае принимается по группе из нескольких маркеров, что эквивалентно использованию коллективного решающего правила.

Пороговое значение r выбирается экспериментально по итогам вычисления фрактальных спектров «рандомизированных геномов» (с изменённым случайным образом порядком следования нуклеотидов).

2 Фрактальные спектры бактериальных геномов

Пусть $\Sigma = \{a, c, g, t\}$ — исходный алфавит, S — последовательность, составленная из элементов Σ, $N = |S|$ — длина последовательности S. Периодичностью будем называть последовательность, полученную тандемным k-кратным повторением базового фрагмента (мономера) M ($k \geq 2$). Для представления периодичности удобно использовать краткую запись $S = M^k$. Например, $S = tcaagtcaagtcaag = (tcaag)^3$ или $S = M^3$, где $M = tcaag$. Иногда периодичность заканчивается неполным периодом, например $S = agtagtagtagtag$, где $M = agt$. Тогда формально можно использовать запись $S = (M)^k$, где $k = |S|/|M|$ не является целым числом (в нашем примере $k = 14/3$).

В бактериальных геномах с длинами порядка 10^6 нуклеотидных пар (н.п.) содержится множество периодичностей с разной длиной периода ($|M| \geq 1$), кратностью повторения ($k \geq 2$) и длиной L ($L \geq 2|M|$). Многие из них носят случайный характер и должны быть отфильтрованы. Простейший способ фильтрации – введение ограничения на длину рассматриваемых периодичностей $L \geq r$. Так, при пороговом значении $r = 16$ число периодичностей сокращается уже до нескольких сот.

Следующий шаг — наложение ограничений на структуру самого периода: будем предполагать, что это палиндром или комплементарный палиндром. Обычный палиндром – это фрагмент, одинаково читаемый в обоих направлениях (например, *caggac*). Комплементарный палиндром удовлетворяет этому свойству лишь при переименовании элементов алфавита в одном из двух прочтений по схеме $a \leftrightarrow t$, $c \leftrightarrow g$, соответствующей известному в молекулярной биологии отношению комплементарности. Так, комплементарным палиндромом является цепочка *catatg*.

Элемент самоподобия проявляется в том, что повторение палиндрома (случай «а» ниже) или комплементарного палиндрома (случай «б») приводит к «усилению» закономерности: возникает новый палиндром (соответственно, комплементарный палиндром) вдвое большей длины. Здесь повторяющиеся фрагменты подчёркнуты, расходящиеся стрелки сверху соответствуют палиндрому, сходящиеся — комплементарному

Таблица 1: Фрактальные спектры бактериальных геномов Y.pseudotuberculosis strain MD67 (NZ_CP009757) и Y.pestis Nepal516 (NC_008149). Порог $r = 16$

	Y. pseudotuberculosis			Y. pestis		
№	Позиция	Мономер	Длина	Позиция	Мономер	Длина
1	168841	atgtgtgta	28			
2	592138	ttaaaaaaatt	28	3091331	aatttttttaa	22
3	750327	attggtta	44	2932672	caattaac	68
4	934858	ataaata	16	2782647	tatttat	16
5	1416855	actaatca	57	1724582	actaatca	73
6	1451254	ggtaac	16	1759180	ggtaac	16
7	2163106	gtg	17			
8	2353115	agtttga	28			
9	2433892	aataa	17			
10	2532038	agataga	24	3280973	agataga	17
11	2612070	aaataaa	20	3362261	aaataaa	20
12	2677145	gcaacg	16	3429822	gcaacg	16
13				3643223	tggtgtggt	39
14	2891698	gcttgttcg	20	3643434	gcttgttcg	20
15	3632141	taaccaccaat	36			
16	3930443	cttttc	16			
17	3950750	acaca	17			
18	3988950	tcact	29			
19	4038025	tgaagaagt	111	192840	tgaagaagt	39
20	4049899	taacaat	69			
21	4567259	gggaggg	24			

палиндрому. При кратности повторений выше двух возникают множественные структуры, частично налагающиеся друг на друга.

а) $\ldots \overleftrightarrows{tac\,cat}\ \overleftrightarrows{tac\,cat}\ldots$; б) $\overrightarrow{\overleftarrow{cat\,atg}}\ \overleftarrow{\overrightarrow{cat\,atg}}$

Описанные выше структуры, образуемые точными («совершенными») тандемными повторами симметричных (в широком смысле) фрагментов, мы называем локальными фракталами. При наличии незначительных искажений внутри повторяющихся фрагментов или коротких вставок между ними используем термины «фракталоподобные структуры» или несовершенные локальные фракталы.

Полную совокупность $\Phi(S, r)$ локальных фракталов генома S с длинами $L \geq r$ назовём фрактальным спектром генома при пороге r. Для характерных длин рассматриваемых бактериальных геномов ($N \sim 10^6$н.п.) — это вполне обозримые совокупности, состоящие из нескольких десятков последовательностей. Для иллюстрации в табл. 1 приведены фрактальные спектры двух бактерий, относящихся к разным видам иерсиний: Yersinia pseudotuberculosis и Yersinia pestis.

Фрактальные структуры в таблице 1 упорядочены по позициям их вхождения в геном Y.pseudotuberculosis strain MD67 (NZ_CP009757). Анализ таблицы 1 показывает, что в большинстве случаев кратность повторения мономера в выделяемой периодичности ($k = L/|M|$) не является целым числом. Это означает, что периодичность начинается и (или) заканчивается неполным мономером. Например, строке 12 в левой половине таблицы соответствует периодичность ggcaacggcaacggca, периодом которой можно считать фрагмент ggcaac, не обладающий свойством симметрии. Сдвигом на один символ получаем симметричный мономер gcaacg. Соответствующая ему фрактальная структура начинается последним элементом мономера (g) и заканчивается тремя его начальными элементами (gca).

При сопоставимых длинах геномов Yersinia pseudotuberculosis MD67 и Yersinia pestis Nepal 516 (4729690 и 4534590, соответственно), в первом выявлено значительно больше локальных фракталов, чем во втором (20 против 11). При этом для десяти из одиннадцати фракталов генома Y. Pestis находится аналог (на уровне мономера) в геноме Y. Pseudotuberculosis, отличия наблюдаются лишь по количеству повторений мономера

в каждом фрактале. К аналогам мы относим: точно совпадающие мономеры (таких большинство, см. например, actaatca в обеих половинах таблицы); мономеры, образующие комплементарный инвертированный повтор (см., например, № 4 в обоих геномах); мономеры, переводимые друг в друга путём циклического сдвига одного из них и комплементарной подстановки (см., например, № 3). Единственная периодичность, представленная в геноме Y. Pestis Nepal516 (NC_008149), но отсутствующая в геноме Y. Pseudotuberculosis MD67, образована повторением мономера tggtgtggt. У других представителей вида Y. Pestis из нашей подборки этот фрактал отсутствует, т.е. он малоинформативен в плане классификации. Более того, этот фрактал встречается у двух других представителей вида Y. Pseudotuberculosis из нашей подборки.

Эксперимент с рандомизированными геномами обоих видов показал, что случайное перемешивание символов приводит (при том же пороге отбора по длине $r = 16$) к резкому сокращению числа выявляемых фракталов (в среднем по одному на геном, а иногда и ни одного) и их длин (не более 18 символов). Это говорит о том, что выявляемые в реальных геномах фрактальные структуры с длинами, превышающими пороговое значение, в большинстве случаев не относятся к категории случайных.

3 Сравнение подборок полных геномов Yersinia pseudotuberculosis и Yersinia pestis

Данный этап преследовал своей целью проверку на более объёмном материале (порядка полутора десятков геномов по каждому виду) тех отличий между видами, которые мы выявили при сравнении единичных их представителей. Систематизируем полученные результаты в общем виде.

1) Все представители Yersinia pestis содержат значимо меньшее количество локальных фракталов на геном (порядка 10–11) по сравнению с представителями вида Yersinia pseudotuberculosis (порядка 18–20). Этот фактор, сам по себе, может служить критерием для разделения представителей обоих видов.

2) В геномах обоих видов в качестве мономеров фрактальных структур абсолютно доминируют палиндромы. Тем не менее существует единственная структура с мономером в виде комплементарного палиндрома ggtacc (или accggt), представленная во всех геномах рассматриваемых видов. Из этого следует, что повторы комплементарных палиндромов конкретного организма представляют определённую опасность для нормального его функционирования.

3) Все мономеры, формирующие фрактальные структуры в геномах Yersinia pestis, представлены и во фрактальных структурах геномов Yersinia pseudotuberculosis. Однако, среди мономеров, типичных для представителей вида Yersinia pseudotuberculosis, много таких, которые отсутствуют у представителей вида Yersinia pestis (например, cttttc, atgtgtgta, acaca и др.).

4) Наиболее длинные локальные фракталы у представителей обоих видов формируются с участием мономеров tgaagaagt и actaatca. В то же время многие мономеры порождают локальные фракталы ограниченной длины. Например, встречающиеся почти у всех представителей обоих видов фракталы, построенные на основе мономера ataaata, все имеют длину 16, а на основе мономера ttaaaaaaatt – длину 22.

5) Кратность повторения одного и того же мономера в разных периодичностях может меняться и у представителей одного вида. Например, периодичности на основе мономера agataga имеют кратность от 17/7 до 108/7 для Yersinia pseudotuberculosis. Возможность использования этого параметра для целей классификации требует отдельного изучения.

Итак, детально описанные в данной работе понятия локального фрактала и фрактального спектра служат удобным инструментом для компактного описания геномов достаточно большой длины. Функциональная значимость таких структур не вызывает сомнений и подтверждается многочисленными примерами их участия в основных генетических процессах. Особенно значимой представляется возможность использования фрактальных структур в качестве маркеров для дифференциации на геномном уровне родственных видов или представителей одного вида, что иллюстрируется примером сопоставления близких по структуре бактериальных геномов, существенно отличающихся, тем не менее, по степени патогенности для человека.

Список литературы

[1] Смирнов Г.Б. Механизмы приобретения и потери генетической информации бактериальными геномами // Успехи молекулярной биологии, 2008, Т. 128, № 1, 52–76.

[2] Гусев В.Д., Мирошниченко Л.А., Чужанова Н.А. – Выявление фракталоподобных структур в ДНК-последовательностях // Information Science & Computing. International Book Series, №8: Classification, Forecasting, Data Mining. – ITHEA, Sofia, 2009.— P. 117–123.

[3] Gusev V.D., Miroshnichenko L.A., Titkova T.N., Dzhioev Y.P., Kozlova I.V., Paramonov A.I. — Structured RNA Markers for Genotyping of Tick-Borne Encephalitis Virus // Mathematical Biology and Bioinformatics, 2018, Vol. 13, № 1, P. 13 - 37, Scopus – 0.19, DOI: 10.17537/2018.13.13

[4] Сунцов В.В., Сунцова Н.И. Макро- и микроэволюция в проблеме происхождения и мировой экспансии возбудителя чумы Yersinia pestis // Известия РАН. Серия Биологическая, 2008, № 4, 389–395.

[5] Bercovier H., Mollaret HH, Alonso JM Intra- and interspecies relatidness of Yersinia pestis by DNA hybridization and its relationship to Yersinia pseudotuberculosis // Current Microbiology: 1980, 4(4), 225–229.

[6] Платонов М.Е. Молекулярно - генетическое изучение разнообразия и микроэволюции Yersinia pestis // Дисс. соиск. уч. ст. к.б.н. по спец. 03.02.03, 03.01.03 (микробиология), 2013, 196 с.

Гусев Владимир Дмитриевич — к.т.н., ст. науч.сотр.
Института математики им. С. Л. Соболева СО РАН;
e-mail: gusev@math.nsc.ru;
Мирошниченко Любовь Александровна — к.т.н., ст. науч.сотр.
Института математики им. С. Л. Соболева СО РАН;
e-mail: luba@math.nsc.ru;
Джиоев Юрий Павлович — к.б.н., ведущий науч.сотр.
НИИ Биомедицинских технологий
Иркутского государственного медицинского университета, Иркутск;
e-mail:alanir07@mail.ru.
Дата поступления — 30 апреля 2019 г.

РЕШЕНИЕ ЗАДАЧИ ТЕПЛООБМЕНА КРЫЛА САМОЛЁТА МЕТОДОМ МОНТЕ-КАРЛО

С. А. Гусев[1,2], В. Н. Николаев[3]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный технический университет, 630073, Новосибирск*
[3] *Сибирский научно-исследовательский институт авиации им. С.А. Чаплыгина, 630051, Новосибирск*

УДК 519.24
DOI: 10.24411/9999-016A-2019-10021

Предложен метод статистического моделирования для определения значений температуры в заданных точках крыла самолёта в случае обледенения. Метод основан на вероятностном представлении решения краевой задачи с подвижной границей для параболического уравнения и численном решении стохастических диференциальных уравнений. В работе рассмотрен двумерный случай.
Ключевые слова: обледенение крыла, краевая задача, стохастические диференциальные уравнения, метод Монте-Карло.

Введение

Обледенение самолётов приводит к нежелательным последствиям, таким как ухудшение аэродинамических качеств, потеря устойчивости и снижение управляемости [1]. Можно выделить три группы причин, приводящих к появлению ледяных наростов, и от которых зависит их вид:

- Характеристики окружающей среды, такие как, например, температура, концентрация водяных капель и частиц льда, скорость и направление ветра.
- Параметры режима полёта. Современные самолёты за короткое время могут преодолевать большие расстояния, что нередко приводит к быстрым и значительным изменениям внешних условий.
- Существенное значение имеют геометрические характеристики и покрытие самолёта.

Несмотря на многочисленные исследования в данном вопросе, а также разработанные практические рекомендации и меры, проблема обледенения самолётов остается актуальной и требует дальнейшего всестороннего внимания и исследования.

В данной работе предлагается численно-статистический метод для расчета значений температуры в заданных точках крыла самолёта. Предполагается, что крыло изготовлено из углепластика — материала с низкой теплопроводностью. Поскольку в результате обледенения крыла толщина его стенки увеличивается, температура в точках крыла в зоне ледяного нароста должна с течением времени изменяться иначе, чем температура в случае отсутствия обледенения. Таким образом, если установить датчики температуры на внутренней поверхности крыла, то по полученным измерениям можно диагностировать появление льда на внешней поверхности. Проведение расчётов температуры в заданных точках крыла является важной составной частью исследования процессов его обледенения.

Предполагается, что процессы теплообмена, происходящие в крыле самолёта, достаточно хорошо математически описываются двумерной краевой задачей для уравнения теплопроводности. Предложенный метод основан на вероятностном представлении решения краевой задачи в виде математического ожидания

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-01-00698-А).

ISBN 978-5-901548-42-4

функционала от случайного процесса диффузионного типа. Для получения оценки решения надо с использованием численного метода промоделировать большое количество траекторий этого случайного процесса и на основании полученных данных определить вероятностную оценку решения краевой задачи.

Особенность решаемой задачи состоит в том, что граница области, соответствующая участкам поверхности, на которых происходит обледенение, меняет своё положение с течением времени. В работе предлагается моделировать ледяные наросты с использованием треугольной сетки на участках обледенения. При этом аппроксимацией подвижной границы является ломаная линия, составленная из сторон треугольников этой треугольной сетки. Вводится взаимно однозначное преобразование, отображающее любой подвижный треугольник на неподвижный равнобедренный прямоугольный треугольник, катеты которого направлены вдоль координатных осей прямоугольной системы координат. Этот треугольник мы называем стандартным треугольником. С использованием этого преобразования моделирование траекторий случайного процесса в каждом подвижном треугольнике сводится к их моделированию на стандартном треугольнике.

1 Постановка задачи

Для определения значений температуры в заданных точках крыла самолёта мы решаем следующую краевую задачу

$$\frac{\partial u}{\partial t} = b(x,y)\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right), \quad (t,x,y) \in Q_T \equiv (0,T)\times G(t), \tag{1}$$

$$u(0,x,y) = \phi(x,y), \tag{2}$$

$$\frac{\partial u}{\partial n} + \eta(t,x,y)u + \gamma(t,x,y) = 0, \quad (t,x,y) \in (0,T)\times \Gamma(t), \tag{3}$$

где $G(t) = G_1 \cup G_2(t)$; G_1 — область, изменения переменных x, y в зоне крыла; $G_2(t)$ — область, соответствующая обледенению крыла; $\Gamma(t)$ — граница области $G(t)$; n — единичный вектор внутренней нормали; $b(x,y) = \begin{cases} b_1, \ (x,y) \in G_1 \ , \\ b_2, \ (x,y) \in G_2(t). \end{cases}$

В краевой задаче (1)–(3) коэффициент температуропроводности терпит разрыв на линии соприкосновения льда и крыла. Для краевых задач параболического типа с разрывными коэффициентами в [2] доказаны существование и единственность их обобщенных решений в банаховом пространстве $V_2^{0,1}(Q_T)$ с нормой $|u|_{Q_T} = \max_{0\le t\le T}\|u(t,p)\|_{L_2} + \left(\int\limits_{Q_T} u_p^2 dpdt\right)^{\frac{1}{2}}$ (в нашем случае $p = (x,y)$) .

Также в [2] доказано (Теорема 5.1, Глава III), что обобщённое решение краевой задачи с разрывными коэффициентами устойчиво в относительно вариаций коэффициентов. Опираясь на этот факт, в работе [3] для получения приближенного решения краевой задачи на основе вероятностного представления предложено использовать сглаживание разрывных коэффициентов. В дальнейшем будем предполагать, что в задаче (1)–(3) над коэффициентами параболического оператора уже проведена процедура сглаживания.

2 Решение краевой задачи методом Монте-Карло

Для получения вероятностного представления решения задачи (1)–(3) мы задаём случайный процесс $(X_\cdot, Y_\cdot)$ в области $G(t) \subset \mathbb{R}^2$ $s \in [0,T]$, который является решением системы стохастических дифференциальных уравнений (СДУ)

$$\begin{aligned}
X_s &= x + \int\limits_{T-t}^{s} \sigma(X_v, Y_v)\, dW_{1,v} + k_{1,s},\\
Y_s &= y + \int\limits_{T-t}^{s} \sigma(X_v, Y_v) dW_{2,v} + k_{2,s},\\
k_{i,s} &= \int\limits_{T-t}^{s} n_i(X_v, Y_v)\, d|k_v|, \ (i = 1,2),\\
|k_s| &= \int\limits_{T-t}^{s} \chi_\Gamma(X_v, Y_v)\, d|k_v|,\\
U_s &= \exp\left(\int\limits_{T-t}^{s} \eta_\nu d|k_\nu|\right), \quad V_s = \int\limits_{T-t}^{s} \gamma_v U_v\, d|k_v|,
\end{aligned} \tag{4}$$

где W_1, W_2 — независимые винеровские процессы; $\sigma = \sqrt{2b}$; $|k.|$ — скалярный случайный процесс, $|k_{T-t}| = 0$, его значения возрастают, только при тех $s \in (0,T)$, когда (X_s, Y_s) достигает границы G.

Решение краевой задачи (1)–(3) в точке $(t,x,y) \in Q_T$ можно определить по формуле

$$u(t,x,y) = \mathrm{E}_{T-t,x,y}\big[\varphi(X_T, Y_T)U_T + V_T\big]. \tag{5}$$

где $\mathrm{E}_{T-t,x,y}$ — символ математического ожидания относительно вероятностной меры, соответствующей случайному процессу, выходящему в момент времени $T-t$ из точки (x,y).

Приближенные значения решения краевой задачи (1)–(3) можно получать путем численного моделирования траекторий системы (4) и вычисления среднего значения выражения под знаком математического ожидания в (5). Для численного решения системы СДУ мы используем модифицированный метод Эйлера (см., например, [4]). Применительно к системе СДУ (4) вычисления с шагом интегрирования делаются по схеме:

$$X^{\Delta}_{i+1} = x_i + h^{\frac{1}{2}}\sigma_i\xi_{1,i}, \tag{6}$$

$$Y^{\Delta}_{i+1} = y_i + h^{\frac{1}{2}}\sigma_i\xi_{2,i}, \tag{7}$$

$$x_{i+1} = X^{\Delta}_{i+1} + (\Delta_{i+1}K)n_{1,i+1}\chi_{R^2\setminus\bar{G}}(X^{\Delta}_{i+1}, Y^{\Delta}_{i+1}), \tag{8}$$

$$y_{i+1} = Y^{\Delta}_{i+1} + (\Delta_{i+1}K)n_{2,i+1}\chi_{R^2\setminus\bar{G}}(X^{\Delta}_{i+1}, Y^{\Delta}_{i+1}), \tag{9}$$

$$U_{i+1} = \begin{cases} 1,\ i=0, \\ U_i(1+\eta_i)\chi_{\Gamma}(x_{i+1}, y_{i+1})\Delta_{i+1}K),\ i>0 \end{cases}, \tag{10}$$

$$V_{i+1} = \begin{cases} 0,\ i=0, \\ V_i + \gamma_i\chi_{\Gamma}(x_{i+1}, y_{i+1})U_{i+1}\Delta_{i+1}K),\ i>0 \end{cases}, \tag{11}$$

$$\Delta_{i+1}K = \begin{cases} 0,\ (X^{\Delta}_{i+1}, Y^{\Delta}_{i+1}) \in \bar{G}(t_i), \\ \rho_{n,\bar{G}}(X^{\Delta}_{i+1}, Y^{\Delta}_{i+1}),\ (X^{\Delta}_{i+1}, Y^{\Delta}_{i+1}) \notin \bar{G}(t_i), \end{cases} \tag{12}$$

где $\rho_{n,\bar{G}}(x,y)$ — расстояние от точки (x,y) до её нормальной проекции на $\bar{G}$. Здесь черта над обозначением множества означает его замыкание. В формулах (9)–(15), как обычно, индексы i, указывают на значение функции в соответствующем узле равномерной сетки с шагом на отрезке.

3 Моделирование образования льда треугольниками

Как было сказано во Введении, построение подобласти $G_2(t)$, которая соответствует ледяному наросту, делается с помощью треугольников, координаты вершин которых являются заданными функциями t. Для построения этих треугольников на поверхности крыла вводится сетка по его контуру, и в каждом узле этой сетки вдоль направления внешней нормали задаётся отрезок, как показано на рис.1. Длина каждого такого отрезка есть заданная функция времени, которая определяет движение границы в данном узле. Если в месте положения некоторого узла обледенения не происходит, то длина соответствующего отрезка равна нулю. Концы этих отрезков являются вершинами ломаной линии, с помощью которой мы моделируем поверхность ледяного нароста. В результате соединения концов отрезков на внешней поверхности крыла образуются четырехугольники. Эти четырехугольники диагоналями разбиваются на треугольники, которые таким образом образуют треугольную сетку, см. рис. 2.

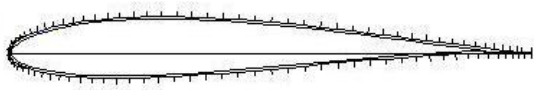


Рис. 1: Разбиение контура крыла на отрезки

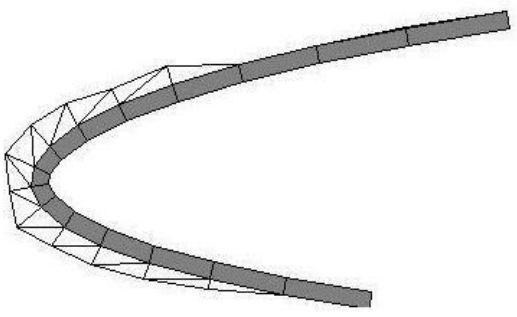


Рис. 2: Моделирование ледяного нароста

4 Моделирование траекторий случайного процесса на движущихся треугольниках

Моделирование траекторий случайного процесса в подобласти $G_2(t)$, которая состоит из треугольников с движущимися вершинами, осуществляется с использованием взаимно однозначного отображения каждого такого треугольника на неподвижный прямоугольный равнобедренный треугольник, который мы называем стандартным. Взаимно однозначное отображение строится следующим образом. Пусть дан треугольник $\mathrm{T} \subseteq G_2(t)$ с координатами вершин (x_1, y_1), (x_2, y_2), (x_3, y_3). Зададим в том же порядке соответствие вершинам треугольника T вершины стандартного треугольника $\overline{\mathrm{T}}$: $(0,0)$, $(0,1)$, $(1,0)$. В вещественном векторном пространстве $\mathbb{R}^3$ определим две плоскости:

1) плоскость $\mathcal{X}$, проходящую через точки
$\mathcal{X}_1 = (0,0,x_1), \quad \mathcal{X}_2 = (0,1,x_2), \quad \mathcal{X}_3 = (1,0,x_3)$;

2) плоскость $\mathcal{Y}$, проходящую через точки
$\mathcal{Y}_1 = (0,0,y_1), \quad \mathcal{Y}_2 = (0,1,y_2), \quad \mathcal{Y}_3 = (1,0,y_3)$.

$$A_x\bar{x} + B_x\bar{y} + C_x x + D_x = 0, \quad A_y\bar{x} + B_y\bar{y} + C_y y + D_y = 0. \tag{13}$$

При этом коэффициенты плоскостей имеют вид:

$$\begin{aligned} A_x &= x_3 - x_1, & A_y &= y_3 - y_1, \\ B_x &= x_2 - x_1, & B_y &= y_2 - y_1, \\ C_x &= -1, & C_y &= -1, \\ D_x &= x_1, & D_y &= y_1. \end{aligned}$$

Из уравнений плоскостей (13) получаем равенства, определяющие взаимно однозначное соответствие между точками (x, y) и $(\bar{x}, \bar{y})$ треугольников T и $\overline{\mathrm{T}}$ соответственно.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_3 - x_1 & x_2 - x_1 \\ y_3 - y_1 & y_2 - y_1 \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \tag{14}$$

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} y_2 - y_1 & x_1 - x_2 \\ y_1 - y_3 & x_3 - x_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix}, \tag{15}$$

где $\Delta = (x_3 - x_1)(y_2 - y_1) - (y_3 - y_1)(x_2 - x_1)$.

4.1 Коэфициенты СДУ на стандартном треугольнике

Для того, чтобы перенести моделирование траекторий случайного процесса с подвижного треугольника на стандартный треугольник, необходимо определить, как изменятся коэффициенты сноса и диффузии при использовании замены переменных, определяемой равенством (15). Для этой цели сделаем замену переменных (14) в уравнении (1).

Обозначим $\begin{pmatrix} f_x(t, \bar{x}, \bar{y}) \\ f_y(t, \bar{x}, \bar{y}) \end{pmatrix}$ правую часть (14). Решение уравнения теплопроводности при $t \in (0, T)$ в точке $(\bar{x}, \bar{y})$ стандартного треугольника будем обозначать как $\bar{u}(t, \bar{x}, \bar{y})$. Тогда, используя взаимно однозначное соответствие между точками треугольников T и $\overline{\mathrm{T}}$, можно определить решение уравнения теплопроводности на $\overline{\mathrm{T}}$ с помощью равенства

$$\bar{u}(t, \bar{x}, \bar{y}) = u(t, f_x(t, \bar{x}, \bar{y}), f_y(t, \bar{x}, \bar{y})). \tag{16}$$

После дифференцирования (16) по t получим

$$\frac{\partial u}{\partial t} = \frac{\partial \bar{u}}{\partial t} - \frac{\partial u}{\partial x}\frac{\partial f_x}{\partial t} - \frac{\partial u}{\partial y}\frac{\partial f_y}{\partial t}. \tag{17}$$

Дифференцируя (16) по $\bar{x}$ и $\bar{y}$, получаем соотношение производных решения в треугольниках T и $\overline{\mathrm{T}}$ по пространственным переменным

$$\begin{pmatrix} \frac{\partial \bar{u}}{\partial \bar{x}} \\ \frac{\partial \bar{u}}{\partial \bar{y}} \end{pmatrix} = \begin{pmatrix} x_3 - x_1 & y_3 - y_1 \\ x_2 - x_1 & y_3 - y_1 \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{pmatrix}. \tag{18}$$

Обращая равенство (18), имеем

$$\begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} y_2 - y_1 & x_1 - x_2 \\ y_1 - y_3 & x_3 - x_1 \end{pmatrix} \begin{pmatrix} \frac{\partial \bar{u}}{\partial \bar{x}} \\ \frac{\partial \bar{u}}{\partial \bar{y}} \end{pmatrix}. \tag{19}$$

Из (19) следует, что в системе СДУ на стандартном треугольнике появляется вектор сноса, компоненты которого получаются из (17) и (19)

$$\frac{1}{\Delta} \begin{pmatrix} \frac{\partial f_x}{\partial t}(y_2 - y_1) + \frac{\partial f_y}{\partial t}(y_1 - y_3) \\ \frac{\partial f_x}{\partial t}(x_1 - x_2) + \frac{\partial f_y}{\partial t}(x_3 - x_1) \end{pmatrix}.$$

Для вывода формул элементов матрицы диффузии на стандартном треугольнике определим вторые производные $\bar{u}$ по $\bar{x}$ и $\bar{y}$ аналогично тому, как это было сделано для первых производных в (18)

$$\begin{pmatrix} \frac{\partial^2 \bar{u}}{\partial \bar{x}^2} & \frac{\partial^2 \bar{u}}{\partial \bar{x} \partial \bar{y}} \\ \frac{\partial^2 \bar{u}}{\partial \bar{y} \partial \bar{x}} & \frac{\partial^2 \bar{u}}{\partial \bar{y}^2} \end{pmatrix} = \begin{pmatrix} x_3 - x_1 & y_3 - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{pmatrix} \begin{pmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial y \partial x} & \frac{\partial^2 u}{\partial y^2} \end{pmatrix} \begin{pmatrix} x_3 - x_1 & x_2 - x_1 \\ y_3 - y_1 & y_2 - y_1 \end{pmatrix}. \tag{20}$$

Из последнего равенства получим выражение вторых производных решения на подвижном треугольнике через вторые производные на стандартном треугольнике

$$\begin{pmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial y \partial x} & \frac{\partial^2 u}{\partial y^2} \end{pmatrix} = \frac{1}{\Delta^2} \begin{pmatrix} y_2 - y_1 & y_1 - y_3 \\ x_1 - x_2 & x_3 - x_1 \end{pmatrix} \begin{pmatrix} \frac{\partial^2 \bar{u}}{\partial \bar{x}^2} & \frac{\partial^2 \bar{u}}{\partial \bar{x} \partial \bar{y}} \\ \frac{\partial^2 \bar{u}}{\partial \bar{y} \partial \bar{x}} & \frac{\partial^2 \bar{u}}{\partial \bar{y}^2} \end{pmatrix} \begin{pmatrix} y_2 - y_1 & x_1 - x_2 \\ y_1 - y_3 & x_3 - x_1 \end{pmatrix}. \tag{21}$$

На основании равенств (1), (19), (21) можем получить уравнение теплопроводности на стандартном треугольнике

$$\begin{aligned} \frac{\partial \bar{u}}{\partial t} = \quad & \frac{1}{\Delta} \left(\frac{\partial f_x}{\partial t}(y_2 - y_1) + \frac{\partial f_y}{\partial t}(y_1 - y_3) \right) \frac{\partial \bar{u}}{\partial \bar{x}} + \frac{1}{\Delta} \left(\frac{\partial f_x}{\partial t}(x_1 - x_2) + \frac{\partial f_y}{\partial t}(x_3 - x_1) \right) \frac{\partial \bar{u}}{\partial \bar{y}} \\ & + \frac{b_2}{\Delta^2} \left[(x_2 - x_1)^2 + (y_2 - y_1)^2 \right] \frac{\partial^2 \bar{u}}{\partial \bar{x}^2} + \frac{b_2}{\Delta^2} \left[(x_3 - x_1)^2 + (y_3 - y_1)^2 \right] \frac{\partial^2 \bar{u}}{\partial \bar{y}^2}. \end{aligned} \tag{22}$$

Следовательно, матрица коэффициентов в винеровском интеграле соответствующей системы СДУ на стандартном треугольнике имеет вид

$$\frac{\sqrt{2b_2}}{\Delta} \begin{pmatrix} ((x_2 - x_1)^2 + (y_2 - y_1)^2)^{\frac{1}{2}} & 0 \\ 0 & ((x_3 - x_1)^2 + (y_3 - y_1)^2)^{\frac{1}{2}} \end{pmatrix}.$$

4.2 Численная схема на стандартном треугольнике

При попадании траектории случайного процесса в любой подвижный треугольник, все расчеты по схеме Эйлера в его внутренних точках автоматически переносятся на стандартный треугольник. Принадлежность точки $(\bar{x}, \bar{y})$ стандартному треугольнику определяется с помощью простых неравенств:

$$\bar{x} > 0, \quad \bar{y} > 0, \quad \bar{x} + \bar{y} < 1\,.$$

Формулы (6), (7) при вычислениях на стандартном треугольнике с учётом изменений вектора сноса и матрицы диффузии имеют вид

$$X_i^{\Delta} = x_i + \frac{1}{\Delta_i} \left((f_{x,i+1} - f_{x,i})(y_{2,i} - y_{1,i}) + (f_{y,i+1} - f_{y,i})(y_{1,i} - y_{3,i}) \right) + h^{\frac{1}{2}} \frac{\sqrt{b_2}}{\Delta_i} \left[(x_{2,i} - x_{1,i})^2 + (y_{2,i} - y_{1,i})^2 \right]^{\frac{1}{2}},$$

$$Y_i^{\Delta} = y_i + \frac{1}{\Delta_i} \left((f_{x,i+1} - f_{x,i})(x_{1,i} - x_{2,i}) + (f_{y,i+1} - f_{y,i})(x_{3,i} - x_{1,i}) \right) + h^{\frac{1}{2}} \frac{\sqrt{b_2}}{\Delta_i} \left[(x_{3,i} - x_{1,i})^2 + (y_{3,i} - y_{1,i})^2 \right]^{\frac{1}{2}}.$$

Заключение

Предложен метод расчета значений температуры крыла самолёта в случае его обледенения, основанный на вероятностном представлении решения двумерной краевой задачи для уравнения теплопроводности и численном решении СДУ. Рассматривается случай, когда крыло, изготовлено из композитных материалов с низкой теплопроводностью. Для получения приближенного значения температуры в заданной точке необходимо промоделировать большое количество траекторий случайного процесса. Моделирование ледяного нароста на поверхности крыла делается с помощью сетки, состоящей из треугольников с подвижными вершинами. Расчёт траекторий случайного процесса на таких треугольниках сводится к вычислению траекторий на неподвижном треугольнике с использованием введённого взаимно однозначного отображения между треугольниками.

Предложенный метод может быть обобщён на трёхмерный случай.

Список литературы

[1] Ratvasky T.P., Barnhart B.P., Lee S. Current methods for modeling and simulating icing effects on aircraft performance. Stab. Control, NASA/TM-2008-215453. Google Scholar

[2] Ладыженская О.А., Солонников В.А., Уральцева Н.Н. Линейные и квазилинейные уравнения параболического типа. М.: Наука, 1967.

[3] Гусев С.А. Применение СДУ к оценке решения уравнения теплопроводности с разрывными коэффициентами// Сиб. журн. вычисл. математики. 2015. Т. 18, № 2. С. 147–161.

[4] Gusev S.A. Using SDE for solving inverse parabolic boundary value problem with a Neumann boundary condition // Russian Journal of Numerical Analysis and Mathematical Modelling. 2007. Vol. 22, № 5. P. 449–470.

Гусев Сергей Анатольевич — д.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН; профессор Новосибирского государственного технического университета;
e-mail: sergey.gusev@sscc.ru;
Николаев Владимир Николаевич — д.т.н., учёный секретарь Сибирского научно-исследовательского института авиации им. С.А. Чаплыгина
e-mail: nikvla50@mail.ru
Дата поступления — 30 апреля 2019 г.

ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЗАДАЧ СО СЛУЧАЙНЫМИ ВХОДНЫМИ ДАННЫМИ

Б. С. Добронец[1], О. А. Попова[1]

[1] *Сибирский федеральный университет, 660074, Красноярск*

УДК 519.24
DOI: 10.24411/9999-016A-2019-10022

В работе рассмотрены вопросы численного моделирования задач со случайными входными данными. Подход основан на использовании вычислительного вероятностного анализа для вычисления вероятностных расширений функций случайных переменных. Оценена вычислительная сложность подобных задач и приведены пути реализации на основе рекурсивного распараллеливания. Приведены численные примеры решений систем линейных алгебраических уравнений со случайными входными данными. Рассмотрены вопросы совместного использования вычислительного вероятностного анализа и метода Монте-Карло.

Ключевые слова: вычислительный вероятностный анализ, вероятностные расширения, системы линейных алгебраических уравнений, Монте-Карло.

Введение

При моделировании в условиях неопределенности необходимо проводить численное исследование моделей со случайными входными данными. Такие модели традиционно изучаются методами Монте-Карло. Метод Монте-Карло является мощным и универсальным подходом, но один из серьезных недостатков, это низкая скорость сходимости.

Начиная с 1960-х годов для моделирования в условиях неопределенности стали разрабатываться методы не связанные с Монте-Карло. Основной подход дающий гарантированные оценки — интервальный анализ. Однако интервальный анализ вычисляют только границы случайных процессов, не рассматривая их внутренние распределения.

В работе разрабатывается подход, который использует вычислительный вероятностный анализ (ВВА) для решения различных задач со стохастической неопределенностью в данных. Основой вычислительного вероятностного анализа являются численные операции над функциями плотности вероятности случайных переменных. Это операции “ $+$ ”, “ $-$ ”, “ $\cdot$ ”, “ $/$ ”, “ $\uparrow$ ”, “ $\max$ ”, “ $\min$ ”, а также бинарные отношения “ $\leq$ ”, “ $\geq$ ” и некоторые другие [1].

Используя арифметику над функциями плотности вероятности и вероятностные расширения, можно построить численные методы, которые позволят решать системы линейных и нелинейных алгебраических уравнений, задач оптимизации со случайными входными параметрами [2]. Далее будем использовать кусочно-полиномиальные модели для представления функций плотности вероятности [3–5] и обозначать функции плотности вероятности случайных переменных a, b, c жирным шрифтом $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$.

Во многих случаях проблемы математического моделирования можно свести к численному анализу функций [6]

$$z = f(x_1, x_2, \ldots, x_n).$$

В тех случаях, когда известны плотности вероятности $x_1, x_2, \ldots, x_n$ для нахождения плотности вероятности случайной величины z можно использовать метод Монте-Карло [7]. Использование ВВА в ряде случаев позволяет значительно сократить время и поднять точность оценки функции плотности вероятности случайной величины $\boldsymbol{z}$. В работе [8] на примере сложения четырех равномерных случайных величин было показано, что ВВА эффективнее метода Монте-Карло, более чем в 10^4 раз.

ISBN 978-5-901548-42-4

1 Вероятностные расширения

Одна из самых важных проблем, с которыми сталкивается ВВА — построение функции плотности вероятности функций от случайных величин.

Начнем с общего случая, когда $(x_1, \ldots, x_n)$ — система непрерывных случайных величин с совместной функцией плотности вероятности $p(x_1, \ldots, x_n)$ и случайная величина z является функцией $f(x_1, \ldots, x_n)$

$$z = f(x_1, \ldots, x_n).$$

Под *вероятностным расширением* функции f будем называть функцию плотности вероятности случайной величины z.

Носителем функции плотности вероятности $\boldsymbol{f}$ будем называть множество

$$\mathrm{supp}(\boldsymbol{f}) = \{x | \boldsymbol{f}(x) > 0\}.$$

Рассмотрим процедуру вычисления вероятностных расширений в одномерном случае. Пусть дана функциональная зависимость

$$z = f(x),$$

где x — случайная величина, $\boldsymbol{x}$ — функция плотности вероятности случайной величины x с носителем $[\underline{x}, \overline{x}]$. Далее $\{x_i(z) \in [\underline{x}, \overline{x}] | i = 1, \ldots, n\}$ - корни уравнения $z = f(x)$.

Пусть необходимо найти функцию плотности вероятности $\boldsymbol{z}$ случайной величины z или вероятностное расширения $\boldsymbol{f}(\cdot, \boldsymbol{x})$ функции $f(x)$. Мы можем представить решение в виде [9]

$$\boldsymbol{f}(\xi, \boldsymbol{x}) = \sum_{i=1}^{n} \frac{\boldsymbol{x}(x_i(\xi))}{|f'(x_i(\xi))|}.$$

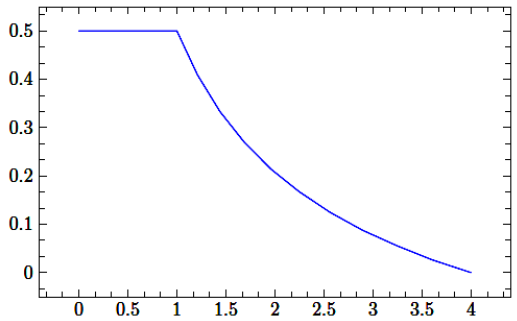


Рис. 1: Вероятностное расширение функции $f = ax^2 + bx$

Пример 1. Рассмотрим построение вероятностного расширения функции

$$f = ax^2 + bx, a \geq 0, b \geq 0,$$

x — случайная переменная распределенная на $[0, 2]$ по треугольному закону

$$\boldsymbol{x}(\xi) = \begin{cases} \xi, & \text{if } \xi \in [0, 1), \\ 2 - \xi, & \text{if } \xi \in [1, 2]. \end{cases}$$

Далее

$$r_1(z) = -\frac{\sqrt{4az + b^2} + b}{2a}, \quad r_2(z) = \frac{\sqrt{4az + b^2} - b}{2a},$$

корни $z = f(r_i)$. Выберем положительный корень $r(\xi) = r_2(\xi) = \frac{\sqrt{4a\xi + b^2} - b}{2a}$, и

$$[ax^2 + bx]' = 2ax + b = \sqrt{4az + b^2}.$$

Окончательно

$$\boldsymbol{f}(\xi, \boldsymbol{x}) = \boldsymbol{x}(r(\xi))/\sqrt{4a\xi + b^2}.$$

Теорема 1. [9] Пусть $\boldsymbol{f}(\cdot, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ вероятностное расширение $f(x_1, x_2, \ldots, x_n)$ и для всех вещественных t функция $\boldsymbol{f}(\cdot, t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ — вероятностное расширение $f(t, x_2, \ldots, x_n)$. Тогда

$$\boldsymbol{f}(z, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \int_{\underline{x}}^{\overline{x}} \boldsymbol{x}_1(t)\boldsymbol{f}(z, t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)dt \tag{1}$$

Замечание. Из теоремы 1 вытекает возможность рекурсивного вычисления вероятностных расширений общего вида, сведением их к вычислению одномерных вероятностных расширений.

Рассмотрим вычисление интеграла (1). Для определенности представим (1) в виде квадратуры

$$\int_{\underline{x}}^{\overline{x}} \boldsymbol{x}_1(t)\boldsymbol{f}(z, t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)dt \approx \sum_{l=1}^{m} \gamma_l \boldsymbol{x}_1(t_l)\boldsymbol{f}(z, t_l, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$$

Далее, для вычисления $\boldsymbol{f}(z, t_l, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ мы также будем использовать квадратуры и так далее. В общем случае, это NP-трудная проблема и для ее решения можно использовать рекурсивное распараллеливание.

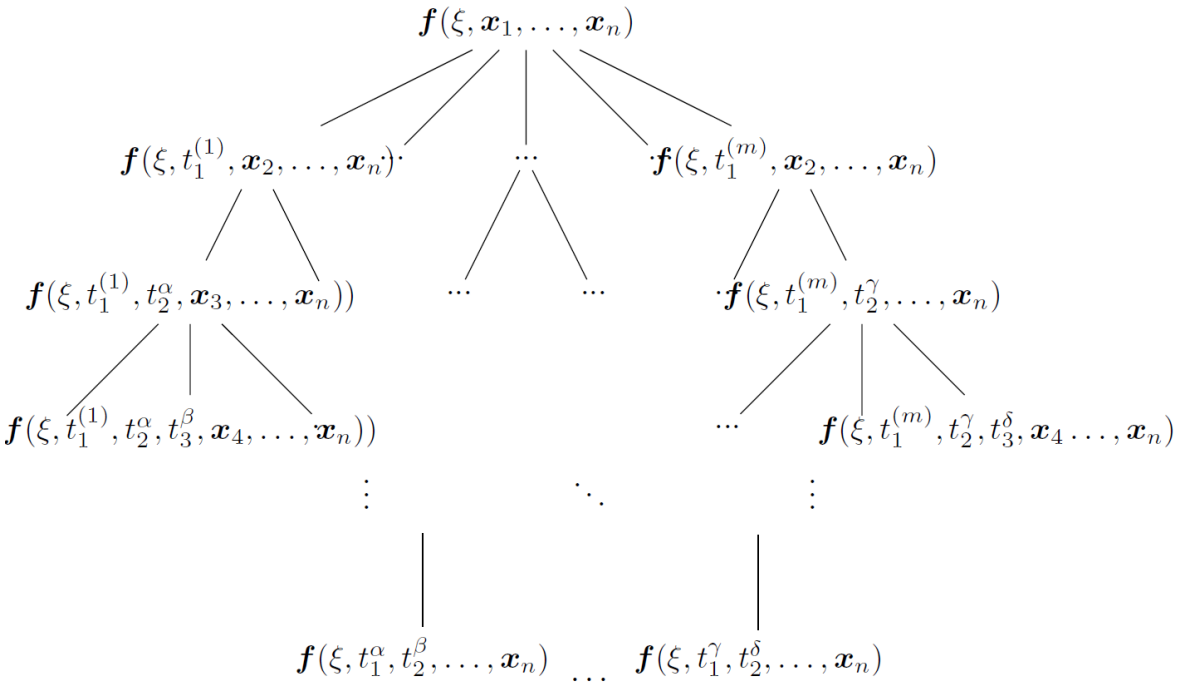


Рис. 2: Дерево параллельно-рекурсивной организации вычислительного процесса

На рис. 2 показано дерево параллельно-рекурсивной организации вычислительного процесса. Таким образом, на нижнем уровне необходимо вычислять вероятностные расширения для функций только одной переменной. Все расчеты на каждом слое независимы и могут быть рассчитаны одновременно.

Пусть даны непрерывных независимые случайные величины x, y с функциями плотности вероятности $\boldsymbol{x}, \boldsymbol{y}$. Плотности, возникающие в результате арифметических операций со случайными величинами

$$(\boldsymbol{x} + \boldsymbol{y})(\xi) = \int_{\underline{x}}^{\overline{x}} \boldsymbol{x}(t)\boldsymbol{y}(\xi - t)dt,$$

если $\mathrm{supp}(\boldsymbol{x}) = [\underline{x}, \overline{x}]$ and $\underline{x} > 0$ тогда

$$(\boldsymbol{x}\boldsymbol{y})(\xi) = \int_{\underline{x}}^{\overline{x}} \frac{1}{t}\boldsymbol{x}(t)\boldsymbol{y}(\frac{\xi}{t})dt,$$

$$(\boldsymbol{y}/\boldsymbol{x})(\xi) = \int_{\underline{x}}^{\overline{x}} t\boldsymbol{x}(t)\boldsymbol{y}(t\xi)dt.$$

Эти арифметические операции полностью совпадают с классическими операциями над функциями плотности вероятности.

Пусть $f(x_1, \ldots, x_n)$ — рациональная функция. Для построения вероятностного расширения функции f мы заменим арифметические операции на вероятностные арифметические операции, а переменные x_1, x_2, $\ldots$, x_n заменяются на функции плотности вероятности их возможных значений. Получившиеся вероятностные расширение $\boldsymbol{f}$ будем называть как *естественное вероятностное расширение.*

Случай 1. [1] Пусть $x_1, \ldots, x_n$ — независимые случайные величины. Если $f(x_1, \ldots, x_n)$ является рациональным выражением, в котором каждая переменная x_i встречается не более одного раза, тогда естественное вероятностное расширение аппроксимирует вероятностное расширение.

Случай 2. [1] Пусть функция $f(x_1, \ldots, x_n)$ может быть заменой переменных, поэтому что $f(z_1, \ldots, z_k)$ является рациональной функцией переменных $z_1, \ldots, z_k$, удовлетворяющих условиям случая 1. Переменная z_i является функцией x_i, $i \in Ind_i$. и Ind_i взаимно не пересекаются. Предположим, для каждого z_i можно построить вероятностное расширение. Тогда естественное расширение $f(z_1, \ldots, z_k)$ будет аппроксимировать вероятностное расширение $f(x_1, \ldots, x_n)$.

2 Системы линейных алгебраических уравнений

В качестве одного из примеров численного моделирования рассмотрим решение системы линейных алгебраических уравнений

$$Ax = b, \tag{2}$$

где $A = (a_{ij})$ — случайная матрица и $b = (b_i)$ — случайный вектор соответственно. Предположим, что случайная матрица A и вектор b имеют независимые компоненты с плотностями вероятности $\boldsymbol{A} = (\boldsymbol{a}_{ij})$, $\boldsymbol{b} = (\boldsymbol{b}_i)$ соответственно

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{pmatrix}.$$

Носитель множества решений может быть представлен в виде [2]

$$\mathcal{X} = \{x | Ax = b, A \in \mathrm{supp}(\boldsymbol{A}), b \in \mathrm{supp}(\boldsymbol{b})\}.$$

Построим вероятностное расширение $\boldsymbol{x}_1(\cdot, \boldsymbol{A}, \boldsymbol{b})$

$$\boldsymbol{x}_1(\cdot, \boldsymbol{A}, \boldsymbol{b}) = \frac{\begin{vmatrix} \boldsymbol{b}_1 & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{b}_n & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{vmatrix}}{\begin{vmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{vmatrix}}$$

или

$$\boldsymbol{x}_1(\xi, \boldsymbol{A}, \boldsymbol{b}) = \int_{\Omega} \boldsymbol{a}_{12}(t_{12}) \ldots \boldsymbol{a}_{nn}(t_{nn}) \frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})}(\xi) dt_{12} \ldots dt_{nn}, \tag{3}$$

где $\Delta_i(t_{12}, \ldots, t_{nn}) \in R$ миноры из метода Крамера для решения СЛАУ, $t_{ij} \in \mathrm{supp}(\boldsymbol{a}_{ij})$, $\Omega = \mathrm{supp}(\boldsymbol{a}_{12}) \times \ldots \mathrm{supp}(\boldsymbol{a}_{nn})$. Выражение

$$\left(\frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})} \right)(\xi)$$

вычисляется используя вероятностные арифметики.

Рассмотрим вопрос построения совместной функции распределения p_x вектора $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$. Рассмотрение начнем со случая, когда матрица A является детерминированной:

$$A\boldsymbol{x} = \boldsymbol{b}.$$

Тогда

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}.$$

Пусть для определенности совместная функция распределения $p(b_1, \ldots, b_n)$ вектора $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ задана на $\Omega = [\underline{b}_1, \overline{b}_1] \times \ldots \times [\underline{b}_n, \overline{b}_n]$. Ω — параллелепипед с ребрами $e_i, i = 1, \ldots, n$. Возьмем $e_i, i = 1, \ldots, n$ в качестве базисных, начало координат выберем в точке $O = (\underline{b}_1, , \ldots, \underline{b}_n)$. Тогда матрица A^{-1} как аффинное преобразование преобразует Ω в Ω', $e_i, i = 1, \ldots, n$ в $e'_i, i = 1, \ldots, n$ и O в O'. Таким образом, значение плотности вероятности $p(b_1, \ldots, b_n)$ в базисе $\{O, e_i, i = 1, \ldots, n\}$ преобразуется в значение $\alpha p(b_1, \ldots, b_n)$ в базисе $\{O', e'_i, i = 1, \ldots, n\}$, где

$$\alpha(A) = \frac{|\Omega|}{|\hat{\Omega}|}.$$

Для вычисления плотности вероятности в некоторой точке $x = (x_1, \ldots, x_n)$ необходимо найти ее координаты в базисе $\{O'e'_i, i = 1, \ldots, n\}$: $x' = (x'_1, \ldots, \hat{x}'_n)$. Заметим, что

$$x'(A) = A(x - O').$$

Следовательно

$$p_x(x_1, \ldots, x_n) = \alpha(A)p(x'(A)).$$

Окончательно получаем

$$p_x(x_1, \ldots, x_n) = \int_\Omega \boldsymbol{a}_{11}(t_{11}) \ldots \boldsymbol{a}_{nn}(t_{nn})\alpha(A(t))p(x'(A(t)))dt_{11} \ldots dt_{nn}. \tag{4}$$

Пример 2. Рассмотрим систему линейных алгебраических уравнений

$$Ax = b, \tag{5}$$

Пусть $A = (a_{ij})$ — случайная матрица $n = 2$. Элементы матрицы $\boldsymbol{A}$ независимы и распределены по треугольному закону, $\boldsymbol{a}_{11}, \boldsymbol{a}_{22}$ распределены на интервале $[2, 4]$, a_{21}, a_{12} распределены на отрезке $[-1, 1]$. Вектор $\boldsymbol{b}$ состоит из независимых компонент $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, распределенных по треугольному закону на отрезке $[0, 2]$.

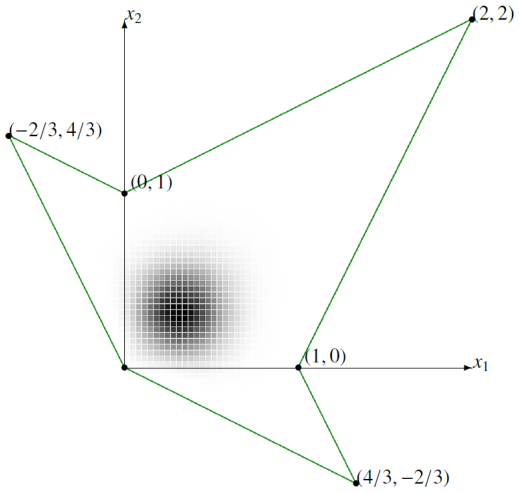


Рис. 3: Граница множества решений и совместная функция плотности вероятностей вектора (x_1, x_2)

Вектор (x_1, x_2) решение (5)

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}},$$

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

На рис. 3 показана совместная плотность вероятности вектора (x_1, x_2). Значение вероятности представлено оттенками серого. Сплошная линия — граница множества решений.

В этом случае для для вычисления x_1 заменим a_{22}, a_{12} на t_{22}, t_{12}. Вычислим интеграл численно для разных ξ

$$\boldsymbol{x}_1(\xi) = \int_\Omega \boldsymbol{a}_{22}(t_{22})\boldsymbol{a}_{12}(t_{12}) \left(\frac{t_{22}\boldsymbol{b}_1 - t_{12}\boldsymbol{b}_2}{\boldsymbol{a}_{11}t_{22} - t_{12}\boldsymbol{a}_{21}} \right)(\xi) dt_{22}dt_{12}.$$

и получим функцию плотности вероятности первой компоненты вектора решений x_1.

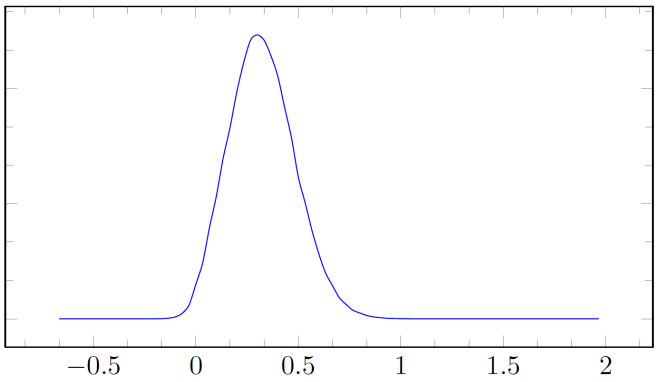


Рис. 4: Функция плотности вероятности x_1

На рис. 4 показана функция плотности вероятности первой компоненты вектора решений x_1. Плотность вероятности случайной величины x_1 имеет носитель $[-2/3, 2]$. Тем не менее, вне интервала $[-0.1, 1]$ плотность вероятности пренебрежимо мала.

3 Совместное использование метода Монте-Карло и вычислительного вероятностного анализа

Как известно, при переходе к сложным многомерным задачам эффективность метода Монте-Карло [7] возрастает. Принято считать, что для размерностей $D \leq 3$ предпочтительней использовать квадратурные формулы. Для размерностей больше 10 метод Монте-Карло не имеет конкурентов.

Заметим, что размерность D задачи (2) $D = (n+1)n$. Использование вычислительного вероятностного анализа понижает размерность на $2n$ до $(n-1)n$.

В тех случаях, когда численное вычисление интегралов (3),(4) затруднительно, воспользуемся методом Монте-Карло совместно с вычислительным вероятностным анализом. Для этих целей, например, при вычислении $\boldsymbol{x}_1(\xi, \boldsymbol{A}, \boldsymbol{b})$ (3) реализуем N выборочных значений $\zeta_i = (t^i_{12}, \ldots, t^i_{nn})$, $i = 1, \ldots, N$ случайного вектора $\zeta = (a_{12}, \ldots, a_{nn})$ согласно плотностей $\boldsymbol{a}_{12}, \ldots, \boldsymbol{a}_{nn}$. Для каждого вектора ζ_i вычислим

$$I_i = t^i_{12} \cdot \ldots \cdot t^i_{nn} \frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})}(\xi), \quad i = 1, \ldots, N. \tag{6}$$

Следовательно, согласно методу Монте-Карло [7]

$$\boldsymbol{x}_1(\xi, \boldsymbol{A}, \boldsymbol{b}) \approx \bar{I} = \frac{1}{N} \sum_i^N I_i.$$

Таким образом, совместное использование Монте-Карло и ВВА позволит уменьшить время расчетов и поднять точность.

Заключение

Предложенный подход позволяет решить задачу вычисления функции плотности вероятности в процессах моделирования со случайными входными данными. Для этих целей мы предлагаем использовать параллельно-рекурсивную организацию вычислительного процесса. Таким образом, важная проблема вычисления вероятностных расширений может быть решена в рамках параллельного рекурсивного программирования. Это

открывает множество возможностей для изучения различных моделей со случайными входными данными. Быстрые и точные вычисления основаны на свойствах числовых арифметических процедур над кусочно-полиномиальными моделями, разработанными в рамках вычислительного вероятностного анализа.

Список литературы

[1] Dobronets B.S., Popova O.A. Numerical probabilistic analysis under aleatory and epistemic uncertainty // Reliable Computing.2014. V. 19. P. 274–289.

[2] Dobronets B., Popova O. Numerical Probabilistic Approach for Optimization Problems // Scientific Computing, Computer Arithmetic, and Validated Numerics. Lecture Notes in Computer Science 9553. 2016. Springer International Publishing, Cham, P. 43–53.

[3] Dobronets B.S., Popova O.A. Improving the accuracy of the probability density function estimation // Journal of Siberian Federal University, Mathematics & Physics. 2017. V. 10. iss. 1. P. 16–21.

[4] Dobronets B.S., Popova O.A. Piecewise Polynomial Aggregation as Preprocessing for Data Numerical Modeling // IOP Conf. Series: Journal of Physics: Conf. Series 1015. 2018. 032028 doi :10.1088/1742-6596/1015/3/032028

[5] Dobronets B.S., Popova O.A. Improving reliability of aggregation, numerical simulation and analysis of complex systems by empirical data // IOP Conf. Series: Materials Science and Engineering. 354. 2018. 012006 doi:10.1088/1757-899X/354/1/012006

[6] Dobronets B.S., Popova O.A. The numerical probabilistic approach to the processing and presentation of remote monitoring data // Journal of Siberian Federal University. Engineering & Technologies. 2016 V. 9, iss. 7. P. 960–971.

[7] Михайлов, Г. А. Статистическое моделирование. Методы Монте-Карло : учеб. пособие для бакалавриата и магистратуры / Г. А. Михайлов, А. В. Войтишек. — М. : Издательство Юрайт, 2018. – 371 с. — (Серия : Бакалавр и магистр. Академический курс).

[8] Dobronets B.S., Krantsevich A.M., Krantsevich N.M. Software Implementation of Numerical Operations on Random Variables // Journal of Siberian Federal University, Mathematics & Physics. 2013. V. 6. iss. 2. P. 168–173.

[9] Dobronets B.S., Popova O.A. Computational Aspects of Probabilistic Extensions // Tomsk State University Journal of Control and Computer Science. 2019. (to appear)

Добронец Борис Станиславович — д.ф.-м.н., профессор
Сибирского федерального университета, Институт космических
и информационных технологий, кафедра систем искусственного интеллекта;
e-mail: BDobronets@yandex.ru;
Попова Ольга Аркадьевна — к.т.н., доцент
Сибирского федерального университета, Институт космических
и информационных технологий, кафедра систем искусственного интеллекта;
e-mail: OlgaArc@yandex.ru;
Дата поступления — 30 апреля 2019 г.

ИСПОЛЬЗОВАНИЕ КОНТРОЛЛЕРА LEAP MOTION ДЛЯ ПРИКЛАДНЫХ СИСТЕМ РАСПОЗНАВАНИЯ ЯЗЫКА ЖЕСТОВ

Д. Г. Еникеев[1], С. А. Мустафина [1]

[1] *Стерлитамакский филиал Башкирского государственного университета, 453103, Стерлитамак*

УДК 004.93'1
DOI: 10.24411/9999-016A-2019-10023

Слабослышащие люди и люди с полной потерей слуха для передачи информации используют язык жестов, основанный на самостоятельной форме изложения, не связанной со словесной. Здоровые люди находят такой способ общения затруднительным и неизбежно возникает потребность в переводчике. Современное развитие технологий распознавания и обучения видеообразов позволили создавать системы распознавания жестов. Они призваны обеспечить общество глухих и слабослышащих людей более удобным способом общения. Задачей устройства является перевод языка жестов в письменную либо аудио-речь. В представленной работе нами приведены основные свойства языка жестов, его параметры и отличия от других способов коммуникации. Также проведен обзор наиболее популярных аппаратных способов сбора данных о жестах и их обработке. В работе предлагается использование контроллера Leap motion [1] от одноименной компании в системе распознавания жестов, так как такой контроллер позволяет избавиться от проблем с процессом распознавания рук пользователя и выделением основных характеристик.

Ключевые слова: язык жестов, система распознавания, классификация, методы сбора данных, компьютерное зрение.

Введение

Люди общаются друг с другом, чтобы обмениваться своими идеями, мыслями и опытом. Но, к сожалению, это не касается глухонемых. Они не могут создать разговор со здоровыми людьми и изолируются от общества. Язык жестов помогает людям с ограниченными возможностями беседовать между собой. Данный способ общения состоит из различных жестов, образованных позами рук, движениями и ориентацией рук, тела, или выражениями лица. Использовать язык возможно только для тех, кто прошел специальную подготовку. Поэтому разработанная система распознавания должна играть роль переводчика языка жестов для разрушения барьеров общения с глухими людьми.

Существует много мифов о языках жестов, например, что язык жестов универсален, имеет одинаковую грамматическую структуру, и люди разных стран понимают с помощью языка жестов друг друга. Но реальные факты таковы: язык жестов для большинства стран является уникальным; диалекты жестового языка каждой страны отличаются от региона к региону; языки жестов не будут полностью поняты иностранцами, пересечения в словах минимальны; в жестовом языке также много слов, как и в любом другом речевом.

1 Этапы распознавания жестов

Разработка технической системы распознавания языка жестов состоит из следующих этапов:

Первый этап. Создание блока сбора информации. Это устройство, которое отвечает за захват жеста. В случае подходов, основанных на видео, это камера, а в случае устройств на основе датчиков — это перчатка данных или трекер движения.

ISBN 978-5-901548-42-4

Второй этап. Создание блока предварительной обработки жестов. Это блок, отвечающий за то, чтобы сделать информацию о жестах полезной с точки зрения извлечения признаков. Такой блок расширит полезные данные и избавит от ненужных шумов.

Третий этап. Создание блока извлечения характеристик. Это блок, который собирает все компоненты жеста и сохраняет их в кодовом векторе.

Четвертый этап. Сопоставление с шаблоном. В этом блоке кодовый вектор сравнивается с существующими векторами в базе данных.

Пятый этап. Классификация жестов. На основе результатов сопоставления с шаблоном будет проведена классификация жестов. Этот блок классифицирует жест в соответствии с ближайшим совпадением, найденным в сопоставлении с шаблоном.

Шестой этап. Блок распознавания жестов. Этот конечный блок полностью распознает жест и выдаст соответствующий вывод.

2 Способы сбора информации в системе распознавания жестов

Подходы и методы распознавания языка жестов появились за десять лет после распознавания аудио-речи. Связано это с историческим зазором между передачей, обработкой аудио- и видеоинформации. Первые научные публикации в области распознавания языка жестов появились в начале 90-х годов. Большинство приложений, представленных в работах, не работали в режиме реального времени и требовали до 20 секунд задержки для завершения обработки и классификации жеста. Существует небольшое количество опубликованных работ, в которых приводятся подробности об оборудовании и разрешении камеры. Как правило, большинство предлагаемых подходов предполагают использование профессионального оборудования, оптимальное размещение камеры, низкий уровень шума и высокое разрешение. Магнитные или оптические маркеры на руках и лице облегчают определение ручной конфигурации и выражения лица.

Исторически первым аппаратным способом сбора данных являются «умные перчатки». Устройства содержали в себе датчики в перчатке, регистрирующие угол наклона фаланг относительно ладони, поворот руки к горизонту. В дальнейшем, развитие технологий микроэлектроники позволило добавить акселерометры для измерения картины ускорений во время выполнения жеста. Так, например, в работе [2] используют сенсоры изгиба и акселерометры на каждой руке, которые используются для определения их наклона в пространстве. В работе [3] использует 5DTDatagloves — перчатки с 5 магнитными датчиками, отслеживающими одиннадцати свойств для одной руки: позиции в трехмерном пространстве, поворот ладони, изгиб пальцев. Преимуществами этого метода сбора данных является большой объем регистрируемой информации и его точность. Такой способ сбора имеет большую точность, однако ношение дополнительного оборудования является неприемлемым в использовании в повседневной жизни.

Постепенно с развитием видеоаппаратуры сформировались стандартные системы распознавания, основанные на использовании камер видимого диапазона. Варианты типов камер и их расположения могут быть различны. В [4] исследователи используют веб-камеру ноутбука и MultiSim для записи и обработки данных. В работе [5] использует камеру с частотой 60 кадров в секунду и программный фильтр кожи. Благодаря использованию камер значительный опыт в области распознавания объектов может быть перенесен в область языка жестов. Второй положительной стороной является обширная база данных жестов. Однако решение имеет существенные недостатки, связанные со способом сбора информации о жесте. На качество съемки влияет свойства камеры, такие как разрешение и светочувствительность, формат выходного файла. Прямое влияние точности распознавания от камеры увеличивает стоимость и комплексность системы. Кроме того, влияние оказывают внешние условия съемки: освещение, угол и расстояние между объектом и камерой, что непосредственно отражается на сценариях использования. С точки зрения процесса распознавания алгоритмы ограничены двухмерным пространством.

В качестве методов сбора данных также используются специализированное оборудование, разработанное для определения жестов. Компания Microsoft в 2010 году выпустила устройство Kinect. Оно оснащено RGB-камерой, датчиком глубины и стерео микрофоном и способно отслеживать движения тела пользователя. Использование Microsoft Kinect дает лучшие результаты [6], чем традиционные одиночные камеры, имеющие недостатки в условиях окружающей среды. Тем не менее, оно не поддерживает распознавание формы руки, и поскольку язык жестов обычно включает различные формы руки, схожие движения нельзя различить.

3 Моделирование жестов

Моделирование относится к использованию моделей - физического, математического или иного логического представления системы, объекта, явления или процесса в качестве основы для моделирования. Относительно систем распознавания жестов, существует два типа моделирования.

2D моделирование. В случае 2D моделирования данные двухмерного изображения захватываются с одной камеры. Алгоритмы сегментации и манипулирования изображениями используются для извлечения информации. Она является единственным источником для классификации жеста. Это привело к разработке цветных маркеров или цветных перчаток на руках, которые непосредственно записывают жест и игнорируют фон. Методы 2D-моделирования основаны на алгоритмах компьютерного зрения для извлечения информации о жесте, а не на специальном оборудовании.

3D моделирование влечет за собой захват жеста или знака в трехмерном пространстве. Такой тип нивелирует неоднозначность в жестах при помощи дополнительной координаты и показывает повышенную точность в распознавании. Однако повышение количества получаемой информации влечет за собой повышенную нагрузку на процессоры обработки и выделения информации.

4 Предлагаемое решение

В данной работе нами предлагается использовать контроллер Leap Motion [1]. Устройство использует три инфракрасных светодиода для подсветки зоны распознавания и две камеры. Рабочая зона представляет собой перевернутую пирамиду объемом около 8 кубических футови дальностью распознавания около 60 см. Устройство интерпретирует пространство как трехмерную прямоугольную систему координат с центром между двух инфракрасных камер. Leap Motion отслеживает руки, пальцы, фаланги и объекты-указатели, такие как карандаш или письменная ручка. Встроенное в драйвер устройства программное обеспечение отвечает за задачи распознавания и последующего препроцессинга. В результате распознавания API выделяет характеристики для руки, пальцев, костей и жестов. В работе [7] были проведены эксперименты по измерению точности определения координат руки в пространстве. В статике устройство может обеспечивать среднюю точность в 0,7 мм, повторяемость результатов крайне велика — девиация между сериями экспериментов составила 0,17 мм. При динамическом перемещении, возможно, достичь точности менее чем 2,5 мм (среднее значение - 1,2 мм).

Заключение

В работе было рассмотрены основные свойства жестового языка, был проведен обзор способов сбора данных для систем распознавания жестов. Предложенное решение Leap Motion является перспективной аппаратной платформой для создания системы распознавания языка жестов. Наличие в устройстве инфракрасных камер для сбора видеосигнала нивелирует влияние освещенности и позволяет анализировать трехмерную картину жеста. Процедура препроцессинга и выделения характеристик, выполняемые встроенным программным обеспечением, упрощает создание приложений.

Список литературы

[1] The Leap motion company // URL: www.leapmotion.com

[2] Lokhande P. Data Gloves for Sign Language Recognition System // International Journal of Computer Applications. 2015. P. 11–14.

[3] Thang P.Q., Dung N.D., Thuy N.T. A comparison of SimpSVM and RVM for Sign Language Recognition. // the International Conference on Machine Learning and Soft Computing - ICMLSC '17. 2017. P. 98–104.

[4] Raheja J.L., Mishra A., Chaudary A. Indian Sign Language Recognition Using SVM //Pattern Recognition and Image Analysis archive. 2016. V. 26, iss. 2. 2016. P. 434–441.

[5] SubhaRajam P., Balakrishnan G. Sign Language Recognition System For Deaf And Dumb People // International Journal of Engineering Research and Technology (IJERT). 2013. P. 382–387.

[6] Huang J., Zhou W., Li H., Li W. Sign language recognition using 3D convolutional neural networks // IEEE International Conference; 2015. P. 1–6.

[7] Weichert F., Bachmann D., Bartholom?us R., Fisseler D. Analysis of the Accuracy and Robustness of the Leap Motion Controller //Sensors (Basel). 2013. P. 6380–6393.

Еникеев Данияр Гадиевич — аспирант
Стерлитамакского филиала Башкирского государственного университета;;
e-mail: cooldeepband@gmail.com.
Мустафина Светлана Анатольевна — д.ф.-м.н.,
декан факультета математики и информационных технологий
зав. кафедрой математического моделирования, профессор;
Стерлитамакский филиал Башкирского государственного университета;
e-mail: mustafina_sa@mail.ru.
Дата поступления — 1 сентября 2018 г.

АДАПТИВНЫЕ ФОРМУЛЫ ЧИСЛЕННОГО ДИФФЕРЕНЦИРОВАНИЯ ПРИ НАЛИЧИИ ПОГРАНИЧНОГО СЛОЯ

А. И. Задорин[1], В. П. Ильин[2]

[1] *Институт математики имени С.Л. Соболева СО РАН, 630090, Новосибирск*

[2] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

УДК 519.653
DOI: 10.24411/9999-016A-2019-10024

Исследован вопрос численного дифференцирования функций одной переменной с большими градиентами в пограничном слое. Строятся формулы численного дифференцирования, погрешность которых равномерна по большим градиентам функции в пограничном слое. Построение таких формул основано на выделении с точностью до множителя сингулярной составляющей, отвечающей за большие градиенты функции, и построении разностной формулы для производной, точной на сингулярной составляющей. Сингулярная составляющая рассматривается как функция общего вида. В частности, такая составляющая может соответствовать экспоненциальному пограничному слою и наличию логарифмической особенности. Получены оценки погрешности построенных формул для вычисления первой и второй производных по значениям функции в узлах сетки. Приведены результаты вычислений.

Ключевые слова: функция с большими градиентами, формулы численного дифференцирования, оценка погрешности.

Введение

Классические полиномиальные формулы численного дифференцирования неприемлемы в случае функций с большими градиентами в пограничном слое [1]. Большие градиенты в области пограничного слоя имеют решения сингулярно возмущенных задач [2], моделирующих различные конвективно-диффузионные процессы с преобладающей конвекцией. Задача построения формул численного дифференцирования для таких функций является актуальной.

Пусть функция $u(x)$ имеет представление

$$u(x) = p(x) + \gamma\Phi(x),\ x \in [a, b], \tag{1}$$

где $p(x)$ — регулярная составляющая с ограниченными производными до некоторого порядка, $\Phi(x)$ — известная сингулярная составляющая, имеющая большие градиенты в области пограничного или внутреннего слоя. Функция $p(x)$ и постоянная γ не заданы.

В частности, декомпозиция (1) справедлива для решения сингулярно возмущенной краевой задачи [2], при этом $\Phi(x) = e^{-mx/\varepsilon}$, где $m > 0$, $\varepsilon \in (0, 1]$, $x \in [0, 1]$.

В данной работе исследуются формулы численного дифференцирования, точные на сингулярной составляющей $\Phi(x)$. В случае $\Phi(x) = e^{-mx/\varepsilon}$ такие формулы исследовались в [3].

Под C и C_j будем подразумевать положительные постоянные, не зависящие от сингулярной составляющей $\Phi(x)$, ее производных и от шага сетки. В случаях экспоненциального пограничного слоя и логарифмической особенности эти постоянные не зависят от малого параметра ε и шага сетки. Различные величины можем ограничивать одной постоянной C.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 16-29-15122) и программы 1.1.3 фундаментальных исследований СО РАН (проект 0314-2019-0009).

ISBN 978-5-901548-42-4

1 Формула подгонки к сингулярной составляющей

Пусть функция $u(x)$ представима в виде (1) и задана в узлах равномерной сетки. Применение к функции с большими градиентами классических интерполяционных формул, формул численного дифференцирования и интегрирования может приводить к существенным погрешностям [1, 4].

Рассмотрим формулу подгонки к сингулярной составляющей $\Phi(x)$, предложенную в [5]. Обозначим за $I(u,x)$ оператор интерполяции, численного дифференцирования или интегрирования. Пусть $I^h(u,x)$ — классическая формула для приближения $I(u,x)$, основанная на интерполяции Лагранжа. Тогда

$$|I(u,x) - I^h(u.x)| \le |I(p,x) - I^h(p,x)| + |\gamma||I(\Phi,x) - I^h(\Phi,x)|. \tag{2}$$

Зададим адаптивную формулу, точную на составляющей $\Phi(x)$:

$$I_\Phi(u,x) = \frac{I^h(u,x)}{I^h(\Phi,x)} I(\Phi,x). \tag{3}$$

Предполагается, что $I^h(\Phi,x) \neq 0$. Этого ограничения можно добиться за счет выбора регулярной составляющей $p(x)$. Формула (3) применялась в [5] в задаче численного интегрирования.

Несложно получить оценку

$$\Big|I_\Phi(u,x) - I(u,x)\Big| \le \Big|I^h(p,x) - I(p,x)\Big| + |I^h(p,x)|\Big|\frac{I(\Phi,x) - I^h(\Phi,x)}{I^h(\Phi,x)}\Big|. \tag{4}$$

Из (2) и (4) следует, что формула подгонки $I_\Phi(u,x)$ является более точной, если $|I^h(p,x)| \ll |I^h(\Phi,x)|$.

2 Адаптивные формулы численного дифференцирования

2.1 Формула с двумя узлами для вычисления первой производной

Пусть Ω^h - равномерная сетка интервала $[a,b]$:

$$\Omega^h = \{x_n :\ x_n = a + nh, n = 0,1,\dots,N,\ x_N = b\},\ \Delta_n = [x_{n-1}, x_n].$$

Предполагаем, что функция $u(x)$ вида (1) задана в узлах сетки Ω^h и $u_n = u(x_n)$.

Покажем неэффективность применения классической формулы для производной в случае функции вида (1). Зададим $\Phi(x) = \ln x, x \in [\varepsilon, 1]$. При $\varepsilon = h\ \ \varepsilon|(u_1 - u_0)/h - u'(\varepsilon)| = 1 - \ln 2$. Итак, при $h \to 0$ относительная погрешность классической формулы для первой производной не уменьшается, если $\varepsilon = h$. Задача построения специальных формул численного дифференцирования для функций вида (1) актуальна.

Получим адаптивную формулу численного дифференцирования на основе формулы (3). Зададим формулу для производной: $I^h(u,x) = (u_n - u_{n-1})/h,\ x \in [x_{n-1}, x_n]$. Предполагая, что $\Phi'(x) \neq 0$, из (3) получим формулу, точную на сингулярной составляющей $\gamma\Phi(x)$:

$$u'(x) \approx L'_{\Phi,2}(u,x) = \frac{u_n - u_{n-1}}{\Phi_n - \Phi_{n-1}} \Phi'(x),\ x \in [x_{n-1}, x_n]. \tag{5}$$

Пусть

$$|\Phi'(x)| \le B_n,\ x \in [x_{n-1}, x_n].$$

Лемма 1. *Пусть для некоторой постоянной G_n справедлива оценка:*

$$\frac{\int\limits_{x_{n-1}}^{x_n} |\Phi''(s)|\, ds}{B_n|\Phi_n - \Phi_{n-1}|} \le G_n. \tag{6}$$

Тогда

$$\Big|\frac{L'_{\Phi,2}(u,x) - u'(x)}{B_n}\Big| \le G_n \int\limits_{x_{n-1}}^{x_n} |p'(s)|\, ds + \frac{1}{B_n} \int\limits_{x_{n-1}}^{x_n} |p''(s)|\, ds,\ x \in [x_{n-1}, x_n]. \tag{7}$$

Доказательство. Учитывая, что формула (5) является точной на $\Phi(x)$, получаем

$$L'_{\Phi,2}(u,x)-u'(x)=\frac{p_n-p_{n-1}}{\Phi_n-\Phi_{n-1}}\Big[\Phi'(x)-\frac{\Phi_n-\Phi_{n-1}}{h}\Big]+\Big(\frac{p_n-p_{n-1}}{h}-p'(x)\Big). \tag{8}$$

Учитывая оценку

$$\Big|\Phi'(x)-\frac{\Phi_n-\Phi_{n-1}}{h}\Big|\le\int\limits_{x_{n-1}}^{x_n}|\Phi''(s)|\,ds,$$

из (8) получаем (7). Лемма доказана.

Замечание 1. Остановимся на случае экспоненциального погранслоя, когда $\Phi(x)=e^{-mx/\varepsilon},\ \varepsilon,m>0,\ x\in[0,1]$. Зададим $B_n=m/\varepsilon$, тогда в (6) $G_n=1$. Из (7) получаем

$$\varepsilon\Big|L'_{\Phi,2}(u,x)-u'(x)\Big|\le\int\limits_{x_{n-1}}^{x_n}\Big[m|p'(s)|+\varepsilon|p''(s)|\Big]\,ds,\ x\in[x_{n-1},x_n]. \tag{9}$$

В случае сингулярно возмущенной задачи $p'(x),\varepsilon p''(x)$ — ограничены, поэтому получаем оценку относительной погрешности порядка $O(h)$ равномерно по параметру ε.

Замечание 2. Остановимся на случае $\Phi(x)=\ln x,\ x\in[\varepsilon,1]$. Зададим $B_n=1/x_{n-1}$ и получим

$$\frac{\int\limits_{x_{n-1}}^{x_n}|\Phi''(s)|\,ds}{B_n|\Phi_n-\Phi_{n-1}|}=\frac{1-x_{n-1}/x_n}{\ln\frac{x_n}{x_{n-1}}}\le 1.$$

Получаем, что условие (6) выполнено при задании $G_n=1$. Итак, в этом случае получаем оценку погрешности (9) при задании $m=1$.

2.2 Разностная формула для второй производной

Классическая разностная формула для вычисления второй производной по значениям функции в трех узлах сетки имеет вид:

$$u''(x)\approx\frac{u_{n+1}-2u_n+u_{n-1}}{h^2},\ x\in[x_{n-1},x_{n+1}]. \tag{10}$$

Можно показать, что в случае функции вида (1) относительная погрешность формулы (10) может быть порядка $O(1)$. На основе формулы (3) получим разностную формулу, точную на сингулярнгой составляющей $\gamma\Phi(x)$:

$$u''(x)\approx L''_{\Phi,3}(u,x)=\frac{u_{n+1}-2u_n+u_{n-1}}{\Phi_{n+1}-2\Phi_n+\Phi_{n-1}}\Phi''(x),\ x\in[x_{n-1},x_{n+1}]. \tag{11}$$

Предполагаем, что

$$\Phi''(x)\ne 0, x\in(x_{n-1},x_{n+1}),\ |\Phi''(x)|\le D_n,\ x\in[x_{n-1},x_{n+1}]. \tag{12}$$

Лемма 2. *Пусть для некоторой постоянной G_n справедлива оценка:*

$$\frac{h}{D_n}\frac{\int\limits_{x_{n-1}}^{x_{n+1}}|\Phi'''(s)|\,ds}{|\Phi_{n+1}-2\Phi_n+\Phi_{n-1}|}\le G_n. \tag{13}$$

Тогда

$$\Big|\frac{L''_{\Phi,3}(u,x)-u''(x)}{D_n}\Big|\le\frac{3G_n}{2}\int\limits_{x_{n-1}}^{x_{n+1}}|p''(s)|\,ds+\frac{3}{2D_n}\int\limits_{x_{n-1}}^{x_{n+1}}|p'''(s)|\,ds,\ x\in[x_{n-1},x_{n+1}]. \tag{14}$$

Доказательство. Учитывая декомпозицию (1) и формулу (11), получаем:

$$L''_{\Phi,3}(u,x) - u''(x) = L''_{\Phi,3}(p,x) - p''(x) =$$

$$= (p_{n+1} - 2p_n + p_{n-1})\Big[\frac{1}{\Phi_{n+1} - 2\Phi_n + \Phi_{n-1}}\Phi''(x) - \frac{1}{h^2}\Big] + \Big(\frac{p_{n+1} - 2p_n + p_{n-1}}{h^2} - p''(x)\Big). \tag{15}$$

Запишем первое слагаемое формулы (15) в виде

$$Q_1 = (p_{n+1} - 2p_n + p_{n-1})\frac{\Phi''(x) - (\Phi_{n+1} - 2\Phi_n + \Phi_{n-1})/h^2}{\Phi_{n+1} - 2\Phi_n + \Phi_{n-1}}. \tag{16}$$

Учитывая (13), из (16) получаем

$$\Big|\frac{Q_1}{D_n}\Big| \le \frac{3G_n}{2}\int\limits_{x_{n-1}}^{x_{n+1}} |p''(s)|\,ds. \tag{17}$$

Для второго слагаемого в (15) имеем

$$\Big|\frac{p_{n+1} - 2p_n + p_{n-1}}{h^2} - p''(x)\Big| \le \frac{3}{2}\int\limits_{x_{n-1}}^{x_{n+1}} |p'''(s)|\,ds. \tag{18}$$

Из (15), (17), (18) получаем (14). Лемма доказана.

Замечание 3. По аналогиями с замечаниями 1, 2 на основании оценки (14) можно показать, что в случаях $\Phi(x) = e^{-mx/\varepsilon}$ и $\Phi(x) = \ln x$ для некоторой постоянной C справедлива оценка погрешности:

$$\varepsilon^2\Big|L''_{\Phi,3}(u,x) - u''(x)\Big| \le Ch.$$

2.3 Формула с тремя узлами для первой производной

Выпишем полиномиальную формулу с тремя узлами для вычисления производной:

$$L'_3(u,x) = \frac{u_{n+1} - u_{n-1}}{2h} + \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2}(x - x_n). \tag{19}$$

Несложно показать, что относительная погрешность формулы (19) порядка $O(1)$ в случае функции с погранслойной составляющей $\Phi(x) = e^{-x/\varepsilon}$, при $\varepsilon \le h$. Итак, применение формулы (19) может приводить к существенным погрешностям при наличии пограничного слоя.

Выпишем формулу с тремя узлами для производной, точную на функции $\Phi(x)$:

$$L'_{\Phi,3}(u,x) = \frac{u_{n+1} - u_{n-1}}{2h} + \frac{u_{n+1} - 2u_n + u_{n-1}}{\Phi_{n+1} - 2\Phi_n + \Phi_{n-1}}\Big[\Phi'(x) - \frac{\Phi_{n+1} - \Phi_{n-1}}{2h}\Big],\ x \in [x_{n-1}, x_{n+1}]. \tag{20}$$

Лемма 3. *Пусть*

$$\Phi''(x) \ne 0,\ |\Phi'''(x)| \le q|\Phi''(x)|,\ x \in (x_{n-1}, x_{n+1}). \tag{21}$$

Тогда

$$\Big|L'_{\Phi,3}(u,x) - u'(x)\Big| \le \frac{7}{4}h\int\limits_{x_{n-1}}^{x_{n+1}} \Big[q|p''(s)| + |p'''(s)|\Big]\,ds. \tag{22}$$

Доказательство. Представим погрешность в виде

$$L'_{\Phi,3}(u,x) - u'(x) = L'_{\Phi,3}(p,x) - p'(x) = \Big(L'_{\Phi,3}(p,x) - L'_3(p,x)\Big) + \Big(L'_3(p,x) - p'(x)\Big) \tag{23}$$

и оценим каждое слагаемое.

Оценим первое слагаемое в (23). Несложно получить:

$$L'_{\Phi,3}(p,x) - L'_3(p,x) = \frac{p_{n+1} - 2p_n + p_{n-1}}{\Phi_{n+1} - 2\Phi_n + \Phi_{n-1}}F_n, \tag{24}$$

где

$$F_n = \Big(\Phi'_n - \frac{\Phi_{n+1} - \Phi_{n-1}}{2h}\Big) + (x - x_n)\Big(\Phi''_n - \frac{\Phi_{n+1} - 2\Phi_n + \Phi_{n-1}}{h^2}\Big) + \int\limits_{x_n}^{x} (x - s)\Phi'''(s)ds. \tag{25}$$

Учитывая разложения в ряд Тейлора с остаточным членом в интегральной форме, получаем:

$$|F_n| \le \frac{7}{4} \int\limits_{x_n}^{x_{n+1}} (x_{n+1} - s)|\Phi'''(s)|\, ds + \frac{7}{4} \int\limits_{x_{n-1}}^{x_n} (s - x_{n-1})|\Phi'''(s)|\, ds, \tag{26}$$

$$\Phi_{n+1} - 2\Phi_n + \Phi_{n-1} = \int\limits_{x_n}^{x_{n+1}} (x_{n+1} - s)\Phi''(s)\, ds + \int\limits_{x_{n-1}}^{x_n} (s - x_{n-1})\Phi''(s)\, ds. \tag{27}$$

Учитывая (26), (27) и условие (21), из (24) получаем

$$\Big|L'_{\Phi,3}(p, x) - L'_3(p, x)\Big| \le \frac{7}{4} qh \int\limits_{x_{n-1}}^{x_{n+1}} |p''(s)|\, ds. \tag{28}$$

Для второго слагаемого в (23) несложно получить оценку:

$$\Big|L'_3(p, x) - p'(x)\Big| \le \frac{7}{4} h \int\limits_{x_{n-1}}^{x_{n+1}} |p'''(s)|\, ds. \tag{29}$$

Учитывая оценки (28), (29), получаем (22), что доказывает лемму.

Из леммы 3 следует, что в случаях $\Phi(x) = e^{-mx/\varepsilon}$ и $\Phi(x) = \ln x,\ x \ge \varepsilon > 0$ справедлива оценка погрешности:

$$\varepsilon\Big|L'_{\Phi,3}(u, x) - u'(x)\Big| \le Ch \int\limits_{x_{n-1}}^{x_{n+1}} \Big[|p''(s)| + \varepsilon|p'''(s)|\Big]\, ds. \tag{30}$$

3 Применение формул в двумерном случае

Пусть для достаточно гладкой функции $u(x, y)$ в области $\Omega = [0, 1] \times [0, 1]$ справедливо представление:

$$u(x, y) = p(x, y) + \gamma(y)\Phi(x), \tag{31}$$

где $p(x, y)$ — регулярная составляющая с ограниченными производными до некоторого порядка, $\Phi(x)$ — известная функция с большими градиентами. Функции $p(x, y), \gamma(y)$ в декомпозиции (31) не заданы. Такая декомпозиция справедлива для решения эллиптической задачи с пограничным слоем по переменной x [6].

Пусть Ω^h — прямоугольная сетка исходной области Ω:

$$\Omega^h = \{(x_i, y_j),\ x_i = ih_1,\ i = 0, 1, \ldots, N_1,\ x_j = jh_2,\ j = 0, 1, \ldots, N_2.$$

Зададим ячейку $K_{i,j}$: $K_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$. Предполагаем, что функция $u(x, y)$ вида (31) задана в узлах равномерной сетки Ω^h и построим формулу численного дифференцирования в ячейке $K_{i,j}$.

На основе дифференцирования двумерного интерполянта, точного на $\Phi(x)$, строим формулу для вычисления производной, которая является точной на $\Phi(x)$:

$$\frac{\partial u}{\partial x}(x, y) \approx L'_{\Phi.x}(u, x, y) = \Big(u_{i+1,j+1} - u_{i,j+1} - u_{i+1,j} + u_{i,j}\Big)\frac{\Phi'(x)}{\Phi_{i+1} - \Phi_i} \times \frac{y - y_j}{h_2} +$$

$$+\Big(u_{i+1,j} - u_{i,j}\Big)\frac{\Phi'(x)}{\Phi_{i+1} - \Phi_i},\ (x, y) \in K_{i,j}. \tag{32}$$

Лемма 4. *Пусть* $|\Phi'(x)| \le B_i, x \in [x_i, x_{i+1}]$ *и для некоторой постоянной* G_i *справедлива оценка:*

$$\frac{\int\limits_{x_i}^{x_{i+1}} |\Phi''(s)|\,ds}{B_i|\Phi_{i+1} - \Phi_i|} \le G_i.$$

Тогда

$$\frac{1}{B_i}|L'_{\Phi,x}(u,x,y) - u'_x(x,y)| \le 3h_1 \max_{(s,t)} |p'_x(s,t)|G_i + \frac{2}{B_i}\Big(\max_{s,t}|p''_{xx}(s,t)|h_1 + \max_{s,t}|p''_{xy}(s,t)|h_2\Big). \tag{33}$$

Замечание 4. Пусть $\Phi(x) = e^{-mx/\varepsilon}$ или $\Phi(x) = \ln x$. Тогда в (33) $B_i = C/\varepsilon$.

Рассмотрим уравнение Пуассона в кольце:

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \phi^2} = f(r,\phi), \tag{34}$$

$$u(\varepsilon,\phi) = \Psi_1(\phi),\ u(1,\phi) = \Psi_2(\phi), \tag{35}$$

где $\varepsilon \le r \le 1,\ 0 \le \phi \le 2\pi$. Функции $f(r,\phi), \Psi_1(\phi), \Psi_2(\phi)$ — достаточно гладкие.

К задаче (34)–(35) приводится, например, задача об установившейся фильтрации жидкости к скважине, при этом ε — параметр, пропорциональный относительному диаметру скважины. Решение задачи (34)–(35) имеет большие градиенты у границы $r = \varepsilon$ и представимо в виде

$$u(r,\phi) = p(r,\phi) + \gamma(\phi)\ln r,$$

где функции $p(r,\phi), \gamma(\phi)$ имеют ограниченные первые производные по своим аргументам. Следовательно, для вычисления производной по r можно применить формулу (32). При этом для некоторой постоянной C справедлива оценка погрешности:

$$\varepsilon|L'_{\Phi,r}(u,r,\phi) - u'_r(r,\phi)| \le C(h_1 + h_2),\ (r,\phi) \in K_{i,j},\ i,j = 0,1,\ldots,N-1.$$

4 Результаты численных экспериментов

Рассмотрим функцию с погранслойной составляющей экспоненциального вида

$$u(x) = cos(\pi x) + e^{-x/\varepsilon},\quad 0 \le x \le 1,\ \varepsilon > 0.$$

Таблица 1: Погрешность полиномиальной формулы (19)

ε	N					
	10	10^2	10^3	10^4	10^5	10^6
1	$5.04e-2$	$5.07e-4$	$5.07e-6$	$5.07e-8$	$5.39e-10$	$3.27e-10$
10^{-1}	$2.06e-2$	$1.36e-3$	$1.63e-5$	$1.66e-7$	$1.67e-9$	$3.65e-11$
10^{-2}	$5.14e-4$	$2.37e-2$	$1.37e-3$	$1.63e-5$	$1.66e-7$	$1.67e-9$
10^{-3}	$5.14e-5$	$2.24e-6$	$2.37e-2$	$1.36e-3$	$1.63e-5$	$1.66e-7$
10^{-4}	$5.14e-6$	$5.17e-8$	$2.27e-6$	$2.37e-2$	$1.37e-3$	$1.63e-5$
10^{-5}	$5.14e-7$	$5.17e-9$	$5.17e-11$	$2.27e-6$	$2.37e-2$	$1.37e-3$

Рассмотрим формулы для производных с тремя узлами. В табл. 1 приведена норма погрешности

$$\Delta = \varepsilon \max_n |u'(x_n) - L'_3(u,x_n)|,\ 0 \le n \le N,$$

для полиномиальной формулы численного дифференцирования (19). Погрешность неравномерна по параметру ε.

В табл. 2 приведена норма погрешности построенной неполиномиальной формулы для производной (20). Подтверждается второй порядок точности по шагу сетки, что соответствует оценке (30).

Таблица 2: Погрешность формулы (20), точной на погранслойной составляющей

ε	N					
	10	10^2	10^3	10^4	10^5	10^6
1	$5.39e-2$	$5.42e-4$	$5.42e-6$	$5.42e-8$	$5.75e-10$	$7.35e-11$
10^{-1}	$1.66e-2$	$1.72e-4$	$1.72e-6$	$1.72e-8$	$1.74e-10$	$3.13e-11$
10^{-2}	$4.80e-3$	$1.59e-4$	$1.64e-6$	$1.65e-8$	$1.65e-10$	$3.80e-12$
10^{-3}	$4.81e-4$	$4.93e-5$	$1.60e-6$	$1.64e-8$	$1.65e-10$	$1.85e-12$
10^{-4}	$4.81e-5$	$4.93e-6$	$4.93e-7$	$1.59e-8$	$1.64e-10$	$1.66e-12$
10^{-5}	$4.81e-6$	$4.93e-7$	$4.93e-8$	$4.93e-9$	$1.59e-10$	$1.65e-12$

Заключение

Разработаны адаптивные формулы численного дифференцирования функции одной переменной с большими градиентами. По построению формулы являются точными на сингулярной составляющей, отвечающей за большие градиенты функции. Построенная формула первого порядка точности обобщена на двумерный случай. Получены оценки погрешности разработанных формул. Показано, что в случае экспоненциального пограничного слоя и логарифмической особенности эти оценки равномерны по малому параметру. Результаты вычислительных экспериментов согласуются с полученными оценками.

Список литературы

[1] Задорин А. И. Метод интерполяции для задачи с пограничным слоем // Сибирский журнал вычисл. математики. 2007. Т. 10, № 3. С. 267–275.

[2] Шишкин Г. И. Сеточные аппроксимации сингулярно возмущенных эллиптических и параболических уравнений. Екатеринбург: УрО РАН, 1992.

[3] Zadorin A., Tikhovskaya S. Formulas of numerical differentiation on a uniform mesh for functions with the exponential boundary layer // International Journal of Numerical Analysis and Modeling. 2019. V. 16, № 4. P. 590–608.

[4] Zadorin A. I. Interpolation formulas for functions with large gradients in the boundary layer and their application // Modeling and Analysis of Information Systems. 2016. V. 23, № 3. P. 377–384.

[5] Гурьева Я. Л., Ильин В. П. Модифицированный метод конечных объемов для сингулярно возмущенных краевых задач // Вычислительные технологии. Спец. вып.: Труды Международной конференции RDAMM. - 2001. 2001. Т. 6, № 2. С. 245–252.

[6] Roos H. G., Stynes M., Tobiska L. Numerical Methods for Singularly Perturbed Differential Equations, Convection-Diffusion and Flow problems. Berlin: Springer, 2008.

Задорин Александр Иванович — д.ф.-м.н., вед. науч. сотр. Института математики им. С.Л. Соболева СО РАН;
e-mail: zadorin@ofim.oscsbras.ru;
Ильин Валерий Павлович — д.ф.-м.н., гл. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: ilin@sscc.ru.
Дата поступления — 30 апреля 2019 г.

О МОДЕЛЯХ СЕЙСМОГЕНЕРИРУЮЩИХ ГЕОЛОГИЧЕСКИХ СТРУКТУР

М. В. Зарецкая[1], О. М. Бабешко[1], В. В. Лозовой[2]

[1] *Кубанский государственный университет, 350040, Краснодар*
[2] *Южный научный центр РАН, 344006, Ростов-на-Дону*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10025

Изучение причин возникновения сильных сейсмических событий на территории Краснодарского края позволило выявить факторы, оказывающие влияние на их возникновение на территории интенсивной эксплуатации месторождений углеводородного сырья. К ним относятся, в том числе, грязевулканические структуры. Сложность сформулированной задачи определяет подходы к её решению — оценка напряженно-деформированного состояния и резонансного поведения сложно структурированной, разномасштабной, разнотипной геологической среды вулканической постройки под действием внешних и внутренних факторов делает необходимым привлечение математического аппарата, основанного на топологическом подходе: теории блочных структур и метода блочного элемента. В работе предложено для моделирования вулканических построек применять блочный элемент в форме треугольной пирамиды, позволяющей приближенно описывать любые выпуклые многогранные области и области с криволинейными границами.

Ключевые слова: сейсмогенерирующая структура, грязевулканическая постройка, блочная структура, топологический подход, дифференциальный метод факторизации.

Введение

Освоение ресурсов нефти и газа акватории Азовского и Черного морей — одно из перспективных направлений увеличения минерально-сырьевой базы Краснодарского края и юга России. Геолого- геофизическое изучение глубинного строения коры Земли этих территорий началось еще с 50-х годов XX в. Проведен обширный комплекс исследований, включающий грави-, магнито- и электроразведку. Обобщение полученного материала позволило констатировать тот факт, что на шельфе можно выделить геологические и геофизические структуры, ответственные за реализацию сейсмического потенциала. В качестве первого источника сейсмической опасности можно рассматривать вулканические структуры, при этом процесс катастрофического извержения и сильные сейсмические явления могут сопровождать друг друга или проявляться независимо.

Так как Краснодарский край относится к регионам высокой сейсмической опасности, то изучение сейсмогенерирующих структур, ответственных за реализацию сейсмического потенциала региона, закономерностей развития техногенной сейсмичности и ее взаимосвязи с промышленным воздействием на верхние слои литосферы, является актуальной задачей. Кроме того, возможность активизации наведенной сейсмичности должна учитываться при разработке шельфовых месторождений углеводородов, так как техническая сложность и, как следствие, стоимость ликвидации последствий экологических катастроф, например, в морских условиях на порядок превышают сложность и стоимость аналогичных работ на суше.

1 Модель

Для получения достоверных результатов при решении задачи оценки напряженности нефтегазоносных провинций требуется построить модель геологической среды, наиболее приближенную к естественным грязевулканическим постройкам. Выделение блочных элементов в исследуемых природных объект выполняется,

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований и администрации Краснодарского края (код проекта 19-41-230002).

ISBN 978-5-901548-42-4

как правило, по результатам экспериментального исследования строения коры Земли. Для определения глубинного строения грязевулканических построек важны как результаты отдельных экспериментальных геофизических исследований, так и совместная интерпретация результатов исследований (методом микросейсмического зондирования (ММЗ), методом магнитотеллурического зондирования (МТЗ)), изотопной и гидрохимической геотермометрии, что позволит предложить рабочую гипотезу о моделировании исследуемого геологического объекта совокупностью блочных элементов.

На территории Краснодарского края грязевой вулканизм широко представлен на шельфе Азовского моря, на Таманском полуострове, также грязевые вулканы обнаружены на Керченском полуострове и на нижней части крымского континентального склона. Анализируя результаты экспериментальных геофизических исследований, можно выделить четыре типа структур грязевых вулканов.

Первый тип — конические в поперечном сечении структуры, довольно правильные, округлые в плане, диаметром 700-800 м и высотой в несколько десятков метров.

Второй тип — вулканы с четко выраженными кальдерами обрушения по системе концентрических сбросов. Они имеют разнообразную форму и диаметр до 1 км.

Третий тип — структура с правильной округлой формой, совершенно плоским сводом и сильно разжиженными продуктами извержения. Его диаметр достигает 1100 м (грязевой вулкан Двуреченского).

К четвертому типу относятся не собственно грязевые вулканы, а трещинные излияния по системе параллельных трещин, ориентированных в широтном направлении. Ширина потоков достигает 1,5 км, а длина — 4 км.

Для первых двух типов модель предполагает наличие полости, ограниченной некоторой замкнутой поверхностью произвольной конфигурации и расположенной в многослойной среде. Как правило, рассматривается цилиндрическая поверхность. В отдельных случаях форма полости может быть близка к сферической или эллипсоидальной.

В третьем случае может быть принята упругая модель, в которой верхняя часть вулканической постройки моделируется гибкой плитой на упругом основании. Это одна из простейших моделей, позволяющая учесть реальные резонансные свойства геофизической среды.

В качестве модели вулканической постройки четвертого типа допустимо принять две полуплоскости с параллельными границами, удаленные друг от друга на некоторое расстояние, на линейно деформируемом основании.

2 Численно-аналитические методы исследования

Для исследования процессов в структурированной геофизической среде грязевулканической постройки, блоки которой являются разномасштабными и разнотипными по свойствам и структуре, применяется топологический математический аппарат, включающий в себя теорию блочных структур, дифференциальный метод факторизации и метод блочного элемента [1, 2].

Метод блочного элемента разработан для исследования и решения граничных задач для систем дифференциальных уравнений любого порядка. Его достоинство состоит в универсальности и способности выявлять составляющие решений, обладающие сингулярностями различного уровня, которые сложно выявить иными методами.

При применении метода блочного элемента граничная задача для каждого блока средствами внешнего анализа погружается в топологическую структуру таким образом, чтобы она стала эквивалентной функциональному уравнению для элементов алгебраического кольца. Дальнейшие исследования граничных задач состоит в факторизации коэффициента функционального уравнения, который может быть как функцией, так и матрицей-функцией, вычислении многомерного формы-вычета Лере и построении псевдодифференциального уравнения, содержащего все типы граничных условий, допустимых граничной задачей. Объединение блочных элементов, для которых построены свои псевдодифференциальные уравнения, производится построением фактор-топологий соответствующих прямых произведений топологических пространств. Из псевдодифференциальных уравнений извлекаются интегральные уравнения, отвечающие тому или иному типу взаимодействия блоков.

Ниже рассматривается блочный элемент в форме произвольной треугольной пирамиды, который позволяет рассматривать граничные задачи в любых областях, в том числе с криволинейными границами, для дифференциальных уравнений и их систем, имеющих также переменные коэффициенты. Блочный элемент в форме произвольной пирамиды занимает область Ω с границей $\partial\Omega$. Одновременно он является ограниченным объемным элементом, имеющим наименьшее количество граней, которые будем обозначать $\partial\Omega_k$,

$k = 1, 2, 3, 4$. Последнее обстоятельство позволяет уменьшить количество псевдодифференциальных уравнений, возникающих при решении задачи. При построении блочного элемента принято расположение локальных систем координат с началом в двух вершинах, причем в одной сосредоточено три системы, оставшаяся, основная, имеет начало координат в отдельной вершине треугольной пирамиды. Две оси локальных систем координат лежат в плоскостях граней, а третья ось во всех системах направлена по внешней нормали к соответствующим граням.

Расположим основную декартову прямоугольную локальную систему координат x^1o^1, (x_1^1, x_2^1, x_3^1) на нижней грани $\partial\Omega_1$. Поместим начало системы координат x^1o^1, в одной из трех вершин грани $\partial\Omega_1$, направив орт $\mathbf{e}_3^1$ вертикально вниз, соответственно орт $\mathbf{e}_1^1$ — вдоль линии пересечения граней $\partial\Omega_1$, $\partial\Omega_2$, взяв орт $\mathbf{e}_2^1$ в виде $\mathbf{e}_2^1 = \mathbf{e}_3^1 \times \mathbf{e}_1^1$. Тогда в системе x^1o^1 уравнение плоскости $\partial\Omega_1$ принимает вид $x_3^1 = 0$.

В результате плоскости остальных граней пирамиды в координатах x^1o^1 описываются уравнениями

$$\partial\Omega_2 : \; c_2 x_2^1 + d_2 x_3^1 = 0, \quad c_2 > 0;$$

$$\partial\Omega_3 : \; b_3 x_1^1 + c_3 x_2^1 + d_3 x_3^1 = 0, \quad b_3 > 0, \quad e_3 < 0;$$

$$\partial\Omega_4 : \; b_4 x_1^1 + c_4 x_2^1 + d_4 x_3^1 = 0, \quad b_4 > 0.$$

Постоянные коэффициенты b_n, c_m, d_k определяются размерами пирамиды. Приняв вершину схождения боковых граней $\partial\Omega_2$, $\partial\Omega_3$, $\partial\Omega_4$ в качестве начал координат локальных систем x^ko^k, (x_1^k, x_2^k, x_3^k), $k = 2, 3, 4$, построим орты $\mathbf{e}_1^k$, $\mathbf{e}_2^k$, $\mathbf{e}_3^k$, $k = 2, 3, 4$, лежащие вдоль координатных осей x_1^k, x_2^k, x_3^k соответственно. Во всех указанных системах координат орты $\mathbf{e}_1^k$, $k = 2, 3, 4$, расположим вдоль исходящих из этой вершины ребер, по одному на каждое ребро. Производя при принятых построениях довольно простые преобразования, имеем

$$\mathbf{e}_1^2 = \left\{c_{11}^2, c_{12}^2, c_{13}^2\right\} = \{b_{20}, -d_2, c_2\} \frac{1}{\sqrt{c_2^2 + d_2^2 + b_{20}^2}}, \quad b_{20} = b_3^{-1}\left(c_3 d_2 - c_2 d_3 - e_3\right);$$

$$\mathbf{e}_3^2 = \left\{0,\ c_{32}^2,\ c_{33}^2\right\} = \{0,\ -c_2\ ,\ -d_2\ \} \cdot \frac{1}{\sqrt{c_2^2 + d_2^2}};$$

$$\mathbf{e}_2^2 = \mathbf{e}_3^2 \times \mathbf{e}_1^2 = \left\{c_{21}^2,\ c_{22}^2,\ c_{23}^2\right\} = \left\{c_{32}^2 c_{13}^2 - c_{12}^2 c_{33}^2,\ \ c_{11}^2 c_{33}^2,\ \ -c_1^2 c_2^2\right\};$$

$$\mathbf{e}_1^3 = \left\{c_{11}^3, c_{12}^3, c_{13}^3\right\} = \{b_{30}, d_{30}, c_{30}\} \frac{1}{\sqrt{b_{30}^2 + d_{30}^2 + c_{30}^2}};$$

$$b_{30} = -b_4^{-1}\left(\ c_4 d_{30} + d_4 c_{30}\ \right), \quad d_{30} = -\left(d_3 b_3^{-1} + b_4^{-1} d_4 + e_3 b_3^{-1} c_{30}^{-1}\right), \quad c_{30} = c_3 b_3^{-1} - c_4 b_4^{-1}$$

$$\mathbf{e}_3^3 = \left\{c_{31}^3,\ c_{32}^3,\ c_{33}^3\right\} = \{b_3,\ c_3,\ d_3\} \frac{1}{\sqrt{b_3^2 + c_3^2 + d_3^2}};$$

$$\mathbf{e}_2^3 = \mathbf{e}_3^3 \times \mathbf{e}_1^3 = \left\{c_{21}^3, c_{22}^3, c_{23}^3\right\} = \left\{c_{32}^3 c_3^3 - c_{33}^3 c_2^3,\ c_{11}^3 c_{33}^3 - c_{31}^3 c_{13}^3,\ c_{31}^3 c_{12}^3 - c_{11}^3 c_{32}^3\right\};$$

$$\mathbf{e}_1^4 = \{b_{40}, -d_2, c_2\} \cdot \frac{1}{\sqrt{b_{40}^2 + d_2^2 + c_2^2}} = \left\{c_{11}^4, c_{12}^4, c_{13}^4\right\}, \quad b_{40} = b_4^{-1}\left(c_4 d_2 - d_4 c_2\right);$$

$$\mathbf{e}_3^4 = \left\{c_{31}^4, c_{32}^4, c_{33}^4\right\} = \{b_4, c_4, d_4\} \frac{1}{\sqrt{b_4^2 + c_4^2 + d_4^2}};$$

$$\mathbf{e}_2^4 = \mathbf{e}_3^4 \times \mathbf{e}_1^4 = \left\{c_{21}^4, c_{22}^4, c_{23}^4\right\} = \left\{c_{32}^4 c_{13}^4 - c_{33}^4 c_{12}^4, \quad c_{11}^4 c_{33}^4 - c_{31}^4 c_{13}^4, \quad c_{31}^4 c_{12}^4 - c_{11}^4 c_{32}^4\right\}.$$

Формулы перехода между системами координат имеют вид:

$$x^\nu = B^{\nu\tau} x^\tau + x_0^{\nu\tau},$$

$$B^{\nu\tau} = \mathrm{A}_\nu \mathrm{A}_\tau^{-1} \quad A_k = \left\|c_{mn}^k\right\|, \quad B^{\nu\tau} = \left\|b_{mp}^{\nu\tau}\right\|, \quad b_{mk}^{\nu\tau} = c_{mp}^\nu c_{kp}^\tau, \quad \left(B^{\nu\tau}\right)^{-1} = \left(B^{\nu\tau}\right)^* = B^{\tau\nu},$$

$$x_0^{\nu\tau} = \mathrm{A}_\nu\left(x_{\tau 0}^1 - x_{\nu 0}^1\right), \quad x_0^{\nu\tau} = \left\{x_{10}^{\nu\tau}, x_{20}^{\nu\tau}, x_{30}^{\nu\tau}\right\} \quad , \quad x^\tau = \left\{x_1^\tau, x_2^\tau, x_3^\tau\right\}.$$

Здесь $x_{\tau 0}^1$ — векторы начал координат систем $x^\tau o^\tau \tau = 2, 3, 4$ в системе x^1o^1.

Далее для иллюстрации построен блочный элемент для дифференциального уравнения:

$$\mathbf{Q}(\partial x_1^1, \partial x_2^1, \partial x_3^1)\varphi = \left[A_{11}\partial^2 x_1^1 + A_{22}\partial^2 x_2^1 + A_{33}\partial^2 x_3^1 + A\right]\varphi(x_1^1,\ x_2^1,\ x_3^1) = 0.$$

Граничная задача поставлена в области Ω с некоторыми граничными условиями, например, Дирихле или Неймана.

Сформулируем граничную задачу в каждой локальной системе координат $x^\nu o^v$, $\nu = 2, 3, 4$:

$$Q\left(\partial x_1^v, \partial x_2^v, \partial x_3^v\right)\varphi \equiv \left(\mathrm{A}_{ms}^v \partial x_m^v \partial x_s^v + \mathrm{A}\right)\varphi = 0, \quad \mathrm{A}_{pp} c_{mp}^\nu c_{sp}^\nu = \mathrm{A}_{ms}^v.$$

$$\alpha^\nu = B^{\nu\tau}\alpha^\tau, \quad \alpha_m^\tau = b_{mk}^{\tau\nu}\alpha_k^\nu, \quad \alpha^\tau = \{\alpha_1^\tau, \alpha_2^\tau, \alpha_3^\tau\}.$$

Следуя алгоритму дифференциального метода факторизации [1, 2], введем двойное и тройное преобразования Фурье и запишем функциональные уравнения для пирамиды:

$$\mathrm{K}_\nu \Phi_\nu = \iint\limits_{\partial\Omega_v} \left[\mathrm{A}_{33}^v\left(\varphi_{v3} - i\alpha_3^v\varphi_v\right) - i\alpha_1^v \mathrm{A}_{13}^v\varphi_v - i\alpha_2^v \mathrm{A}_{23}^v\varphi_v\right] \exp i\left[\alpha_1^v\eta_1^v + \alpha_2^v\eta_2^v\right] d\eta_1^v d\eta_2^v +$$

$$+ \sum_{\tau=1}^{4}{}' \iint\limits_{\partial\Omega_\tau} \left[\mathrm{A}_{33}^\tau\left(\varphi_{\tau 3} - ib_{3m}^{\tau\nu}\alpha_m^v\varphi_\tau\right) - ib_{1m}^{\tau\nu}\alpha_m^v \mathrm{A}_{13}^\tau\varphi_\tau - ib_{2m}^{\tau\nu}\alpha_m^v \mathrm{A}_{23}^\tau\varphi_\tau\right] \times$$

$$\times D_{\nu\tau}\left(\alpha_1^v, \alpha_2^v, \alpha_3^v, x_1^\tau, x_2^\tau, 0\right) dx_1^\tau dx_2^\tau, \quad \nu = 1, 2, 3, 4.$$

Здесь и далее штрих у знака суммы означает, что член $\tau = \nu$ опущен.

Внешняя форма в локальных системах координат имеет вид

$$\omega^\nu = R^\nu dx_1^v \wedge dx_2^v + Q^v dx_3^v \wedge dx_1^v + \mathrm{P}^\nu dx_2^v \wedge dx_3,$$

$$Q^v = e^{i\langle\alpha^v x^v\rangle}\left[\mathrm{A}_{22}^v\left(\frac{\partial\varphi_v}{\partial x_2^v} - i\alpha_2^v\varphi_v\right) - i\alpha_1^v \mathrm{A}_{12}^v\varphi_v + \mathrm{A}_{23}^v\frac{\partial\varphi_v}{\partial x_3^v}\right],$$

$$R^v = e^{i\langle\alpha^v x^v\rangle}\left[\mathrm{A}_{33}^v\left(\frac{\partial\varphi_v}{\partial x_3^v} - i\alpha_3^v\varphi_v\right) - i\alpha_1^v \mathrm{A}_{13}^v\varphi_v - i\alpha_2^v \mathrm{A}_{23}^v\varphi_v\right],$$

$$\mathrm{P}^v = e^{i\langle\alpha^v x^v\rangle}\left[\mathrm{A}_{11}^v\left(\frac{\partial\varphi_v}{\partial x_1^v} - i\alpha_1^v\varphi_v\right) + \mathrm{A}_{13}^v\frac{\partial\varphi_v}{\partial x_3^v} + \mathrm{A}_{12}^v\frac{\partial\varphi_v}{\partial x_2^v}\right],$$

$$e^{i\langle\alpha^v x^v\rangle} = \exp i(\alpha_1^v x_1^v + \alpha_2^v x_2^v + \alpha_3^v x_3^v), \quad \nu = 1,\ 2,\ 3,\ 4.$$

Определив корни характеристического уравнения

$$\mathrm{K}\left(\alpha^v\right) \equiv -Q\left(-i\alpha^v\right) \equiv \mathrm{A}_{33}^v\left(\alpha_3^v\right)^2 + \mathrm{A}_1 \cdot \alpha_3^v + \mathrm{A}_0 - \mathrm{A} = 0,$$

$$\mathrm{A}_1 = \sum_{m,s=1}^{2}\left[\mathrm{A}_{m3}^\nu\alpha_m^v + \mathrm{A}_{3s}^\nu\alpha_s^v\right], \quad \mathrm{A}_0 = \sum_{m,s=1}^{2}\mathrm{A}_{ms}^v\alpha_m\alpha_s$$

в каждой локальной системе координат, составим псевдодифференциальные уравнения для рассматриваемого блочного элемента:

$$\mathbf{F}^{-1}(x_1^\nu, x_2^\nu)\left\{\iint\limits_{\partial\Omega_v}\left[\mathrm{A}_{33}^v\left(\varphi_{v3} - i\alpha_{3-}^v\varphi_v\right) - i\alpha_1^v \mathrm{A}_{13}^v\varphi_v - i\alpha_2^v \mathrm{A}_{23}^v\varphi_v\right] \exp i\left[\alpha_1^v\eta_1^v + \alpha_2^v\eta_2^v\right] d\eta_1^v d\eta_2^v \times\right.$$

$$\times \sum_{\tau=1}^{4}{}' \iint\limits_{\partial\Omega_\tau}\left[\mathrm{A}_{33}^\tau\left(\varphi_{\tau 3} - i(b_{3m}^{\tau\nu}\alpha_m^v)_-\varphi_\tau\right) - i(b_{2m}^{\tau\nu}\alpha_m^v)_- \mathrm{A}_{13}^\tau\varphi_\tau - i(b_{2m}^{\tau\nu}\alpha_m^v)_- \mathrm{A}_{23}^\tau\varphi_\tau\right] \times$$

$$\left.\times D_{\nu\tau}\left(\alpha_1^v, \alpha_2^v, \alpha_{3-}^v, x_1^\tau, x_2^\tau, 0\right) dx_1^\tau dx_2^\tau\right\} = 0, \quad \nu = 1, 2, 3, 4,$$

$$(b_{pm}^{\tau\nu}\alpha_m^\nu)_- = b_{p1}^{\tau\nu}\alpha_1^\nu + b_{p2}^{\tau\nu}\alpha_2^\nu + b_{p3}^{\tau\nu}\alpha_{3-}^\nu, \quad \alpha_{3\pm}^v = \frac{-\mathrm{A}_1 \pm \sqrt{\mathrm{A}_1^2 - 4\mathrm{A}_{33}^v\left(\mathrm{A}_0 - \mathrm{A}\right)}}{2\mathrm{A}_{33}^v}.$$

Коэффициент A_{33}^v не зависит от α^v, A_0, A_1 — зависят только от α_1^v, α_2^v; α_{3-}^v — корень с отрицательной мнимой частью, $\operatorname{Im}\alpha_{3-}^v \leqslant 0$.

Решение задачи, записанное в локальных системах координат, имеет вид:

$$\varphi_v(x_1^v,x_2^v,x_3^v)=\mathbf{F}^{-1}(x_1^v,x_2^v,x_3^v)\mathrm{K}_v^{-1}\Bigg\{\iint\limits_{\partial\Omega_v}\left[\mathrm{A}_{33}^v\left(\varphi_{v3}-i\alpha_3^v\varphi_v\right)-i\alpha_1^v\mathrm{A}_{13}^v\varphi_v-i\alpha_2^v\mathrm{A}_{23}^v\varphi_v\right]\times$$

$$\times\exp i\left[\alpha_1^v\eta_1^v+\alpha_2^v\eta_2^v\right]d\eta_1^v d\eta_2^v+\sum_{\tau=1}^{4}{}'\iint\limits_{\partial\Omega_\tau}\left[\mathrm{A}_{33}^\tau\left(\varphi_{\tau3}-ib_{3m}^{\tau\nu}\alpha_m^v\varphi_\tau\right)-\right.$$

$$\left.-ib_{1m}^{\tau\nu}\alpha_m^v\mathrm{A}_{13}^\tau\varphi_\tau-ib_{2m}^{\tau\nu}\alpha_m^v\mathrm{A}_{23}^\tau\varphi_\tau\right]\times D_{\nu\tau}\left(\alpha_1^v,\alpha_2^v,\alpha_3^v,x_1^\tau,x_2^\tau,0\right)dx_1^\tau dx_2^\tau\Bigg\}.$$

Заключение

Вулканическая постройка — сложный геологический объект, включающий области как с прямолинейными, так и криволинейными границами. Для их моделирования возможно введение блочных элементов частного вида, например, слоя или полупространства с вырезанным цилиндром, но и универсального блочного элемента в форме треугольной пирамиды, позволяющей приближенно описывать любые выпуклые многогранные области и области с криволинейными границами. Рассмотренный блочный элемент может быть самостоятельным упругим телом, моделирующим конус вулканической постройки. В этом случае решения псевдодифференциальных уравнений и их производные могут иметь особенности в вершинах пирамиды и на ее ребрах, как и само решение граничной задачи. Если же блочный элемент возник как результат разбиения на контактирующие блоки некоторой области, то в этом случае указанные особенности взаимно уничтожаются при контакте блоков и не присутствуют в решении, что позволяет упрощать решение псевдодифференциальных уравнений.

Список литературы

[1] Babeshko V.A., Evdokimova O.V., Babeshko O.M., Gorshkova E.M., Zaretskaya M.V., Mukhin A.C., Pavlova A.V. Convergent properties of block elements // Doklady Physics. 2015. V. 60. iss. 11. P. 515–518.

[2] Зарецкая М.В. Математические методы исследования неустойчивых геологических структур // Защита окружающей среды в нефтегазовом комплексе. 2013. № 7. С. 33–38.

[3] Бабешко В.А., Бабешко О.М., Евдокимова О.В. К теории блочного элемента // Доклады академии наук. 2009. Т. 427, № 2. С. 183–186.

Зарецкая Марина Валерьевна — д.ф.-м.н., профессор кафедры математического моделирования;
Кубанский государственный университет;
e-mail: zarmv@mail.ru;
Бабешко Ольга Мефодиевна — д.ф.-м.н., главный научный сотрудник научно-исследовательского центра прогнозирования и предупреждения геоэкологических и техногенных катастроф;
Кубанский государственный университет;
e-mail: babeshko49@mail.ru;
Лозовой Виктор Викторович — к.ф.-м.н., старший научный сотрудник лаборатории математики и механики; Южный научный центр РАН;
e-mail: niva_kgu@mail.ru.
Дата поступления — 30 апреля 2019 г.

АНАЛИЗ ВЛИЯНИЯ СЛУЧАЙНЫХ ШУМОВ НА СОЛИТОННОЕ РЕШЕНИЕ УРАВНЕНИЯ SIN-ГОРДОНА

А. А. Иванов[1], Д. Д. Смирнов[1]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10026

В работе методом численного статистического моделирования исследуется влияние случайных шумов на солитонные решения уравнения sin-Гордона. Приводятся результаты расчётов частотного временного сечения, проведённых на кластере НКС-30Т Сибирского Суперкомпьютерного центра при ИВМиМГ СО РАН

Ключевые слова: стохастические дифференциальные уравнения в частных производных, метод численного статистического моделирования, параллельные алгоритмы, уединённая волна, солитон, уравнение sin-Гордона, суперкомпьютер.

Введение

Изучение динамики солитонов в последнее время привлекает всё большее внимание исследователей. Связано это с тем, что при их помощи можно моделировать различные физические явления. Уравнение sin-Гордона является классическим для моделирования солитонных решений типа кинка и бризера. Такие солитоны описывают многие процессы в физике твёрдого тела [1].

Изучение литературы показало, что с точки зрения моделирования реальных физических процессов, интерес представляется в исследовании влияния случайных физических возмущений на динамику уравнения sin-Гордона. Учёт влияния возмущений приводит к существенному изменению структуры солитонов, которые описываются, как деформируемые частицы. Неоднородную внешнюю силу чаще всего моделируют как функцию, зависящую от координат и времени. Не менее интересным является случай воздействия пространственных модуляций (неоднородности) внутри самой системы. Обнаружено, что в сильно неоднородном случае форма солитонов в модифицированном уравнении sin-Гордона претерпевает сильное изменение, наблюдается возбуждение внутренних солитонных мод и излучение возбуждений, отрывающихся от солитона в виде свободных волн [2].

Зачастую, нелинейные модуляции внутри системы и внешняя сила задаются с помощью детерминированных функций. В настоящей работе внутренние изменения в системе и/или внешние возмущения в уравнении sin-Гордона задаются с помощью случайных функций, что по своей сути является новым подходом в изучении подобного класса задач. Для достижения высокой точности численного решения стохастического уравнения sin-Гордона требуется проводить дискретизацию по пространственным и временным переменным с малыми шагами и моделировать большое количество реализаций, что является трудоемкой задачей, которая решается с помощью суперкомпьютера. Для численного анализа используется частное временное сечение.

1 Детерминированное уравнение sin-Гордона

Уравнение sin-Гордона встречается в нелинейной оптике, в физике конденсированных сред, в магнетизме, в теории сверхпроводящих джозефсоновских контактов, в статистической механике, при описании динамики

Работа выполнена в рамках госзадания ИВМиМГ СО РАН проект № 0315-2016-0002 и при финансовой поддержке Российского фонда фундаментальных исследований (проекты: № 17-01-00698 А, № 18-01-00599 А).

ISBN 978-5-901548-42-4

доменных границ в многослойных ферромагнетиках [1, 3].

Классическое уравнение sin-Гордона записывается в виде:

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - \sin u \tag{1}$$

Решение (1) в виде кинка или топологического солитона имеет вид:

$$u(x,t) = 4\,\mathrm{arctg}(\exp[-\gamma \Delta V_0 (x - V_0 t)]) \tag{2}$$

где $\gamma = \pm 1$ — топологический заряд кинка; V_0 — непрерывный параметр $V_0 \in (0,1)$, определяющий скорость движения кинка; $\Delta V_0 = \frac{1}{\sqrt{1-V_0^2}}$ — ширина кинка, зависящая от скорости.

Решение (1) в виде бризера или динамического солитона имеет вид:

$$u(x,t,\omega) = 4\,\mathrm{arctg}\,\frac{\sqrt{1-\omega^2}\sin\omega t}{\omega\,\mathrm{ch}(\sqrt{1-\omega^2}x)} \tag{3}$$

где ω — частота колебаний бризера; $u_{max} = 4\,\mathrm{arctg}(\frac{\sqrt{1-\omega^2}}{\omega})$ — амплитуда бризера. Оно представляет собой пространственно локализованную осциллирующую нелинейную функцию с неподвижным центром тяжести [2]. При $\omega \longrightarrow 0$, решение «разваливается» на свободный «кинк» и «антикинк». В этом связь динамических двухпараметрических солитонов «бризеров» с однопараметрическими топологическими солитонами «кинками».

На рис. 1а представлен график точного решения уравнения sin-Гордона (2) типа кинк при $V_0 = 0.5$, $\gamma = -1$. На рис. 1б представлен график точного решения уравнения sin-Гордона (2) типа бризер при $\omega^2 = 0.64$.

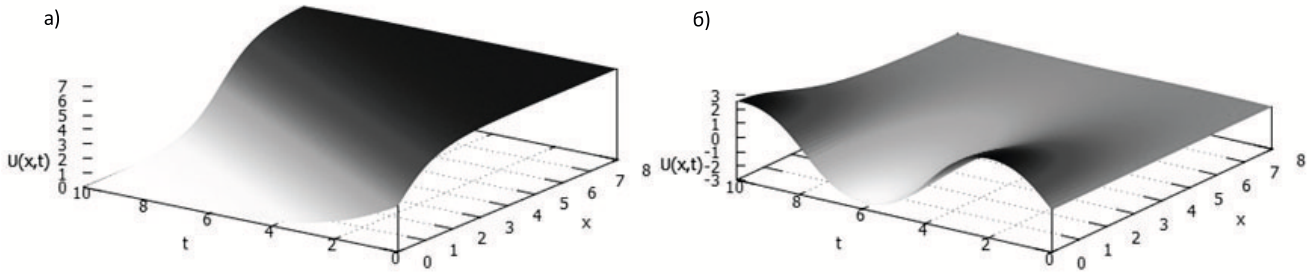


Рис. 1: Точные решения уравнения sin-Гордона типа: кинк (рис. 1а) и бризер (рис. 1б)

2 Стохастическое уравнение sin-Гордона

Начально-краевая задача для стохастического уравнения sin-Гордона имеет вид:

$$\begin{aligned}
&\tfrac{\partial^2 u}{\partial t^2} = \tfrac{\partial^2 u}{\partial x^2} - \sin u + \sigma(x,t,u)\dot{W},\ 0 < x < L;\ 0 < t \le T,\\
&u(x,t)\mid_{t=0} = f(x),\ 0 \le x \le L,\\
&\tfrac{\partial u(x,t)}{\partial t}\mid_{t=0} = g(x),\ 0 \le x \le L,\\
&u(x,t)\mid_{x=0} = p(t),\ 0 < t \le T,\\
&u(x,t)\mid_{x=L} = q(t),\ 0 < t \le T,
\end{aligned} \tag{4}$$

где $\dot{W}$ стандартный белый шум.

Для солитонного решения в виде кинка (2) задаются следующие начальные и краевые условия:

$$\begin{aligned} f(x) &= 4\operatorname{arctg}(\exp(-\gamma\Delta V_0 x)); \\ g(x) &= \tfrac{4\exp(-\gamma\Delta V_0 x)\gamma\Delta V_0 V_0}{1+(\exp(-\gamma\Delta V_0 x))^2}; \\ p(t) &= 4\operatorname{arctg}(\exp(-\gamma\Delta V_0(-V_0 t))); \\ q(t) &= 4\operatorname{arctg}(\exp(-\gamma\Delta V_0(L-V_0 t))). \end{aligned} \tag{5}$$

Для солитонного решения в виде бризера (3) задаются следующие начальные и краевые условия:

$$\begin{aligned} f(x) &= 0; \\ g(x) &= \tfrac{4\sqrt{1-\omega^2}}{\operatorname{ch}(\sqrt{1-\omega^2}x)}; \\ p(t) &= 4\operatorname{arctg}(\tfrac{\sqrt{1-\omega^2}\sin\omega t}{\omega}); \\ q(t) &= 4\operatorname{arctg}(\tfrac{\sqrt{1-\omega^2}\sin\omega t}{\omega\operatorname{ch}(\sqrt{1-\omega^2}L)}). \end{aligned} \tag{6}$$

3 Численный метод решения стохастического уравнения sin-Гордона

Разностная схема для стохастического уравнения sin-Гордона имеет вид:

$$\begin{aligned} &\tfrac{u_j^{k+1}-2u_j^k+u_j^{k-1}}{\tau^2} = \tfrac{u_{j+1}^k-2u_j^k+u_{j-1}^k}{h^2} - \sin u_j^k + \sigma(j,k,u_j^k)\tfrac{\xi_k}{\sqrt{\tau}}; \quad k=1,\ldots,K-1, \quad j=1,\ldots,J-1 \\ &u_j^0 = f_j,\ j=0,\ldots,J \\ &\tfrac{u_j^1-u_j^0}{\tau} = g_j,\ j=0,\ldots,J \\ &u_0^k = p^k,\ k=1,\ldots,K \\ &u_j^k = q^k,\ k=1,\ldots,K \end{aligned} \tag{7}$$

Здесь $\xi_k \in N(0,1)$ — последовательность независимых стандартных нормальных случайных величин, которые моделируются по формуле $\xi_k = \sqrt{-2\ln\alpha_1}\sin 2\pi\alpha_2$ для нечётных k и $\xi_k = \sqrt{-2\ln\alpha_1}\cos 2\pi\alpha_2$ для чётных k, α_1, α_2 — равномерно распределённые случайные величины на отрезке $[0,1]$.

Для исследования влияния случайных шумов на решение уравнения sin-Гордона используется частотное временное сечение (ЧВС). Приведём формальное определение [4]:

На области $X \times U = \{X_0 \le X \le X_{end}; U_0 \le U \le U_{end}\}$ введём сетку $\omega_{h_X h_U} = \{X_i = X_0 + ih_X, i = \overline{(0,N_X)}; U_j = U_0 + jh_U, j = \overline{(0,N_U)}\}$ с шагом $h_X = \frac{(X_{end}-X_0)}{N_X}$ и с шагом $h_U = \frac{(U_{end}-U_0)}{N_U}$. Границы области X_0, X_{end}, U_0, U_{end} и количество узлов сетки N_{X+1}, N_{U+1} вводятся произвольно. ЧВС строится следующим образом: выбирается момент времени t^* и вычисляется количество пар $(x_i, U(x_i,t^*))$, попавших в (X_i, U_j) — узел сетки $\omega_{h_X h_U}$ для всех пространственных координат $x_i, i = \overline{(0,N)}$ и для всех смоделированных траекторий $m = \overline{(1,M)}$. $W = \{W_i j | W_{ij}$ — количество пар $(x_i, U(x_i,t^*)) \in (X_i, U_j), i = \overline{(0,N)}, m = \overline{(1,M)}\}$, где индексы i и j площадки (X_i, U_j) сетки $\omega_{h_X h_U}$ определяются следующим образом: $i = [\frac{(x_i-X_0)}{h_X}]$, $j = [\frac{(U(x_i,t^*)-U_0)}{h_U}]$, [ ] — целая часть числа.

4 Численные эксперименты

Численные эксперименты проводились на кластере НКС-30Т Сибирского Суперкомпьютерного центра при ИВМиМГ СО РАН при следующих значениях параметров: $x \in [0;8]$, временной интервал $t \in [0;10]$, шаг

по пространству $h_x = 10^{-2}$, шаг по времени $h_t = 10^{-6}$, количество моделируемых реализаций $N_{tr} = 10^6$. Заметим, что при таком наборе параметров расчёт одной реализации на персональном компьютере занимает несколько минут. По этой причине данная задача решалась с помощью супер-ЭВМ. В расчётах использовалось 90 ядер кластера. Подготовка численных экспериментов проводилась при помощи комплекса программ SPARDE1D [5].

Тест 1.

На рис. 2 приведены графики временного сечения для детерминированного уравнения sin-Гордона и ЧВС для стохастического уравнения sin-Гордона при $t = 2$ (слева) и при $t = 8$ (справа) для случая кинка (5). Показано влияние аддитивного шума с интенсивностью $\sigma = 0.6$. С увеличением времени динамика отдельных реализаций стохастического уравнения sin-Гордона по сравнению с динамикой точного решения детерминированного уравнения sin-Гордона качественно не изменяется, а имеет место разброс отдельных реализаций относительно точного решения. Дополнительные численные эксперименты показали, что увеличение интенсивности аддитивных шумов увеличивает разброс реализаций.

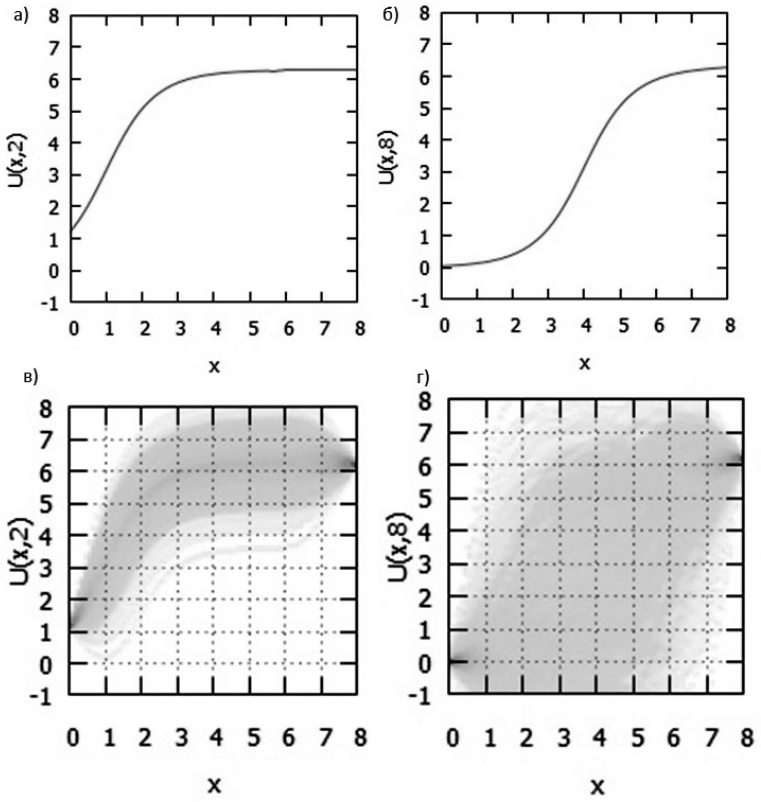


Рис. 2: Временное сечение и ЧВС для уравнения sin-Гордона (рис. 2а, 2в при $t = 2$; рис. 2б, 2г при $t = 8$)

Тест 2

На рис. 3 приведены графики временного сечения для детерминированного уравнения sin-Гордона и ЧВС для стохастического уравнения sin-Гордона при $t = 6$ для случая бризера (6). Показано влияние аддитивного и мультипликативного шумов с интенсивностью $\sigma = 0.3$. Так как в мультипликативном шуме есть зависимость от решения, то начиная с точки по пространству $x = 7$ решение начинает стремиться к нулю, поэтому влияние мультипликативного шума стремиться к нулю, в отличие от аддитивного. Как видно из рис. 3, начиная с точки по пространству $x = 7$, все реализации стохастического уравнения sin-Гордона стягиваются к точному решению детерминированного уравнения sin-Гордона, что кардинально отличается от случая с аддитивным шумом, в котором сохраняется разброс реализаций относительного точного решения. Когда $u(x, 6) > 1$ мультипликативный шум увеличивает амплитуду некоторых реализаций больше, чем аддитивный шум.

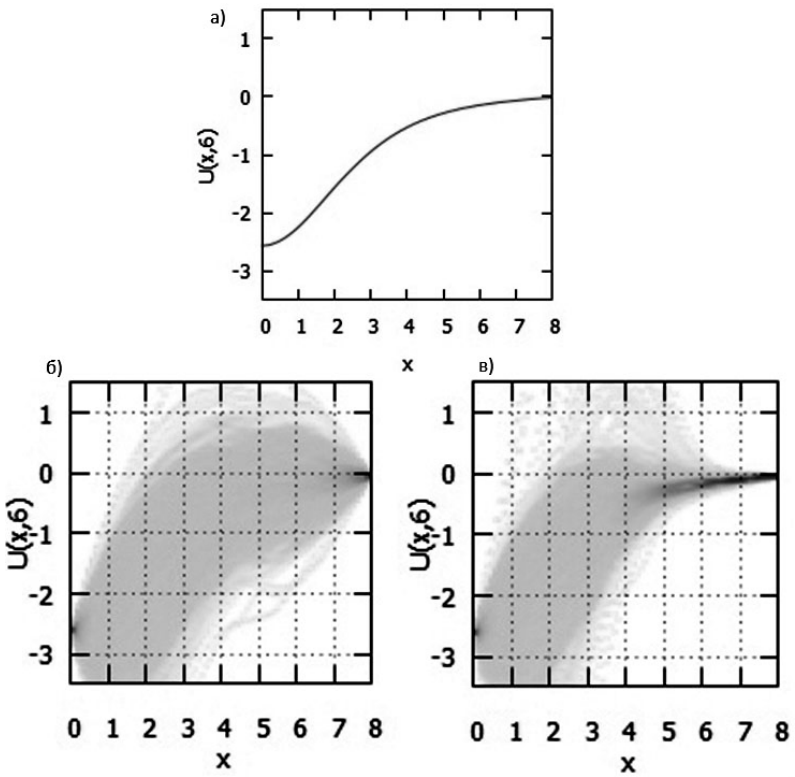


Рис. 3: Временное сечение и ЧВС для уравнения sin-Гордона при $t = 6$ (рис. 3б для аддитивного шума, рис. 3в для мультипликативного шума)

Заключение

Для анализа решений стохастического уравнения sin-Гордона (кинка и бризера) было расчитано ЧВС в различные моменты времени и для разных видов случайных шумов: аддитивного и мультипликативного, которое показывает динамику всех реализаций решения и степень влияния случайных шумов ризличной интенсивности. Графики ЧВС для заданного момента времени показывает распределение всех реализаций решения стохастического уравнения sin-Гордона (7) во всех точках по пространству. Влияние аддитивного шума качественно не изменяет динамику решения, влияя только на разброс реализаций относительно точного решения. Влияние мультипликативного шума почти исчезает при стремлении решения к нулю. В случае, когда решение больше единицы, влияние мультипликативного шума становится больше влияния аддитивного шума. Алгоритмы численного статистического моделирования обладают высокой эффективностью распараллеливания ввиду того, что на разных процессорах моделируются независимые между собой реализации случайного процесса.

Список литературы

[1] А. М. Гумеров, Е. Г. Екомасов, Р. Р. Муртазин, В. Н. Назаров, “Трансформация солитонов уравнения синус-Гордона в моделях с переменными коэффициентами и затуханием”, Ж. вычисл. матем. и матем. физ., 55:4 (2015), 631–640; Comput. Math. Math. Phys., 55:4 (2015), 628–637

[2] Кудряшов Н.А. Нелинейные волны и солитоны. Московский государственный инженерно-физический институт. Физика, 1997.

[3] Екомасов Е.Г. Солитоны модифицированного уравнения синус– Гордона. Учебное пособие. — Уфа: РИЦ БашГУ, 2009. — 94 с.

[4] Артемьев С.С., Марченко М.А., Корнеев В.Д., Якунин М.А., Иванов А.А., Смирнов Д.Д. Анализ стохастических колебаний методом Монте-Карло на суперкомпьютерах. Новосибирск: Издательство СО РАН, 2016. 294 с.

[5] Артемьев С.С., Иванов А.А. Свидетельство о государственной регистрации программы для ЭВМ № 2018617942, "SPARDE1D - программный комплекс для численного анализа стохастических дифференциальных уравнений в частных производных". Дата регистрации: 05.07.2018.

Иванов Александр Александрович — м.н.с. Института вычислительной математики и математической геофизики СО РАН;
e-mail: aaivanov@sscc.ru;
Смирнов Дмитрий Дмитриевич — м.н.с. Института вычислительной математики и математической геофизики СО РАН;
e-mail: smirnovdd@mail.ru;
Дата поступления — 30 апреля 2019 г.

BOUNDARY ELEMENT METHOD IN 3D PROBLEMS OF MATHEMATICAL PHYSICS

V. Ivanov

Institution of Computational Technologies SB RAS, 630090, Novosibirsk

UDC 519.6
DOI: 10.24411/9999-016A-2019-10027

Numerical algorithms of the boundary problems described by Laplace, Poisson, Helmholtz and Maxwell equations based on the integral representations are described. Combinations of single-layer, double-layer and volume potentials are used to represent the solution. Analytical technique of singularity extraction is used to obtain precision and stable numerical solutions for a potential and its high-order derivatives. The numerous examples of test problems and applications to the precision problems of electron optics, accelerator physics and plasma-beam interaction are demonstrated. The original algorithms can be used for the problems of analysis, optimization and synthesis of physical electronic devices. Decomposition algorithms to solve complex three-dimensional problems are presented.

keywords: Boundary Element Method, integral representation, single-layer potential, double-layer potential, numerical approximation, singularity.

Introduction

The mathematical apparatus for describing natural phenomena is usually called the equations of mathematical physics. These equations admit different forms of representation. The most well-known are writing in the form of partial differential equations, variation formulations of problems and integral representations. Thus, the stationary problems of electromagnetism, hydrodynamics, and thermal physics lead to boundary value problems for the Poisson equation

$$\Delta\phi = \rho, \tag{1}$$

where Δ — is the Laplace operator, ϕ — is a potential, but ρ — is a density of field sources. The time-dependent equations of electrodynamics are described by the system of Maxwell's equations

$$\begin{aligned} \frac{\partial \vec{B}}{\partial t} + rot\vec{E} &= 0, \\ \frac{\partial \vec{D}}{\partial t} - rot\vec{H} &= \vec{J}, \\ div\vec{B} &= 0, \\ div\vec{D} &= 0, \end{aligned} \tag{2}$$

Here $\vec{E}$ and $\vec{H}$ — are the strength of electric and magnetic fields, $\vec{D} = \varepsilon\vec{E}$ and $\vec{B} = \mu\vec{H}$ — are the vectors of electric and magnetic induction, $\vec{J}$ — is a current density of excitation field, $c = \frac{1}{\sqrt{\varepsilon\mu}}$ — is a speed of light in vacuum, ε and μ — are the permittivity and permeability of a physical environment.

To obtain the integral representation, we use the second Green's identity

$$\iiint_V (u\Lambda v - v\Lambda u)dV = \oiint_S \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) dS, \tag{3}$$

which is valid for any two harmonic functions u and v. Here, on the right-hand side of the identity, integration is over the boundary S of the volume V occupied by the field, and n is the internal normal vector to this boundary.

ISBN 978-5-901548-42-4

If as the trial function u we choose the potential $\phi(p)$ at some arbitrary point p, and put the function v equal to $v = 1/R_{pq}$, where R_{pq} — is the distance between the observation point p and the source point q, over which the integration in formula (3), then, taking into account the relation $\Delta\frac{1}{R_{pq}} = \delta$ and the property of the δ-function, we obtain the required representation

$$\phi(p) = \oiint \phi(q)\frac{\partial}{\partial n_q}\frac{1}{R_{pq}}dS_q - \oiint \frac{\partial\phi(q)}{\partial n_q}\frac{1}{R_{pq}}dS_q + \iiint \frac{\rho(q)}{R_{pq}}dV_q. \tag{4}$$

In this representation, the first integral is called the double-layer potential, the second is the simple-layer potential, and the third is called the Newtonian potential.

The integral representations for the system of Maxwell equations have the form

$$\begin{aligned} \oint \vec{E}d\vec{l} &= -\iint \frac{\partial\vec{B}}{\partial t}d\vec{S}, \\ \oiint \vec{B}d\vec{S} &= 0, \\ \oint \vec{H}d\vec{l} &= \iint \left(\frac{\partial\vec{D}}{\partial t} + \vec{J}\right)d\vec{S}, \\ \oiint \vec{D}d\vec{S} &= \iiint \rho dV. \end{aligned} \tag{5}$$

The properties of integral equations was represented in fundamental monographs by C. Baker [1]. In the numerical solution of the equation of mathematical physics, they are approximated by some set of orthogonal functions H_m. So that the norm of the difference of the original functions H and the approximated ones satisfies the relation $\|H - H_m\| \approx o(h^m)$, where h — is the characteristic scale of the discretization of the equations, and the number m is called the order of approximation. After sampling, the original equations reduce to a system of linear algebraic equations.

The use of non structured mesh with some geometrical elements on the boundary surface produces the Boundary Element Method (BEM) The technology of development of numerical algorithms of BEM is considered in monographs [2]– [4], and the convergence of these algorithms is investigated in [5]. Consider a mixed boundary-value problem for the Poisson equation (1). Let the Dirichlet boundary condition be given on the part S_1 of the boundary of the domain

$$\phi\,|_{S_1} = U_0(p)\ , p \in S_1, \tag{6}$$

but at the rest part of the boundary S_2 the Neumann condition is given

$$-\frac{\partial\phi}{\partial n}\,|_{S_2} = E_0(p)\ , p \in S_2. \tag{7}$$

Here U_0 and E_0 — are given continuous functions, and n is the inner normal to the boundary. Applying the boundary condition (6) to the integral representation (4), one can obtain the Fredholm integral equation of the first kind with respect to the unknown function σ, but of the second kind equation with respect to the unknown function μ at the boundary

$$\iint_{S_2} \nu(q)\frac{\partial}{\partial n_q}\frac{1}{R_{pq}}dS_q + 2\pi\,\mu(p) + \iint_{S_1}\frac{\sigma(q)}{R_{pq}}dS_q = f_1(p), p \in S_1, \tag{8}$$

$$\iint_{S_2} \nu(q)\frac{\partial}{\partial n_p}\frac{\partial}{\partial n_q}\frac{1}{R_{pq}}dS_q + 2\pi\,\sigma(p) + \iint_{S_1}\sigma(q)\frac{\partial}{\partial n_p}\frac{1}{R_{pq}}dS_q = f_2(p), p \in S_2, \tag{9}$$

where

$$f_1(p) = U_1(p) - \iiint_V \frac{\rho(q)}{R_pq}dV_q, p \in S_1, \tag{10}$$

$$f_2(p) = -E_0(p) - \iiint_V \rho(q)\frac{\partial}{\partial n_p}\frac{1}{R_{pq}}dV_q, p \in S_2. \tag{11}$$

The kernels of the integral equationd have the singularities at $p \to q$. The regularization procedure consists in adding a term to the integrand, the integral of which can be calculated analytically. In this case, the equation in the neighborhood of the singularity S_0, where $p \to q$ becomes

$$\iint_{S_0} \frac{\sigma(q) - \sigma(p)}{R_{pq}} dS_q + \sigma(p) \iint_{S_0} \frac{1}{R_{pq}} dS_q = U_0(p). \tag{12}$$

The integrand in the first integral has no singularity, and the second integral can be calculated analytically when the fragment of the surface S_0 is represented by a second-order surface. Thus, equation (12) is an integral equation of the third kind.

1 Three-dimensional algorithms

When the BEM implementing for three-dimensional problems, a dilemma arises, using structured or non-structured meshes when discretizing the boundaries of the region. We demonstrate the effectiveness of numerical algorithms for both versions [6]– [7]. Let the boundary of the domain consist of m smooth surfaces $S = \bigcup_{i=1}^{m} S_i$, each of which admits a parametric representation

$$v = v_i(\tau, \eta), y = y_i(\tau, \eta), z = z_i(\tau, \eta), \tau \in [\alpha_i, \beta_i], \eta \in [\gamma_i, \delta_i]. \tag{13}$$

Details of algorithms for solving three-dimensional problems are described in the monographs [6-7]. So in the case of a bilinear spline for surface sources

$$\sigma(\tau, \eta) = \{[\sigma_{i+1,j+1}(\tau - \tau_i) + \sigma_{i,j+1}(\tau_{i+1} - \tau)](\eta - \eta_j) + [\sigma_{i+1,j}(\tau - \tau_i) + \sigma_{i,j}(\tau_{i+1} - \tau)](\eta_{j+1} - \eta)\} / (h_\tau h_\eta), \tag{14}$$

where $h_\tau = \tau_{i+1} - \tau_i, h_\eta = \eta_{j+1} - \eta_j$ finally we get the linear algebraic system $\hat{G}X = F$ with the global matrix $\hat{G}$, unknown vector X and right-hand vector F, which includes the boundary conditions. Matrix elements are given with the formula

$$\hat{G}_{i,j} = \frac{1}{h_\tau h_\eta} \sum_{m=1}^{4} \iint \frac{\psi_m(\tau, \eta)}{R_{pq}} J(\tau, \eta) d\tau d\eta. \tag{15}$$

Here ψ_m — are the moments of the bilinear form in the approximation of the field sources, J — Jacobian of the coordimate transformation.

Let us consider the non uniform parallelepiped mesh in Cartesian coordinates $\{x_i\} \times \{y_j\} \times \{z_k\}$ with linear approximation for the space charge [9]– [10]. In order to evaluate the potential:

$$\phi(x, y, z) = \sum_{i,j,k} \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \int_{z_k}^{z_{k+1}} \frac{\rho(x', y', z')}{R(x - x', y - y', z - z')} dx\, dy\, dz. \tag{16}$$

One need the four moments of space charge

$$J_1 = \iiint \frac{x}{R} dx\, dy\, dz, J_2 = \iiint \frac{xy}{R} dxdydz, J_3 = \iiint \frac{xyz}{R} dxdydz, J_4 = \iiint \frac{1}{R} dxdy, dz, \tag{17}$$

where $r^2 = (x - x')^2 + (y - y')^2 + (z - z')^2$. The other integrals are obtained by cyclic change of variables x, y, z. All integrals can be evaluated analytically.

The CPU-time T for the boundary problem solving can be represented with the formula $T = aN^3 + bN^2$, where N — is the dimension of matrix $\hat{G}$, but a and b — are some numerical constants. First term of this formula is a time for solving of the linear system by the Gauss elimination method, and second term is a time for evaluation of the matrix elements. First term dominates for large enough N. To overcome this strong cubic dependence, H.Schwarts suggested the decomposition idea, when a large problem is spitted onto K, overlapped blocks separated each other with the auxiliary boundaries. Then K boundary problems of lower dimensions are solving in the iterative loop, where the boundary condition on the auxiliary bounds are evaluated from the solution of neighbor block at previous iteration

We can make an estimation for the efficiency of decomposition algorithm comparing with the direct solution of the whole problem of dimension N. Let in simplest case all blocks have equal dimensions, i.e. $N = Km$. Then the relative efficiency for large N can be estimated by the formula

$$E = \frac{aN^3 + bN^2}{lK(am^3 + bm^2)} \approx \frac{J^2}{k}, \tag{18}$$

where l — is a total number of iterations to reach the predefined accuracy ε. Publication [11] shows that typical value $l = 3 \div 4$ for $\varepsilon = 0.1\%$.

Another approach is based on the use of an unstructured surface mesh. In this case, first the boundary surface is transformed into a set of geometric elements in a simple form, as in the FEM. Polynomial approximations are set on the surface of the elements, and the natrix elements are calculated. In many cases, integrals over the surface of elements can be calculated analytically. For example, in the works of M. Okon and R. Harrington [12]– [14] analytical expressions for the potential of triangular elements with piecewise constant and linear distribution of the charge density on the surface of an element.

We also give an example of the solution of the equations of electrodynamics in the case of harmonic oscillations. We write down the integral equations for the component H_τ, H_η of the magnetic field [8]:

$$\begin{aligned} H_\tau(\tau_0, \eta_0) &= \frac{1}{\Omega - 4\pi} \oint_S \{G_{\tau_0,\eta} H_\tau(\tau, \eta) - G_{\tau_0,\eta}(\tau, \eta)\}\, dS, \\ H_\eta(\tau_0, \eta_0) &= \frac{1}{\Omega - 4\pi} \oint_S \{G_{\eta_0,\eta} H_\tau(\tau, \eta) - G_{\eta_0,\eta}(\tau, \eta)\}\, dS. \end{aligned} \tag{19}$$

Here $(\tau_0, \eta_0) \in S$ — are the values of variables τ, η at the observation point , Ω — is the solid angle at which the surface S is visible at this point. The tensor Green's function $G_{\alpha,\beta}(\tau_o, \eta_0; \tau, \eta) = [\alpha \nabla \phi]\beta, \phi = cod(jR)/R, k = \omega/c$ is an elementary current component $j = \alpha\delta(r_0)\exp[i(kr_0 - \omega t)], r_0 = (\tau_0, \eta_0)$, oriented in the direction of the vector α and varying in harmonic law with a frequency ω, c — is a speed of light, $R = \|r_0 - r\|$ — is a distance between the observation point r_0 and the integration point $r = (\tau, \eta)$. The system of integral equations (19) can be written in the operator form $[I - L(\omega)]H_t = 0$, where I — is the unit operator, but H_t — is the component of the magnetic field vector tangential to the surface S. With numerical approximation, this equation reduces to a system of linear equations $A(\omega)H_t = 0$ with square matrix A, which approximate the operator $I - L(\omega)$. A nontrivial solution of this system exists if $\det(A) = 0$.

If bilinear approximation is used for an unknown vector-valued function Ht, the boundary collocations method reduces the original problem to a linear system with the matrix elements

$$G_{\alpha,\beta} = \frac{1}{h_\tau H_\eta} \sum_{m=1}^{4} \int_{S_m} (\tau_0, \eta_0; \tau, \eta) J d\tau d\eta. \tag{20}$$

The function $G_{\alpha,\beta}$ has an integrable singularity, for the separation of which the integral (19) is represented in the form

$$\int_{S_m} \psi G_{\alpha,\beta} J d\tau d\beta = \int_{S_m} \psi_m (G_{\alpha,\beta} - \tilde{G}_{\alpha,\beta} J d\tau d\eta + \int_{S_m} \psi_m \tilde{G}_{\alpha,\beta} J d\tau d\eta, \tag{21}$$

where $\tilde{G}_{\alpha,\beta} = \lim_{k\to 0} G_{\alpha,\beta}$ The integrand in the first integral on the right-hand side does not contain a singularity, and the second integral for quadratic surfaces is calculated analytically.

2 Solving practical problems

As a practical problem, let us consider the design of a gun of an 80-megawatt klystron with a sheet beam [16]– [17]. Parameters of the gun: current 250A, accelerating voltage 415 kV, working frequency of klystron 91 GHz. In traditional klystrons with a cylindrical beam, the effect of the space charge creates serious obstacles to constructing devices with a power of more than 75 megawatts, so more powerful klystrons can be built on new principles: multi-beam or with a sheet beam. In multi-beam klystrons, when bundling individual beams into a single bundle, it is difficult to obtain a beam with a small transverse phase volume, therefore, the production of laminar beams has prospects in a plan with a sheet beam. The geometry of the gun of such a klystron is shown in Figure 1.

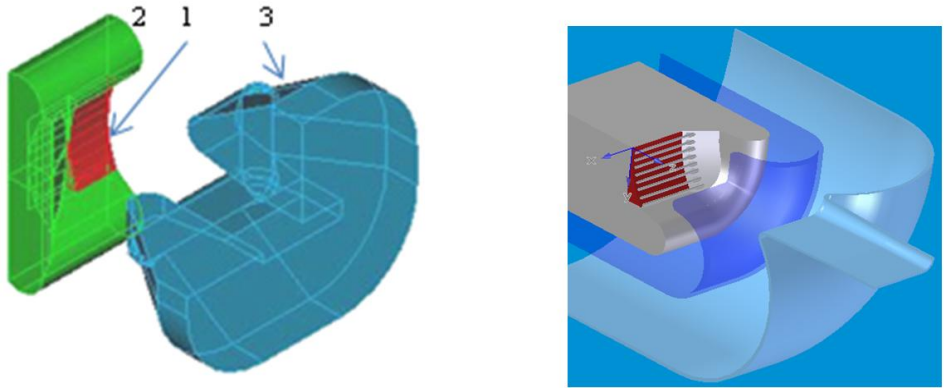


Figure 1: Diode gun on left (1 — cathode, 2 — focusing electrode, 3 — anode) and triode gun (on right) of powerful sheet-beam klystron

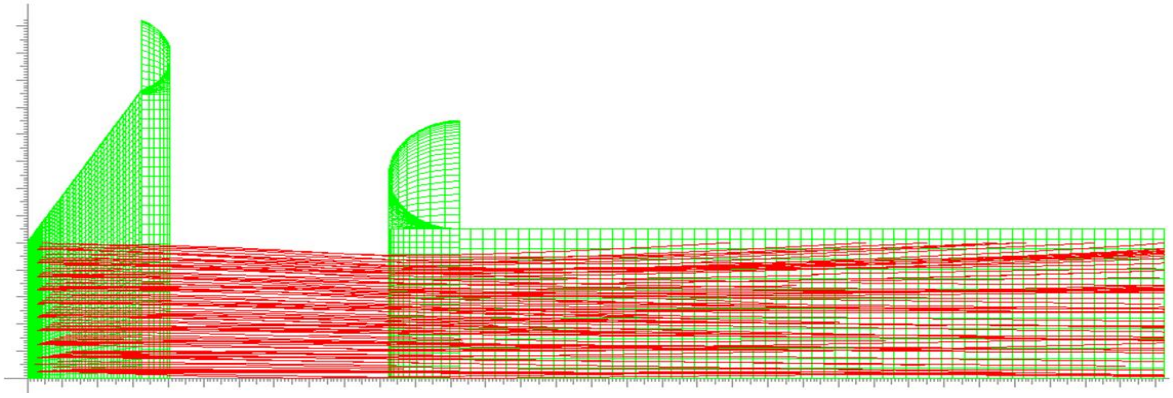


Figure 2: Klystron gun simulation on structured mesh

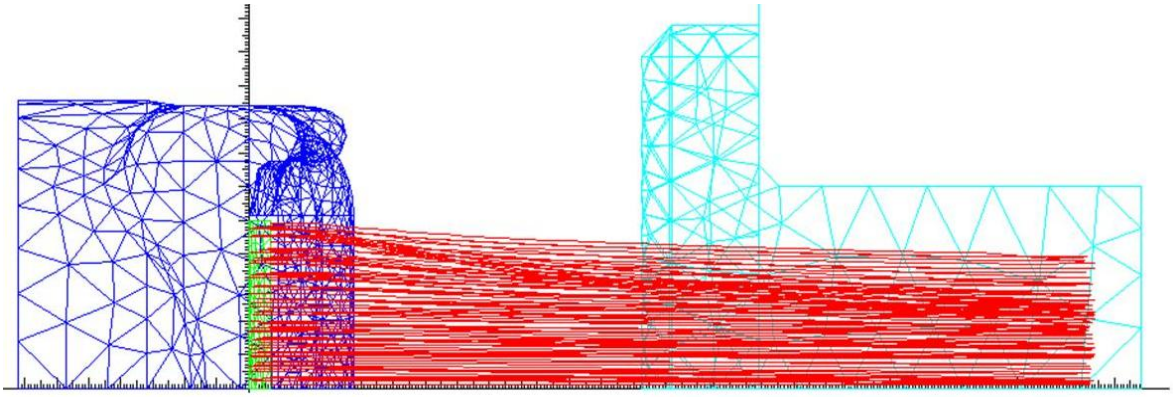


Figure 3: Klystron gun simulation on unstructured mesh

A significant beam current and a high degree of relativism lead to the problem of a self-consistent field in which it is required to take into account the Coulomb field of the beam charge and the self consistent magnetic field of the relativistic beam.The results of the calculations are presented in Figures 2 and 3. The first one is based on a structured grid. In the calculation shown in Figure 3, the electrode surfaces are approximated by a set of triangles that are milled in regions with small radii of curvature. For a beam with a current of 250 A, it took about 10 iterations to obtain current convergence with an accuracy of 0.5

As an example of a practical problem of calculating three-dimensional electrodynamic structures [8], a cell of a biperiodic structure with an inductive coupling was depicted in Fig. 4. The results of experimental measurements of the resonance frequency on the type of oscillations $\pi/2$ corresponded to 430 MHz, and numerous calculations

for a surface partition by 40 elements gave the value of this the value of 434 MHz. When the number of elements was enlarged to 826, the relative error in calculating the natural frequencies was 5 · 10-4

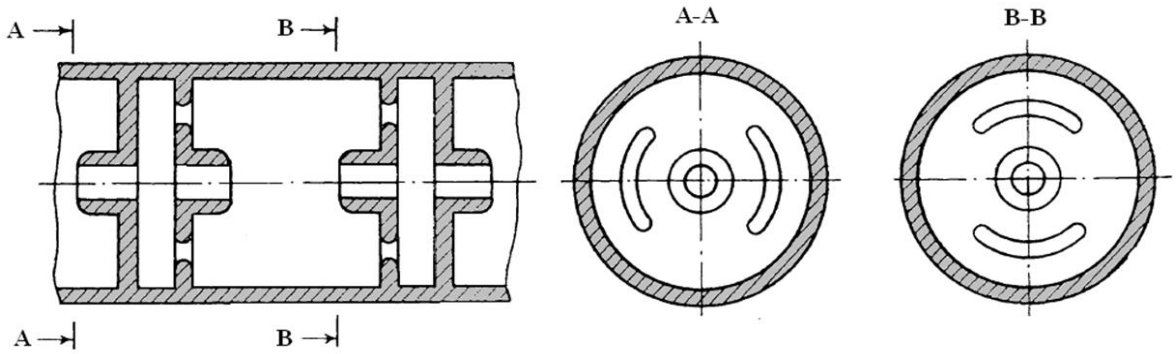


Figure 4: A cell of bi-periodic RF-structure

References

[1] Baker C. T. The Numerical Treatment of Integral Equations. Oxford: Oxford University Press, 1977.

[2] Banerjee P. K. et al., editors. Developments in Boundary Element Methods, Vols 1,2,3, Applied Scien;ce Publishers, London, 1979.

[3] Brebbia C. A, Telles J. C, Wrobel L. C. Boundary Element Techniques. Berlin: Springer-Verlag; 1984.

[4] Mackerle J, Andersson T. Boundary element software in engineering. Advances in Eng. Software. 1984; 6: 66-102.

[5] Wendland W. L. Boundary element methods and their asymptotic convergence. In P. Filippi, editor. Theoretical Acoustics and Numerical Techniques, CISM Courses. Wien, New York: Springer-Verlag; 1983: 135-216.

[6] Ivanov V. CAD methods for the electronic devices. Part 1: methods for calculating physical fields. Novosibirsk: Inst. of Mathematics; 1986. 194 pp. In Russian.

[7] Ivanov V. Computational methods, optimization and synthesis in electron optics. Hmbg: Palmarium Academic Publishing; 2016. 525 pp.

[8] Ivanov V, Teryaev V, Karliner M, Yakovlev V. Application of the method of boundary integral equations for the calculation of high-frequency resonators. Zhurnal Vychislitelnoi Matematiki I Matematicheskoi Fiziki. 1986; 12 :.1900-1905. In Russian.

[9] Ivanov V. Green's Function Technique in Forming of Intensive Beams. Int. J. of Modern Physics A, 2009; 24(5): 869-878.

[10] Ivanov V. Analytical Technique in the Boundary Element Method for 3D Problems of Electron Optics. Microscopy & Microanalysis. 2015; 21(4): 236-241.

[11] Ivanov V. Decomposition method for complex 3D problems in electron optics. J. of Physical Science & Applications. 2015; 5(2): 96-100.

[12] Okon M, Harrington R. F. The potential due to a uniform source distribution over a triangular domain. Int. J. for Numer. Meth. in Engng. 1982; 18: 1401-1419.

[13] Okon M, Harrington, R. F. The potential integral for a linear distribution over a triangular domain. Int. J. Numer. Methods Engng. 1982; 18: 1821-1828.

[14] Okon M, Potential integrals associated with quadratic distribution in a triangular domain, Int. J. Numer. Methods Engng. 1982; 21(2): 197-209.

[15] Ivanov V. The method of analysis of three-dimensional nonstationary flows of charged particle. In: Numerical analysis. In: Proc. Inst. of Mathematics SB RAS. Novosibirsk, Nauka. 1989; 15: 172-187. In Russian.

[16] Ivanov V, Krasnykh A. Analytical computations in solution of 3D problems of physical electronics by BEM. In: Proceedings of the ACES-2002, March 18-22, 2002, Monterey, CA.

[17] Ivanov V, Krasnykh A. Scheitrum G, Sprehn D, Ives L, Miram G. 3D modeling activity for novel high power electron guns at SLAC. In: Proceedings of the Particle Accelerator Conference (PAC 03), Portland, Oregon, 12-16 May 2003: 3312-3314.

Ivanov Valentin Yakovlevich — Doctor of Sciences (Physics and Mathematics),
Institute of Computational Technologies SB RAS;
e-mail: vivanov.48@mail.ru.
Received — 30 April 2019.

MICROCHANNEL AMPLIFIER BEHAVIOR IN STRONG MAGNETIC FIELD

V. Ivanov

Institution of Computational Technologies SB RAS, 630090, Novosibirsk

UDC 519.6
DOI: 10.24411/9999-016A-2019-10028

Microchannel plate photomultiplier tubes (MCP PMT) can work in a high magnetic field and have an excellent time resolution. The influence of the magnetic fields up to 4.5 T on the parameters of several MCP PMTs of different designs was investigated. PMTs with two, three and four MCPs were simulated and tested in magnetic fields. Description of mathematical models for fast photo detectors based on microchannel plates (MCP) in three-dimensional formulation is given [1–6]. The models include calculations of photoelectron collection efficiency in the gap photo cathode — MCP, gain factor of secondary electron cascades in the channels, the particle scattering in the gaps between the plates, taking into account the fringe fields and strong external magnetic fields. Comparisons of numerical and experimental data are given [7–11]. The dependencies of major device parameters vs. of applied voltage, pore size, and magnetic field magnitude have been studied.Dependencies of the time resolution, the gain and the photoelectron collection efficiency on the magnetic field are presented.

Keywords: microchannel plate, Monte-Carlo method, cascade amplification.

Introduction

Microchannel amplifiers (Figure 1) are widely used in astrophysics, medical diagnostics, accelerator physics, and night vision devices. They have many advantages: compactness, high gain, work stability in radiation and in strong magnetic fields. An amplifier consists of a photo cathode, one or more microchannel plates (MCP), anode and high speed electronic circuits that process signals at the output of the amplifier. Theoretical background of microchannel amplifier simulations is described in the monograph [4]. The computer design consists of two independent phases: microscopic and macroscopic.



Figure 1: Electron trajectories passing through a SiO_2 multi-layer structure. Color shows the particle energy.

The microscopic phase should give the emission properties of materials used in the device. First, a parameterized set of the secondary emission yield (SEY) dependencies in two variables, the energy of the primary electron and the angle of incident electrons for lead glass should been developed. This parameterization can be done by

ISBN 978-5-901548-42-4

using results obtained from Monte Carlo calculations or from experimental measurements. As a result, it is necessary to obtain the angular and energy distributions for secondary emission coefficient, as the parameters of the electron flux incident on the surface of the material. Experimental measurements for the required characteristics are associated with considerable time and resources of the measuring complex. It takes about three months for one sample.

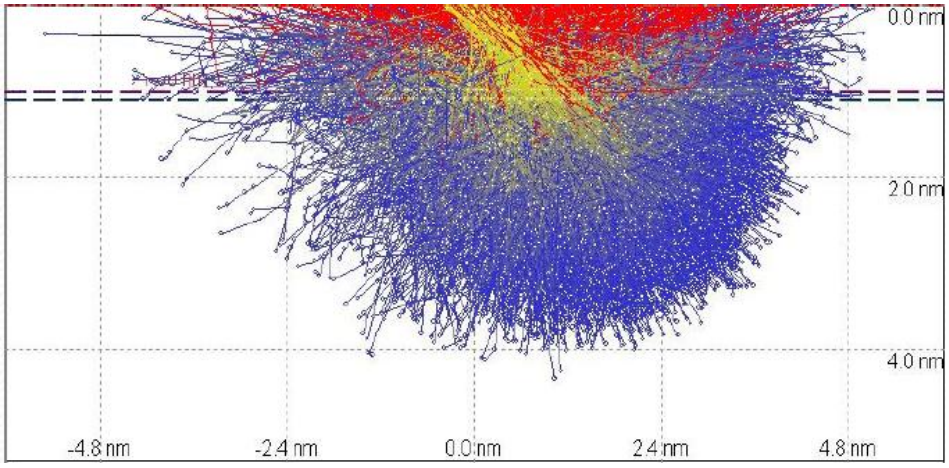


Figure 2: Electron trajectories passing through a SiO_2 multi-layer structure. Color shows the particle energy.

The other way to get these data is computer simulation based on Monte-Carlo algorithms for the numerical models of molecular dynamics [3-4]. An example of numerical simulation with using the code "CASINO" [12] is shown in the Fig. 2, where a beam of 10^5 electrons with random initial data falls on the surface with random positions of atoms at the angle of 45°. Random data should represent the statistical properties of the interaction of a particle beam in the surface layer of matter. Part of the primary particles incident on the surface, experiencing elastic and inelastic collisions with atoms and molecules of the surface layer of matter, will dissipate, penetrating deep into the substance. In this case, the incident particles lose their kinetic energy in inelastic collisions, forming cascades of free electrons. The other part of the electrons is scattered to the side of the surface. If their kinetic energy in this case exceeds the so-called work function of the substance, they are emitted into the vacuum, generating electrons of secondary emission.

The ensembles of secondary particles are described by distribution functions in the angles and emission energies, which can be determined from the results of statistical modelling by means of the averaging procedure. Let N_{in} be the number of primary particles, and N_{out} the total number of secondary particles with their maximal energy E_m. The coefficient of secondary emission is the ratio $\sigma = N_{out}/N_{in}$. The ensembles of secondary particles are described by distribution functions in the angles and emission energies, which can be determined from the results of statistical modelling by means of the averaging procedure. In order to calculate the distribution function of the secondary particles with respect to the energies of their emission, we divide the energy range E_m onto intervals ΔE. We denote by N_i the number of secondary particles trapped in the i-th interval with energies $(E_i, E_{i+1}, \Delta E)$. Then the energy distribution function is represented by the values

$$\sigma(E) = \frac{\sum_{i=1}^{N_{out}} N_i}{N_{out}}. \tag{1}$$

Similarly, the distribution function over the angles of departure can be determined. This function depends on the angle of incidence of the primary particles. The results of the Monte-Carlo simulations for different materials used as secondary emitters are represented in the Fig. 2.

1 Calibration procedure

The SEY dependence can be approximated with semi empirical models. The earliest model was proposed by A.Guest [13]:

$$\frac{\delta(\theta,\varepsilon)}{\delta_m(0,\varepsilon_m} = \left(\frac{\varepsilon}{\varepsilon_m}\sqrt{\cos\theta}\right)^{\beta} \exp\left[\alpha(1-\cos\theta)+\beta\left(1-\frac{\varepsilon}{\varepsilon_m}\sqrt{\cos\theta}\right)\right], \quad (2)$$

where $\delta(\theta,\varepsilon)$ — is a SEY, δ_m — is a maximal value of SEY, ε — is an incident particle energy, ε_m — is an incident energy corresponds to the value δ_m, θ — is an incident angle, and α and β — are the adjustable parameters chosen to fit the experimental data.

Another approximations were suggested by R.Lye and A.Dekker [14], B.Agarval [15], Ito et al. [16] and me other authors [17]. These different approximations are shown in Fig. 3.

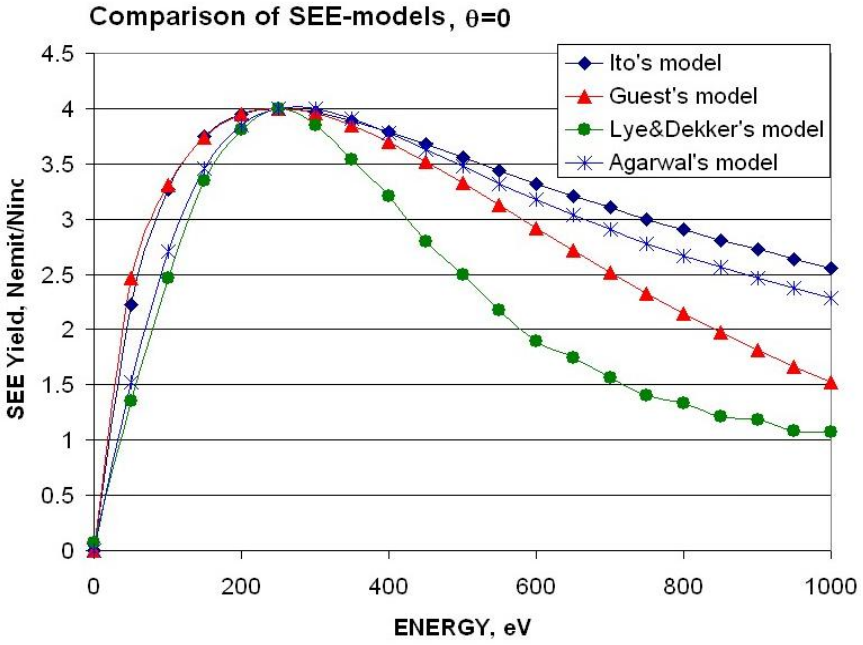


Figure 3: Different appfoximation for SEY vs. incident energy.

The use of analytical approximations instead of time-consuming stages of micro-modelling significantly reduces the calculation time. The coefficients that enter into these formulas can be determined from a comparison of the analytics and the experimental data. This stage is called the calibration procedure of a mathematical model. It is easiest to perform a calibration for a single MCP, since the data for a chevron pair or three plates depend not only on the MCP thickness and the diameter of the channels, but also on the gaps between the MCPs.

2 Numerical simulations

The properties of different MCP assemblies studied numerically are presented in the Table 1. First step was the calibration procedure in order to determine the emission parameters of the MCP channels. The emission coefficient δ_m and corresponding energy ε_m of incoming particles for lead glass can be determined by comparison of computations and experimental data for a single plate. The comparisons were done for single plate 89664, and the results were used for other MCPs.

The calibration results for single MCP 89664 show the optimal fitting between numerical and experimental data was reached for the SEY value δ_m=5.63, and the incident energy ε_m = 350 eV. Those values were used in further MCP simulations.

Fig. 4 shows the numerical simulations and experimental data for different assemblies. It demonstrates a good agreement for all pores with 5 degrees tilt angle, and more difference for 13 degrees angle. In the latter case, there are some features for the MCP 88353, which explains the comparison between the experimental and numerical data. The analysis of these data shows that the calibration procedure gives a good match with the experimental data when the following conditions are met:

a) all MCP items should be made from the same materials;

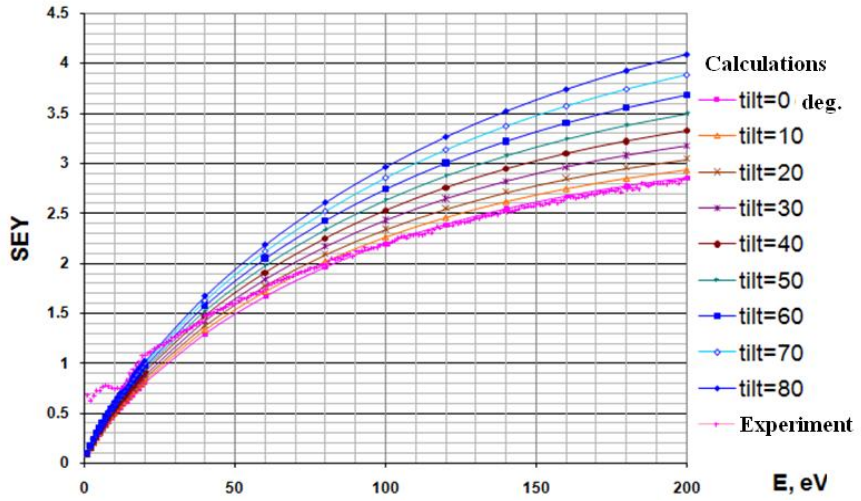


Figure 4: SEY vs. an incident energy and incident angles for 20 nm Al_2O_3.

Table 1: Parameters of the MCP assemblies

Type	Name	Channel diameter D, μm	L/D	Tilt angle, degree
Single	89664	12.57	160	5
Chevron	74	6	50	5
Chevron	88353	7	40	13
Chevron	82015	10	40	5
Z-stack	418	7.5	43	5

b) they must all be of equal purity;

c) when they are manufactured, the same technology should be used.

In other words, all samples must be from the same manufacturer. In our case, this condition was not met for product 88353, which explains the difference between the experimental and numerical data.

More detailed data about the comparison of numerical calculations with experiments for the MCPs in electric fields with no external magnetic fields are presented in publications [7-11].

3 Discussion and summary

In order to provide numerical simulations the code "MCPS" (MCP Simulator) was implemented [4]. Early versions of these codes used the two-dimensional deterministic models for stationary input signals (images) to simulate night vision devices and light amplifiers. It was assumed that the electric field in the microchannel and in the gaps between the electrodes is homogeneous, that is, it changes abruptly at the boundaries of the sub regions. This meant the absence of fringe fields in which the field in the transition regions varies smoothly, which requires the use of complex algorithms for the calculation of such fields.

More sophisticated three-dimensional versions of the code "MCS-3D" (Monte-Carlo Simulator) were used to model multi-plate amplifiers taking into account the fringe fields, saturation and other effects [18]. It uses probabilistic numerical models to simulate properties of photo-emission and secondary electron emission for single photon input signal and for multi-photon arbitrary-shape bunches.

Now we turn our attention to the behavior of MCPs in external uniform magnetic fields. In this case, the trajectories of the secondary electrons are no longer parabolas, as in homogeneous electric fields, but they perform the Larmor rotation. The complexity of the numerical integration of such trajectories is due to the fact that they are emitted at different angles with respect to the direction of the field, so the Larmor radii for them can differ by orders of magnitude. For a fairly accurate integration it is necessary to coordinate the integration step with the value of the minimum radius, which leads to a sharp increase in the amount of computation. Numerical modelling

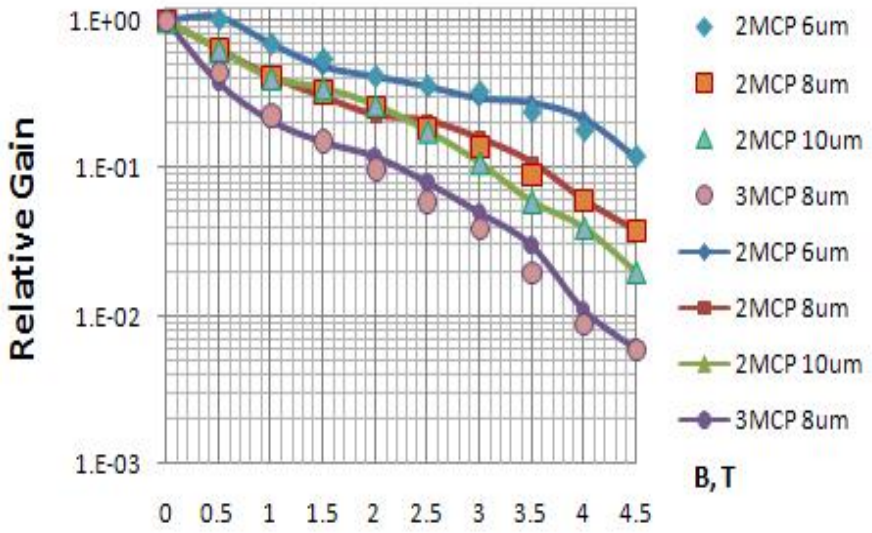


Figure 5: Gain dependence vs. magnetic induction for different MCP assemblies.

for the MCPs in strong magnetic fields up to 5.4 T demonstrates a good agreement for the dependence of gain and transit time spread (TTS) vs. applied voltage and magnetic induction. An example of such a calculation for the considered MCPs is shown in Fig. 5. The gain dependence on the magnetic induction on the MCP orientation relatively the magnetiv field is shown in Fig. 6.

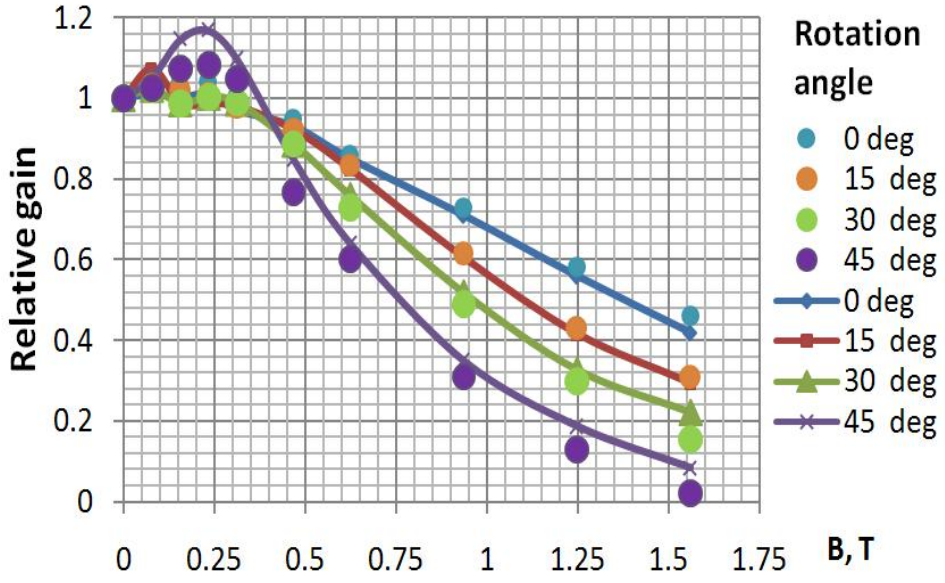


Figure 6: Gain dependence vs. magnetic induction and MCP orientation

Acknowledgments

This work was supported by Russian Science Foundation (Project no.16-12-10221).

References

[1] B.Adams, K.Attenkofer, H.Frish, Z.Insepov, V.Ivanov et al. A Brief Technical History of the Large-Area Picosecond Photodetector (LAPPD) Collaboration. 2016.- 45 pp.

[2] V.Ivanov et al. Numerical simulations of fast photo detectors based on microchannel plates //12th International Conf. "Instrumentation for Colliding Beam Physics" INSTR-2017, February 27 - March 3, 2017, Novosibirsk.

[3] V.Ivanov. The review of mathematical models for three-dimensional problems of electron optics. Int. Conf. on Atomic and Nuclear Physics, July 23-25, 2018, Osaka, Japan.

[4] V.Ivanov. Computer design of microchannel amplifiers. Scholars' Press, 2018.- 220 pp.

[5] V.Ivanov.. Taking into account the fringe fields in microchannel amplifier design//Open Academic J. of Advanced Science and Technology, 1, N1 (2017). P.42-44.

[6] V.Ivanov., I. Turchanocsky. Influence of the fringe fields in microchannel amplifier design// American J. of Modern Physics, 7, 1(2017) pp.31-33.

[7] V. Ivanov, A. Barnyakov, M. Barnyakov, V. Bobrovnikov, I. Ovtin. Numerical simulations of fast photo detectors based on microchannel plates // J of Instrumentation, 12 (2017) P09024.

[8] A. Barnyakov et al. On measurements of photoelectron collection efficiency in MCP based PMTs // J of Instrumentation, 12 (2017) P09036.

[9] V.Ivanov, A.Barnyakov, M.Barnyakov. Calibration procedure in microchannel amplifiers design, NIM A 903 (2018) 170-174.

[10] V.Ivanov, A.Barnyakov, M.Barnyakov, V.Blinov, V.Bobrovsky, I.Ovtin. Methodology of computer design of photodetectors based on microchannel plates. Int. Conf. on Atomic and Nuclear Physics, July 23-25, 2018, Osaka, Japan.

[11] M. Barnyakov, A. Barnyakov, V. Blinov, V. Bobrovnikov, A. Bykov; V. Ivanov, A. Katcin, E. Mamoshkina, I. Ovtin, K. Petrukhin, S. Pivovarov, V. Prisekin, E. Pyata. Development of a picosecond MCP based particle detector, NIM A (2018) 01 057.

[12] P. Hovington et al., CASINO: A new Monte Carlo code in c Language for electron beam interaction. Scanning 19 (1997) 1-14.

[13] A.J. Guest, A computer model of channel multiplier plate performance, Acta Electronica, 14 (1971) pp.79-97.

[14] R.G. Lye, A.J. Dekker, Theory of secondary emission, Phys. Rev. 107 (1957) 977-981.

[15] B.K. Agarwal, Variation of secondary emission with primary electron energy, Proc. Phys. Soc. 71 (1958) 851-852.

[16] M. Ito, H. Kume, K. Oba, Computer analysis of the timing properties in micro channel plate photomultiplier tubes, IEEE Trans. NS-31 (1984) 408-412.

[17] Rodney J., Vaughan M. A New Formula for Secondary Emission Yield // IEEE Transactions on Electric Devices. September 1989. 36, N9. P. 1963-1967.

[18] V.Ivanov, Gain and Time Resolution Simulations in Saturated MCP Pores, 7th Int. Workshop on Ring Imaging Cherenkov Detectors, Cassis, Provence, France.

Ivanov Valentin Yakovlevich — Doctor of Sciences (Physics and Mathematics),
Institute of Computational Technologies SB RAS;
e-mail: vivanov.48@mail.ru.
Received — 30 April 2019.

ОБЗОР И АНАЛИЗ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПРОЦЕССА КАТАЛИТИЧЕСКОГО КРЕКИНГА

Г. И. Исламова[1], И. М. Губайдуллин[1,2]

[1] *Институт нефтехимии и катализа УфИЦ РАН, Уфа, Россия*
[2] *Уфимский государственный нефтяной технический университет, Уфа, Россия*

УДК 51-74+665.64
DOI: 10.24411/9999-016A-2019-10029

В работе представлен обзор восьми математических моделей процесса каталитического крекинга: четырехкомпонентная [1], пятикомпонентная [2], две шестикомпонентные модели [3,4], девятикомпонентная [5], одиннадцатикомпонентная [6], двенадцатикомпонентная [2] и четырнадцатикомпонентная [7] модели. Данные модели отличаются подходом к группировке компонентов, в зависимости от продукта, выбранного целевым, а также в зависимости от необходимой точности модели. Разные модели обуславливаются разным составом сырья, технологическими условиями процесса, различными типами катализаторов и конструкций реакционных аппаратов.

Ключевые слова: Ароматические углеводороды, бензин, газ, газойль, константа равновесия, каталитический крекинг, кокс, математическое моделирование, олефин, парафин.

Процесс каталитического крекинга является одним из важнейших процессов нефтепереработки. В данном процессе получают базовый компонент товарных бензинов, имеющий высокое октановое число и низкое содержание ароматических углеводородов. С каждым годом, ужесточаются требования к товарным бензинам. Например, содержание бензола должно быть не более 1%, а ароматических соединений не более 35%, что ставит новые задачи по углублённому изучению механизма процесса. На основании анализа существующих моделей, экспериментальных лабораторных и заводских данных планируется создание детализированной кинетической модели, которая в дальнейшим будет основой для математического моделирования, модифицирования и проектирования процесса каталитического крекинга.

Наиболее простая модель — четырехкомпонентная (рис. 1) — предложена в 1995 г. группой ученых под руководством Питуайта [1]. В модели процесс каталитического крекинга описан малым количеством компонентов: газ, бензин, вакуумный газойль, коксогенные структуры, что обеспечивает простоту расчета. Такой набор компонентов позволяет учесть основные реакции, протекающие во время процесса: крекинг вакуумного газойля до бензина, газа, а также реакцию коксообразования. Важно учитывать именно побочную реакцию коксообразования, так как коксовые отложения способны дезактивировать катализатор. Учет данной реакции необходим для правильного выбора времени реакции и времени регенерации путем выжига кокса с поверхности катализатора. Однако, существует большое количество и других реакций, происходящих во время процесса. Влияние констант скоростей и энергии активации данных реакций не учитывается в данной модели, это необходимо учитывать при обработке результатов. Следует отметить, что это наиболее простая модель для проведения расчетов процесса.

Реакции уплотнения являются побочными при каталитическом крекинге. В четырехкомпонентной модели учтена только реакция коксообразования, однако помимо кокса существуют и другие побочные тяжелые продукты — каталитические газойли. Группой ученых из Бразилии под руководством Серкьеры [2] была создана пятикомпонентная модель (рис. 2) процесса каталитического крекинга, которая является усовершенствованием четырехкомпонентной модели Питуайта с учетом образования каталитического газойля. Новый компонент — легкое ароматическое масло (LCO), включает в себя все продукты уплотнения, за исключением твердых (кокса). Данная модель становится более сложной для расчетов по сравнению с моделью Питуайта, но учитывает важную реакцию уплотнения. Каталитический газойль составляет до 50% от всех продуктов каталитического крекинга.

ISBN 978-5-901548-42-4

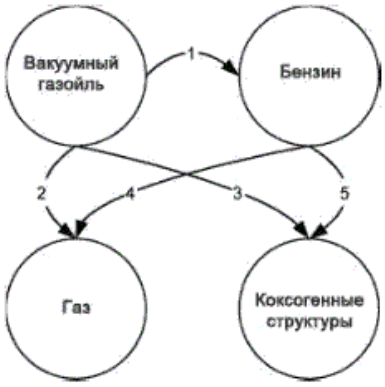


Рис. 1: Схематическое представление четырехкомпонентной модели процесса каталитического крекинга

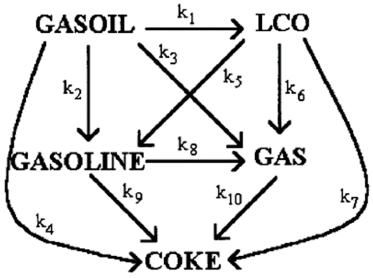


Рис. 2: Схематическое представление пятикомпонентной модели процесса каталитического крекинга

Далее в обзоре представлены шестикомпонентные модели двух групп ученых. В статье [3] описана шестикомпонентая кинетическая модель процесса (рис. 3), которая описывает реакции с участием вакуумного газойля, бензина, легких газообразных продуктов каталитического крекинга (КК), в том числе C_3H_6 и C_2H_4 и двух разновидностей коксогенных структур (КГС1 и КГС2). Как видно, данная модель описывает процесс крекинга, ориентированный на получение именно легких олефинов — этилена и пропилена. Данные вещества являются ценным сырьем для нефтехимии, в частности, сырьем для таких распространенных полимеров, как полиэтилен, полипропилен, полистирол и т.д. Однако, для расчета процесса каталитического крекинга по топливному варианту данная модель может быть недостаточной.

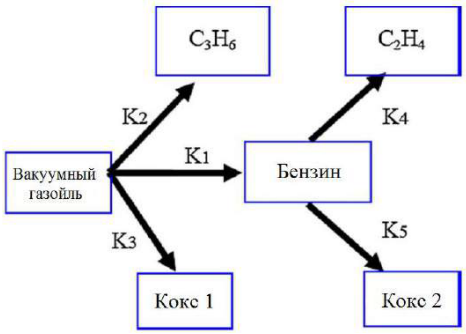


Рис. 3: Схематическое представление шестикомпонентной модели процесса каталитического крекинга с учетом выхода легких олефинов

Для светлых фракций углеводородов (УВ) КК на катализаторе Fe/HZSM-5 была разработана другая шестикомпонентная модель [4]. Схема превращений представлена на рисунке 4. Также отметим, что данная модель особенно хороша для расчета процесса каталитического крекинга с преимущественным получением легких олефинов. Бензин и каталитический газойль объединены в один компонент — жидкие углеводороды, и это делает расчеты каталитического крекинга по топливному варианту по этой модели крайне неточными.

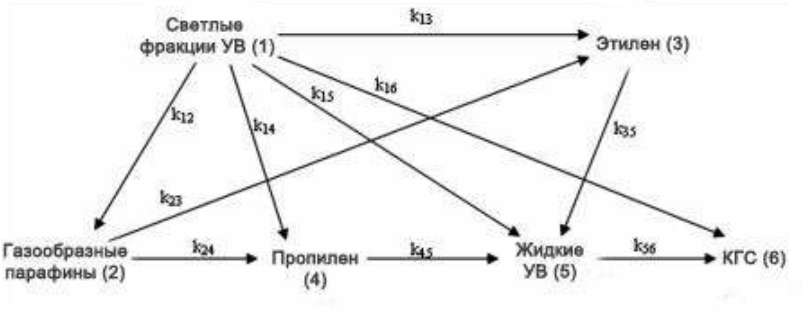


Рис. 4: Схематическое представление шестикомпонентной модели процесса каталитического крекинга

В работе [5] представлена сложная схема превращений с девятью компонентами, учитывающая реакции ароматизации бензина КК. В этой схеме превращений компоненты впервые разделены на группы в зависимости от их строения: н-парафины, изо-олефины, олефины, ароматические углеводороды, коксогенные структуры, легкие углеводороды (C_1–C_4) и водород. В модели учитываются большое количество типов реакций между этими компонентами, например, дегидрирование и циклизации парафинов, изомеризации парафинов и крекинг низкомолекулярных УВ. С целью упрощения, наименее важные реакции были исключены из схемы превращений. Девятикомпонентная модель КК бензинов представлена на рисунке 5 [5]. По сравнению с рассмотренными ранее моделями данная модель является достаточно подробной для расчета процесса каталитического крекинга. Отметим, что изо-парафины и ароматические углеводороды имеют высокое октановое число, и тем самым повышают качество бензинов. Также важно рассчитывать количество ароматических углеводородов и олефинов в получаемом топливе, так как содержание этих компонентов ограничивается российскими и международными нормативными документами.

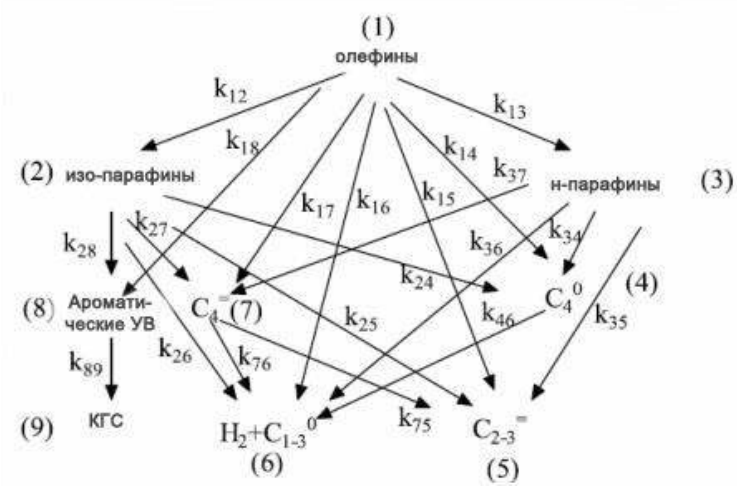


Рис. 5: Схематическое представление девятикомпонентной модели процесса каталитического крекинга

Одиннадцатикомпонентная модель представлена на рисунке 6. По сравнению с девятикомпонентной моделью, здесь группы углеводородов поделены на тяжелые и легкие. Это очень важно, так как это позволит более точно рассчитать качественные показатели получаемых продуктов. В девятикомпонентной модели, напомним, не было подразделения на легкие и тяжелые группы углеводородов, и это затрудняло определение структурного состава бензинов или каталитического газойля, так как не ясно распределение углеводородов по данным продуктам. Кроме этого, следует отметить, что в модели, помимо химических констант, учитываются физические показатели гомологичных углеводородов, например, плотность, теплоемкость, вязкость, теплопроводность, теплота образования и молекулярные массы. Так как модель с одиннадцатью компонентами содержит одну единственную группу, описывающую и кокс, и легкие УВ газы, их характеристики были определены как средневзвешенное этих разновидностей. Подтверждено содержание кокса 30% и 70% легкого УВ газа [6].

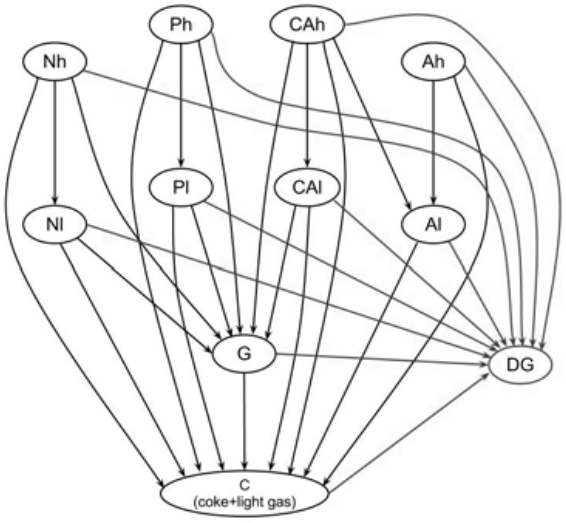


Рис. 6: Схематическое представление одиннадцатикомпонентной модели процесса каталитического крекинга: Ah — Heavy aromatic (тяжелые ароматические углеводороды), Al — Light aromatic (легкие ароматические углеводороды), C — Coke and light gases (кокс и легкие газы), CAh — Heavy aromatic substituents (тяжелые ароматические заместители), CAl — Light aromatic substituents (легкие ароматические заместители), DG — Dry gas (сухой газ), G – Gasoline (легкий бензин), Nh — Heavy naphthenic (тяжелые нафтены), Nl — Light naphthenic (легкие нафтены), Ph — Heavy paraffinic (тяжелые парафины), Pl — Light paraffinic (легкие парафины)

Далее представлена двенадцатикомпонентная модель рассматриваемого нами процесса. Отметим, что она разработана уже упоминавшейся ранее группой ученых под руководством Серкьеры [2]. По сути, пятикомпонентная модель является допустимым упрощением данной двенадцатикомпонентной модели, согласно экспериментальным данным. Однако, можно предположить, что упрощение допустимо не при всех условиях и не при любом виде катализаторов, в связи с чем применять пятикомпонентную модель следует с осторожностью, либо в качестве первого приближения при расчетах. По сравнению с предыдущей моделью здесь газ и кокс разделен на разные компоненты, что дает дополнительную степень свободы при расчетах. Разделение продуктов реакции по массе и структуре позволяет оценивать качество и свойства получаемых продуктов.

14-компонентная кинетическая модель [7], используемая для описания реакций каталитического крекинга, представлена на рисунке 8. Данная модель наиболее полно (из представленных в обзоре) позволяет моделировать и рассчитывать процесс каталитического крекинга. В данной модели считается, что светлые углеводородные фракции состоят из неароматических (парафины, нафтены) и ароматических углеводородов. В процессе каталитического крекинга ароматические углеводороды могут легко превратиться в парафиновые углеводороды. Ароматические углеводороды не могут крекироваться с образованием газообразных продуктов, т.к. реакции дециклизации ароматических углеводородов не протекают в рабочих условиях каталитического крекинга тяжелого сырья [7]. Жидкие углеводороды являются продуктами олигомеризации, ароматизации и конденсации. Поскольку температура процесса каталитического крекинга высока, некоторые газообразные продукты с высокой молекулярной массой могут подвергаться реакциям вторичного крекинга. Кроме того, олефины могут превращаться в жидкие углеводороды путем олигомеризации и ароматизации.

Рассмотренные модели отличаются числом компонентов и необходимых констант равновесия химических превращений между компонентами. Анализ моделей приведен в таблице 1.

Как видно, чем больше компонентов, тем более сложный расчет модели. Кроме этого, существует сложность в виде большого количества необходимых констант равновесия. Данные константы могут быть получены только экспериментальным путем, и отличаются для разных видов катализаторов. Также константы равновесия зависят от температуры и давления процесса (для реакций с участием газовой фазы). Кроме этого, простые модели без учета структурных компонентов не позволяют оценить качество получаемой

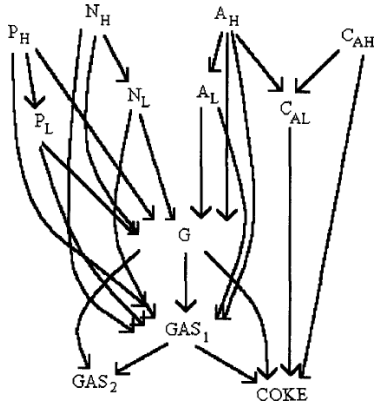


Рис. 7: Схематическое представление двенадцатикомпонентной модели процесса каталитического крекинга: A_H — тяжелые ароматические углеводороды, A_L — легкие ароматические углеводороды, C_{AH} — тяжелые ароматические заместители, C_{AL} — легкие ароматические заместители, N_H — тяжелые нафтены, N_L — легкие нафтены, P_H — тяжелые парафины, P_L — легкие парафины, GAS_1 — первичные газообразные продукты, GAS_2 — вторичные газообразные продукты, G — легкий бензин, COKE — кокс

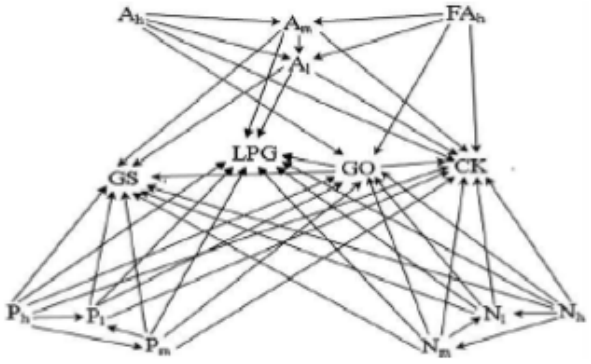


Рис. 8: Схематическое представление четырнадцатикомпонентной модели процесса каталитического крекинга: A_H — тяжелые ароматические углеводороды, A_M — средние ароматические углеводороды, A_L — легкие ароматические углеводороды, F_{AH} — тяжелые ароматические углеводороды в смолах и асфальтенах, N_H — тяжелые нафтены, N_M — средние нафтены, N_L — легкие нафтены, P_H — тяжелые парафины, P_M — средние парафины, P_L — легкие парафины, GO — бензин, LPG — сжиженный нефтяной газ, DG — сухой газ, CK — коксогенные структуры

Таблица 1: Сравнение моделей

Количество компонентов	Количество констант равновесия	Целевой компонент модели	Возможность оценить качество бензина
Четырехкомпонентная модель	5	Бензин	Отсутствует
Пятикомпонентная модель	10	Бензин	Отсутствует
Шестикомпонентная модель (по ист. [3])	5	Этилен, пропилен	Отсутствует
Шестикомпонентная модель (по ист. [4])	10	Этилен, пропилен	Отсутствует
Девятикомпонентная модель	20	Бензин	Существует
Одиннадцатикомпонентная модель	29	Бензин	Существует
Двенадцатикомпонентная модель	23	Бензин	Существует
Четырнадцатикомпонентная модель	46	Бензин	Существует

продукции, что является существенным недостатком. Но, для первичного приближения, применение этих моделей может быть оправдано и целесообразно.

Таким образом, можно утверждать, что в настоящее время существует большое количество различных подходов к моделированию процесса каталитического крекинга. Разные модели обуславливаются составом сырья, технологическими условиями процесса, различными типами катализаторов и конструкций реакционных аппаратов. В настоящее время является актуальной проблема моделирования процесса каталитического крекинга.

Список литературы

[1] Pitault I., Forissier M., Bernard J.R., Determination of kinetics constants of catalytic cracking by modeling microactivity test. // Can. J. Chem. Eng. — 1995. — Vol.73. — P.498–504.

[2] H. S. Cerqueira Mathematical modeling and simulation of catalytic cracking of gasoil in a fixed bed: Coke formation / E.C. Biscaia Jr., E. F. Sousa-Agular // Applied Catalysis A: General 164 (1997) 35–45

[3] Behjata Y. CFD analysis of hydrodynamic, heat transfer and reaction of three phase riser reactor / Shahhosseinia S. // Chemical engineering research and design. — №89. — 2011. — P.978–989.

[4] Sadeghzadeh J., Farshi A., Forsat K., A mathematical modeling of the riser reactor in industrial FCC unit // Petrol.Coal. — 2008. — Vol.50 — P.15–24.

[5] Kang X. An Introduction to the lump kinetics model and reaction mechanism of FCC gasoline / Guo X. // Energy Sources. — Part A, 35:1921–1928. — 2013. — P.343–358

[6] Barbosa A.C. Three dimensional simulation of catalytic cracking reactions in an industrial scale riser using a 11-lump kinetic / Lopes G.C. // Chemical engineering transactions. — Vol. 32. — 2013. — P.637 — 642.

[7] Zhang, Numerical simulation on catalytic cracking reaction in two-stage riser reactors // China University of Petroleum. — Beijing, China. — 2005. — P.129–153.

Исламова Гульшат Ильнуровна — аспирант Института нефтехимии и катализа УФИЦ РАН;
e-mail: gulshat.islamova.2017@mail.ru;
Губайдуллин Ирек Марсович — д.ф.-м.н., с.н.с., зав. лабораторией математической химии Института нефтехимии и катализа УФИЦ РАН;
Уфимский государственный нефтяной технический университет;
e-mail: irekmars@mail.ru.
Дата поступления — 30 мая 2019 г.

DIRECT PROBLEM FOR THE HYPERBOLIC FIRST-ORDER SYSTEM OF ACOUSTIC WAVES PROPAGATION

S. I. Kabanikhin[1,2,3], D. V. Klyuchinskiy[1,2], I. M. Kulikov[1], N.S. Novikov[1,2], M.A. Shishlenin[1,2,3]

[1] *Institute of Computational Mathematics and Mathematical Geophysics SB RAS, 630090, Novosibirsk, Russia*
[2] *Novosibirsk State Univerity, 630090, Novosibirsk, Russia*
[3] *Sobolev Institute of Mathematics, 630090, Novosibirsk, Russia*

UDC 517.968
DOI: 10.24411/9999-016A-2019-10030

We investigate the mathematical modeling of the 2D acoustic waves propagation in homogenious and heterogenious areas. The hyperbolic first order system of partial differential equations is considered and solved by the Godunov method of the first order of approximation and MUSCL-Hancock method of the second order of approximation.
Keywords: acoustics, tomography, first-order hyperbolic system, direct problem, inverse problem, Godunov method, MUSCL-Hancock method

Introduction

In this paper we consider the direct problem for the hyperbolic system of the two dimensional acoustic wave propagation. This problem is related to ultrasound tomography [4, 5], aimed at the detection of inclusions in the soft human tissue, which is connected with the development of methods and instruments for early breast cancer detection. An important issue is the construction of effective algorithms for solving the inverse problem, and, therefore, direct problem due to their tightly connected efficiency.

We use the hyperbolic first-order system to describe the propagation of the acoustic waves, written in the form of conservation laws of continuum mechanics. This allows to propose more realistic model from the physical point of view. We consider two numerical algorithm for solving direct problems: the classical Godunov scheme of the first-order approximation [6, 7, 10] and Monotone Upstream–centred Scheme for Conservation Laws of the second-order approximation investigated by Leer [8] and modified by Hancock [9]. MUSCL-Hancock scheme is based on Godunov approach and the main difference from last one is that it uses piece-wise linear approximation of values inside grid cells instead of piece-wise constant approximation by Godunov method. The numerical methods, based on the Godunov approach, allows to find the balance between mathematical modeling of physical process and the effective numerical realization.

The work has been supported by RSCF under grant 19-11-00154 "Creation of conceptionally new mathematical models of acoustic tomography in medicine. Numerical methods, high-performance computing and software systems".

ISBN 978-5-901548-42-4

1 Direct problem

1.1 Main equations

The propagation of the acoustic waves in the 2D medium is described by the following system of integral conservation laws

$$\oint \rho u dxdy + pdydt = 0, \tag{1}$$

$$\oint \rho v dxdy + pdxdt = 0, \tag{2}$$

$$\oint pdxdy + \rho c^2 \left(udydt + vdxdt\right) = 0. \tag{3}$$

Here $u = u(x, y, t), v = v(x, y, t)$ are components of velocity vector, with respect to x and y respectively, $p = p(x, y, t)$ is the acoustic pressure. The parameters of the system describes the properties of the medium: $\rho(x, y)$ denotes the density of the medium, and $c(x, y)$ is the speed of the wave propagation.

Solution $u(x, y, t)$, $v(x, y, t)$ and $p(x, y, t)$ of integral equations (1)-(3), in general, can be both smooth and discontinuous types. In our problem we will consider only smooth acoustic waves.

1.2 Model of acoustic wave propagation from the source

Let us consider the following direct problem for acoustic wave propagation from the source

$$\frac{\partial u}{\partial t} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0, \quad (x, y) \in \Omega, 0 < t \leq T \tag{4}$$

$$\frac{\partial v}{\partial t} + \frac{1}{\rho}\frac{\partial p}{\partial y} = 0, \quad (x, y) \in \Omega, \tag{5}$$

$$\frac{\partial p}{\partial t} + \rho c^2 \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = \theta_\Omega(x, y) I(t), \quad (x, y) \in \Omega; \tag{6}$$

$$u, v, p|_{(x,y)\in\partial\Omega} = 0. \tag{7}$$

$$u, v, p|_{t=0} = 0 \tag{8}$$

Here $\Omega = (x, y) \in [0, L] \times [0, L]$. Function θ_Ω describes the location of the source, which has the form of Ricker wavelet:

$$I(t) = \left(1 - 2\left(\pi\nu_0\left(t - \frac{1}{\nu_0}\right)\right)^2\right) e^{-\pi\nu_0\left(t - \frac{1}{\nu_0}\right)} \tag{9}$$

For solving the direct problem (4)—(8), in the next section we consider numerical method, proposed by S.K. Godunov in [7]. The main element of this method is the solving of Riemann problem, which consists of an initial value problem composed of a conservation equation together with piece-wise constant data having a single discontinuity. This problem can be solve explicitly or by using some approximation [10].

1.3 Godunov finite-difference scheme

Let us introduce the mesh in $\mathbb{R}^2$, provided by the lines $x = x_i$ and $y = y_j$, and denote by h_x, h_y parameters of the mesh. We will also denote by pair of indices $(i + 1/2, j + 1/2)$ the grid cell $I_{i+1/2,j+1/2}$, corresponded to points (x_i, y_j), (x_i, y_{j+1}), (x_{i+1}, y_j) ,(x_{i+1}, y_{j+1}). We suppose, that values of functions $u(x, y, t)$, $v(x, y, t)$, $p(x, y, t)$, which describes the state of the medium, are constant in each cell, and denoted this values as $u_{i+1/2,j+1/2}$, $v_{i+1/2,j+1/2}$, $p_{i+1/2,j+1/2}$. Now let us consider the ratio (1) for the cell $I_{i+1/2,j+1/2}$:

$$\rho\left(u^{i+1/2,j+1/2} - u_{i+1/2,j+1/2}\right) h_x h_y + \int_{t_0}^{t_0+\tau}\int_{y_j}^{y_{j+1}} \left[p\left(x_{i+1}, y, t\right) - p\left(x_i, y, t\right)\right] dydt = 0, \tag{10}$$

The upper index indicates the values of u, v, p for the moment $t_0 + \tau$ (τ is the time step), and the lower index corresponds to the values of functions at the moment t_0. We approximate the functions $p(x_i, y, t), p(x_{i+1}, y, t)$,

corresponded to values at the boundary of the cell, by some constant values, which we denote as $P_{i,j+1/2}$ and $P_{i+1,j+1/2}$ respectively. Thus, we can obtain:

$$\rho\left(u^{i+1/2,j+1/2}-u_{i+1/2,j+1/2}\right)h_xh_y+\tau h_y\left(P_{i+1,j+1/2}-P_{i,j+1/2}\right)=0. \tag{11}$$

We use the conservation laws (2), (3) to obtain two more ratios:

$$\rho\left(v^{i+1/2,j+1/2}-v_{i+1/2,j+1/2}\right)h_xh_y+\tau h_x\left(P_{i+1/2,j+1}-P_{i+1/2,j}\right)=0, \tag{12}$$

$$\left(p^{i+1/2,j+1/2}-p_{i+1/2,j+1/2}\right)h_xh_y+\tau\rho c^2h_y\left(U_{i+1,j+1/2}-U_{i,j+1/2}\right)+$$
$$+\tau\rho c^2h_x\left(V_{i+1/2,j+1}-V_{i+1/2,j}\right)=\tau\theta_{\Omega_{i+1/2,j+1/2}}I_k. \tag{13}$$

where I_k - approximation of the source on time step t_0. We use the solution of the problem of "decay of discontinuity" to obtain values U, P at the boundary $x=x_i$

$$U=U_{i,j+1/2}=\frac{u_{i-1/2,j+1/2}+u_{i+1/2,j+1/2}}{2}-\frac{1}{\rho c}\frac{p_{i+1/2,j+1/2}-p_{i-1/2,j+1/2}}{2} \tag{14}$$

$$P=P_{i,j+1/2}=\frac{p_{i-1/2,j+1/2}+p_{i+1/2,j+1/2}}{2}-\rho c\frac{u_{i+1/2,j+1/2}-u_{i-1/2,j+1/2}}{2} \tag{15}$$

The same ratios for V and P can be obtained at the boundary $y=y_j$

$$V=V_{i+1/2,j}=\frac{v_{i+1/2,j-1/2}+v_{i+1/2,j+1/2}}{2}-\frac{1}{\rho c}\frac{p_{i+1/2,j+1/2}-p_{i+1/2,j-1/2}}{2} \tag{16}$$

$$P=P_{i+1/2,j}=\frac{p_{i+1/2,j-1/2}+p_{i+1/2,j+1/2}}{2}-\rho c\frac{v_{i+1/2,j+1/2}-v_{i+1/2,j-1/2}}{2} \tag{17}$$

Therefore, according to (11), (12), (13), we can write the final ratios for computing the state of the medium for "upper" layer $t=t_0+\tau$:

$$u^{i+1/2,j+1/2}=u_{i+1/2,j+1/2}-\frac{\tau}{\rho h_x}\left(P_{i+1,j+1/2}-P_{i,j+1/2}\right), \tag{18}$$

$$v^{i+1/2,j+1/2}=v_{i+1/2,j+1/2}-\frac{\tau}{\rho h_y}\left(P_{i+1/2,j+1}-P_{i+1/2,j}\right), \tag{19}$$

$$p^{i+1/2,j+1/2}=p_{i+1/2,j+1/2}-\frac{\tau}{h_x}\rho c^2\left(U_{i+1,j+1/2}-U_{i,j+1/2}\right)-$$
$$-\frac{\tau}{h_y}\rho c^2\left(V_{i+1/2,j+1}-V_{i+1/2,j}\right)-\tau\theta_{\Omega_{i+1/2,j+1/2}}I_k. \tag{20}$$

As was mentioned previously, we solve the problem in the domain $\Omega=(x,y)\in[0,L]\times[0,L]$. We would like to have somehow modeled non-reflecting conditions near the boundaries $x=0, y=0, x=L, y=0$. To model this situation we do not set exact boundary conditions at "big" values $P_{0,j+1/2}, U_{0,j+1/2}$, $P_{N,j+1/2}, U_{N,j+1/2}$, $P_{i+1/2,0}, U_{i+1/2,0}$, $P_{i+1/2,N}, U_{i+1/2,N}$. Instead of this we just rewrite "small" values in the outside cells of the area in the next way:

$$u_{1/2,j+1/2}=u_{3/2,j+1/2},\quad u_{N-1/2,j+1/2}=u_{N-3/2,j+1/2}$$
$$v_{1/2,j+1/2}=v_{3/2,j+1/2},\quad v_{N-1/2,j+1/2}=v_{N-3/2,j+1/2}$$
$$p_{1/2,j+1/2}=p_{3/2,j+1/2},\quad p_{N-1/2,j+1/2}=p_{N-3/2,j+1/2}$$

for boundaries $x=0$ and $x=L$, and the equivalent formulas for boundaries $y=0$ and $y=L$.

These conditions guaranties us an attenuation of waves near the domain boundaries and physically means, that waves just leave the considered area.

Here values U, V, P are defined by (14)—(17). The stability of the scheme can be achieved by using the CFL condition [7, 10]:

$$\tau\left(\frac{C_0}{h_x}+\frac{C_0}{h_y}\right)<1$$

The constant C_0 here is the upper-bound estimate for the speed of wave propagation in the computational domain.

1.4 MUSCL-Hancock finite-difference scheme

In this section for solving (4)—(8) we use the MUSCL-Hancock finite-difference scheme of the second order. This scheme has a classical "predictor-corrector" type. The idea of the method is to use piece-wise linear values of variables in cells instead piece-wise constant values as in classical Godunov method.

For simplicity, let us represent our system in conservative form in 1D case without right-hand side

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0, \tag{21}$$

Here is U - is a vector of three components $U = (u, v, p)$, $F(U)$ - nonlinear function of U. The MUSCL-Hancock approach achieves a second order extension of the Godunov first order upwind method if the intercell flux F_i is computed according to the following steps:

Step 1. Predictor step. Data reconstruction.

In the data reconstruction step, data cell average values $U^n_{i+1/2}$ are locally replaced by piece–wise linear functions in each cell $I_{i+1/2} = [x_i, x_{i+1}]$, namely

$$U_{i+1/2}(x) = U^n_{i+1/2} + \frac{(x - x_{i+1/2})}{\Delta x}\Delta_i$$

$\Delta_{i+1/2}$ is a suitably chosen slope vector of $U_{i+1/2}(x)$ in cell $I_{i+1/2}$. The extreme points $x = 0$ and $x = \Delta x$, in local coordinates, correspond to the intercell boundaries x_i and x_{i+1} in global coordinates. The values $U_{i+1/2}(x)$ at the extreme points are

$$U^L_{i+1/2} = U^n_{i+1/2} - \frac{1}{2}\Delta_{i+1/2}$$

$$U^R_{i+1/2} = U^n_{i+1/2} + \frac{1}{2}\Delta_{i+1/2}$$

where super indexes L and R denotes left and right extreme point, respectively. These values, also, are usually called boundary extrapolated values. Note, that U and $\Delta_{i+1/2}$ are vectors of three components of the Euler equations and thus there are six extrapolated values.

Step 2. Corrector step. Evolution.

For each cell $I_{i+1/2}$, the boundary extrapolated values $U^L_{i+1/2}$ and $U^R_{i+1/2}$ are corrected (evolved) by a time $\frac{\Delta t}{2}$, according to

$$\overline{U}^L_{i+1/2} = U^L_{i+1/2} + \frac{1}{2}\frac{\Delta t}{\Delta x}\left[F(U^L_{i+1/2}) - F(U^R_{i+1/2})\right]$$

$$\overline{U}^R_{i+1/2} = U^R_{i+1/2} + \frac{1}{2}\frac{\Delta t}{\Delta x}\left[F(U^L_{i+1/2}) - F(U^R_{i+1/2})\right]$$

Note that this evolution step is entirely contained in each cell $I_{i+1/2}$, as the intercell fluxes are evaluated at the boundary extrapolated values of each cell. At each intercell boundary i there are two fluxes, namely $F(U^R_{i-1/2})$ and $F(U^L_{i+1/2})$, which are in general distinct. This does not really affect the conservative character of the overall method, as this step is only an intermediate step; the intercell flux F_i to be used is yet to be evaluated.

Step 3. The Riemann Problem.

To compute the intercell flux F_i one now solves the conventional Riemann problem with data

$$U_L \equiv \overline{U}^R_{i-1/2}, \quad U_R \equiv \overline{U}^L_{i+1/2}$$

to obtain the similarity solution $U_i(x/t)$. The intercell flux F_i is now computed in exactly the same way as in the Godunov first order upwind method from previous section, namely

$$F_i = F(U_i(0))$$

Here $U_i(0)$ denotes the value of $U_i(x/t)$ at $x/t = 0$, i.e. the value of $U_i(x/t)$ along the t-axis. There the ten possible wave patterns in the solution of the Riemann problem that must be taken into account when computing the Godunov flux.

A possible choice for the slope vector $\Delta_{i+1/2}$ is

$$\Delta_{i+1/2} = \frac{1}{2}(1 + \omega)\Delta_i + \frac{1}{2}(1 - \omega)\Delta_{i+1}$$

where

$$\Delta_i = U^n_{i+1/2} - U^n_{i-1/2}, \quad \Delta_{i+1} = U^n_{i+3/2} - U^n_{i+1/2}$$

and $\omega \in [-1, 1]$.

Also, in the cases of high values of gradient (discontinuous solutions) a technology of slope limiter can be used. There are a lot of different limiter in literature, we present only two of them.

$$\Delta_{i+1/2} = \max(0, \min(\beta\Delta_i, \Delta_{i+1}), \min(\Delta_i, \beta\Delta_{i+1})), \quad if\ \Delta_{i+1} > 0$$
$$\Delta_{i+1/2} = \min(0, \max(\beta\Delta_i, \Delta_{i+1}), \max(\Delta_i, \beta\Delta_{i+1})), \quad if\ \Delta_{i+1} < 0$$

Here $\beta = 1$ corresponds to $MINMOD$ limiter, and $\beta = 2$ to $SUPERBEE$ limiter. In our problem there are no discontinuous solutions, that's why the choose of special limiter has a very little influence on the final solution.

Step 4. Conservation laws.

Computed "big" values of flux F_i are used in finite-difference conservation laws to obtain "small" values of vector $U = (u, v, p)$ inside cells at time step $t + \tau_0$

$$U^i = U_i - \frac{\tau}{h}(F_{i+1} - F_i)$$

2 Numerical results

In this section we present results of solving direct problem on the grid 400×400 points. Physical size of area $[0 : 0.3] \times [0 : 0.3]$. CFL number was taken equal to 0.5. We consider heterogenious area, which simulates some object inside a water. This object has two insertions with higher density and velocity than object has. Fig. 1 shows the distribution of density and velocity in this area.

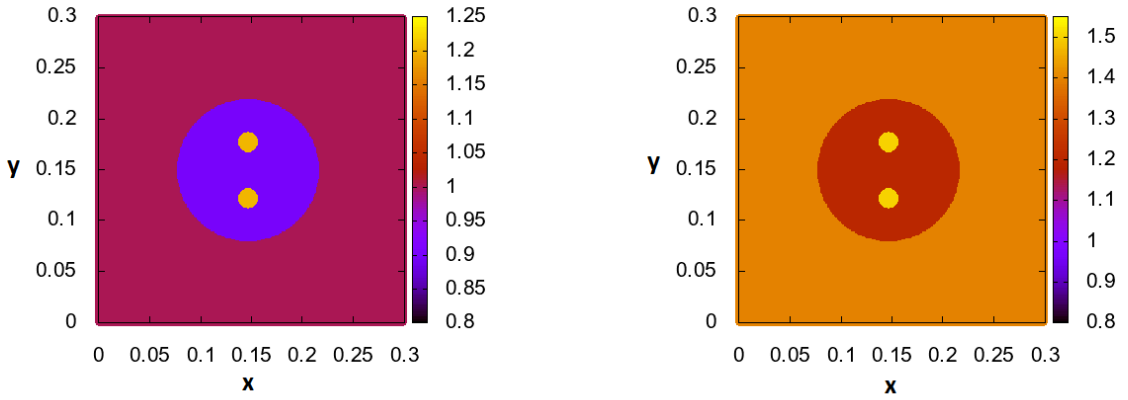


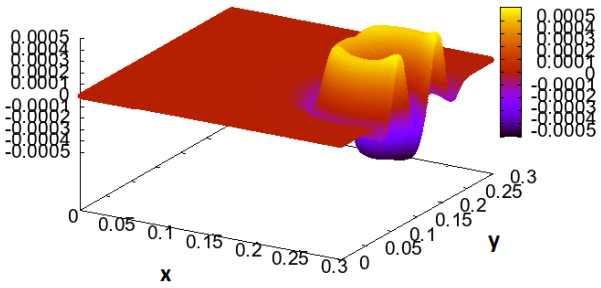
Figure 1: Density and velocity distribution in the area.

Fig. 2 shows the modeling results of pressure for different time points.

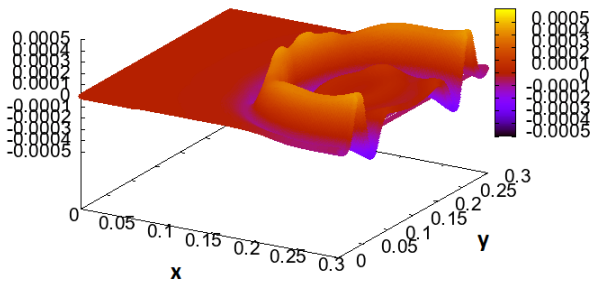
Conclusion

We have developed a direct solver for the hyperbolic first-order system of acoustics equations via Godunov method and Muscl-Hancock methods (with two types of limiters). Numerical experiments show wave propagation from the source without oscillations and quite similar results for these methods due to smooth type of wave. Each finite-difference scheme has been developed via OpenMP multi-threading technology. The results of speedup is good enough, but its description is beyond the scope of this article.

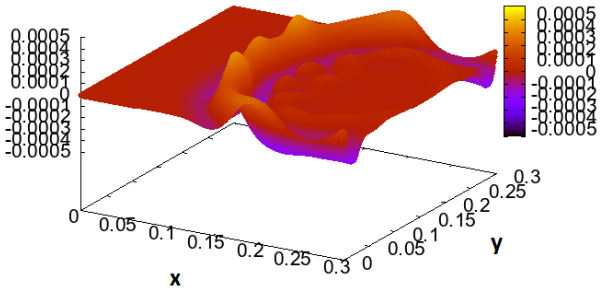
We have considered the mathematical model of acoustic tomograph in the form of the system of hyperbolic equations of the first order because in this form written conservation laws of continuum mechanics. It allows us to control the preservation of the basic invariants during the solution of direct and inverse problems. This is important for solving unstable problems, as the conservation laws of the main invariants are the only criterion



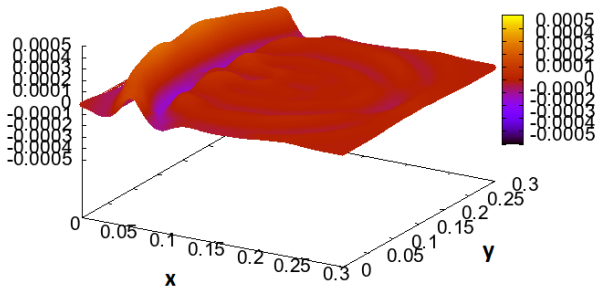
$t = 0.085$



$t = 0.135$



$t = 0.195$



$t = 0.285$

Figure 2: Acoustic wave. Pressure at different time points.

of the well-posedness of the solution. Using the problem formulation in the hyperbolic system we control the numerical solution and we guarantee that numerical solution is close to the physical solution of the process.

The method of the solution of the direct problem for the acoustic tomograph which is based on the the two dimensional hyperbolic system of acoustic equations is developed in [15]. The approach for direct problem solution, proposed in this article, will be used to solve coefficient inverse problem for acoustics wave propagation.

References

[1] Kang H., Tanuma K., Inverse problems for scalar conservation laws, Inverse Problems, **21**, P. 1047–1059 (2005)

[2] Berres S, Bürger R., Coronel A., Sepúlveda M., Numerical identification of parameters for a strongly degenerate convection-diffusion problem modelling centrifugation of flocculated suspensions, Applied Numerical Mathematics, **52**, P. 311–337 (2005)

[3] Holden H., Priuli F.S., Risebro N.H., On an inverse problem for scalar conservation laws, Inverse Problems, **30**, No. 035015, 35pp (2014)

[4] Burov V.A., Vecherin S.N., Morozov C.M., Rumyantseva O.D., Modeling of the exact solution of the inverse scattering problem by functional methods, Acoustic Journal, **56**, 4, p. 516-536 (2010) (in Russian).

[5] Burov V.A., Voloshinov V.B., Dmitriev K.V., Polikarpova N.V., Acoustic waves in metamaterials, crystals, and anomalously refracting structures, Advances in Physical Sciences, **54**, 1165–1170 (2011)

[6] Godunov S.K., Differential method for numerical computation of noncontinuous solutions of hydrodynamics equations, Matematicheskiy Sbornik, **47**, 3, p. 271-306 (1959) (in Russian)

[7] Godunov S.K., Zabrodin A.V., Ivanov M.Y., Kraikov A.N., Prokopov G.P.: Numerical solution for multidimensional problems of gas mechanics, Moscow, Nauka, 1976.

[8] B. van Leer. MUSCL, A New Approach to Numerical Gas Dynamics. In Computing in Plasma Physics and Astrophysics, Max–Planck–Institute Plasma Physik, Garchung, Germany, April 1976.

[9] B. van Leer. On the Relation Between the Upwind–Differencing Schemes of Godunov, Enguist-Osher and Roe. SIAM J. Sci. Stat. Comput., 5(1):1–20, 1985.

[10] Godunov S.K., Kulikov I.M., Computation of Discontinuous Solutions of Fluid Dynamics Equations with Entropy Nondecrease Guarantee, Computational Mathematics and Mathematical Physics **54**, 6, p. 1008–1021 (2014)

[11] S.I. Kabanikhin, D.B. Nurseitov, M.A. Shishlenin, B.B. Sholpanbaev, Inverse problems for the ground penetrating radar, Journal of Inverse and Ill-Posed Problems, **21**, No. 6. P. 885–892 (2013)

[12] S.I. Kabanikhin, M.A. Shishlenin, Quasi-solution in inverse coefficient problems, Journal Inverse and Ill-Posed Problems, **16**, No. 7. P. 705–713 (2008)

[13] S.I. Kabanikhin, M.A. Shishlenin, Numerical Methods for Solving Inverse Hyperbolic Problems, Computational Methods for Applied Inverse Problems (Ed. by Y. Wang, A. Yagola, and C. Yang). Berlin–Boston–Beijing: De Gruyter and Higher Education Press, P. 369–393 (2012)

[14] S.I. Kabanikhin, M.A. Shishlenin, Acoustic sounding methods of linearization and conversion of the wave field , Siberian Electronic Mathematical Reports, **7**, P. C.199–C.206 (2010) Mathematical Notes. **28** (4).P. 723—727 (1981)

[15] Kulikov I.M., Novikov N.S., Shishlenin M.A. Mathematical modelling of ultrasound wave propagation in two dimensional medium: direct and inverse problem. Siberian Electronic Mathematical Reports, **12**, Pp. C.219-C.228 (2015)

Kabanikhin Sergey — Dr. Sc., senior fellow of the Institute of Computational Mathematics and Mathematical Geophysics SB RAS; e-mail: ksi52@mail.ru;
Klyuchinskiy Dmitriy — student of the Novosibirsk State University; e-mail: dmitriy_klyuchinskiy@mail.ru;
Kulikov Igor — Dr. Sc., senior fellow of the Institute of Computational Mathematics and Mathematical Geophysics SB RAS; e-mail: kulikov@ssd.sscc.ru;
Novikov Nikita — PhD, senior fellow of the Institute of Computational Mathematics and Mathematical Geophysics SB RASy; e-mail: novikov-1989@yandex.ru;
Shishlenin Maxim — Dr. Sc., Head of the laboratory of the Institute of Computational Mathematics and Mathematical Geophysics SB RAS; Novosibirsk State University; e-mail: mshishlenin@ngs.ru.
Received — April 30, 2019

МЕТОД ОПОРНЫХ ВЕКТОРОВ РЕШЕНИЯ ЗАДАЧ ОБ ИСТОЧНИКЕ ДЛЯ ПАРАБОЛИЧЕСКИХ УРАВНЕНИЙ

С. И. Кабанихин[1,2], О. И. Криворотько[1,2], Е. А. Ануфриенко[2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10031

В работе приведен и проанализирован алгоритм метода опорных векторов в случае гауссова ядра в разложении неизвестных функций в задаче об источнике для параболического уравнения с граничными условиями типа Неймана по дополнительной информации о решении прямой задачи в фиксированный момент времени. Решение задачи об источнике неединственно. Задача об источнике с граничными условиями типа Неймана сведена к задаче с граничными условиями типа Дирихле, для которой был разработан алгоритм регуляризации на основе введения априорной информации о решении. Приведены и проанализированы результаты тестовых расчетов.

Ключевые слова: метод опорных векторов, задача об источнике, параболические уравнения, обратная задача, оптимизация, регрессия, регуляризация, машинное обучение, социальные процессы.

Введение

Задачи определения источников по некоторой дополнительной информации о процессах, описываемые параболическими уравнениями, играют важную роль в математической физике, биологии, медицине, социологии, распознавании образов и т.д. Как правило, такие задачи являются некорректными, а именно, их решение может быть неустойчивым и/или неединственным. В качестве регуляризации решение задачи об источнике сводится к решению задачи минимизации функционала А.Н. Тихонова. В работе предложен метод опорных векторов (*Support Vector Machine, SVM*) для решении задачи регрессии в смысле наименьших квадратов [1], являющийся одним из эффективных методов машинного обучения [2–4]. SVM состоит в сведении многоэкстремальной задачи к задаче квадратичного программирования, имеющей единственное решение. При соответствующем выборе функции ядра SVM является реализацией метода искуственных нейронных сетей с тем лишь преимуществом, что метод SVM автоматически определяет число нейронов скрытого слоя, равное числу опорных векторов [5, 6].

В работе [3] был разработан алгоритм метода опорных векторов для задачи определения источника в уравнении конвекции-диффузии в случае граничных условий Дирихле. В данной работе приведен и проанализирован алгоритм метода опорных векторов в случае гауссова ядра в разложении неизвестных функций источника и решения тестовой начально-краевой параболической задачи с граничными условиями типа Неймана. В этом случае показано, что метод опорных векторов с точностью до постоянной определеяет функцию источника, поэтому применяется численная регуляризация для получения единственного решения.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 18-31-20019).

ISBN 978-5-901548-42-4

1 Математическая постановка задачи об источнике

Задача определения источника $H(x)$ в начально-краевой задаче для параболического уравнения

$$\begin{cases} u_t = du_{xx} + H(x)S(x,t) + F(x,t), & (x,t) \in D; \\ u(x,t_0) = f(x), & l \le x \le L; \\ u_x(l,t) = g_1(t),\ u_x(L,t) = g_2(t), & t_0 \le t \le T; \\ u(x,t^*) = \omega(x), & t_0 \le t^* \le T, l \le x \le L. \end{cases} \tag{1}$$

с граничными условиями типа Неймана заключается в определении функций источника $H(x)$ и решения начально-краевой задачи $u(x,t)$ в области $D = (l,L) \times (t_0,T)$ по заданным функциям $f(x)$, $g_1(t)$, $g_2(t)$ и дополнительным измерениям $\omega(x)$. Граничные условия в задаче (1) описывают непрохождение потока через границу области и часто используются в постановках задач распространения информации в онлайн социальных сетях [7], ценных бумаг в экономике [8] и т.п. Отметим, что дополнительную информацию (последнее уравнение) можно задавать измерениями в фиксированной точке пространства.

Введем новые функции $U(x,t) = u(x,t) + \dfrac{(x-L)^2}{2(L-l)}g_1(t) - \dfrac{(x-l)^2}{2(L-l)}g_2(t)$, $\tilde{f}(x) = f(x) + \dfrac{(x-L)^2}{2(L-l)}g_1(t_0) - \dfrac{(x-l)^2}{2(L-l)}g_2(t_0)$, $\tilde{F}(x,t) = F(x,t) + \dfrac{(x-L)^2}{2(L-l)}g_1'(t) - \dfrac{(x-l)^2}{2(L-l)}g_2'(t) + d\dfrac{g_2(t)-g_1(t)}{L-l}$, $\tilde{\omega}(x) = \omega(x) + \dfrac{(x-L)^2}{2(L-l)}g_1(t^*) - \dfrac{(x-l)^2}{2(L-l)}g_2(t^*)$ и перепишем задачу (1) с однородными граничными условиями

$$\begin{cases} U_t = dU_{xx} + H(x)S(x,t) + \tilde{F}(x,t), & (x,t) \in D; \\ U(x,t_0) = \tilde{f}(x), & l \le x \le L; \\ U_x(l,t) = U_x(L,t) = 0, & t_0 \le t \le T; \\ U(x,t^*) = \tilde{\omega}(x), & t_0 \le t^* \le T, l \le x \le L. \end{cases} \tag{2}$$

Введем новые функции $v(x,t) = U_x(x,t)$, $g(x) = \tilde{f}'(x)$, $w(x) = \tilde{\omega}'(x)$ и перепишем задачу об источнике (2) в следующем виде:

$$\begin{cases} v_t = dv_{xx} + H'(x)S(x,t) + H(x)S_x(x,t) + \tilde{F}_x(x,t), & (x,t) \in D; \\ v(x,t_0) = g(x), & l \le x \le L; \\ v(l,t) = v(L,t) = 0, & t_0 \le t \le T; \\ v(x,t^*) = w(x), & t_0 \le t^* \le T, l \le x \le L. \end{cases} \tag{3}$$

Решение задачи об источнике (3) может быть сведено к решению задачи минимизации функционала А.Н. Тихонова с параметром регуляризации τ [9]:

$$J(H) = \|v(\cdot,t^*;H) - w(\cdot)\|^2_{L_2} + \tau\|H\|^2_{L_2}, \quad \tau \ge 0. \tag{4}$$

2 Метод опорных векторов в задачах регрессии

Рассмотрим задачу (3) с непрерывной функцией $g(x)$ на ограниченной области D, удовлетворяющей граничным условиям, т. е. $g(l) = g(L) = 0$. Метод состоит из трёх ключевых шагов:
1) Задание точек расчетной области.
2) Построение приближенных решений.
3) Вычисление параметров регрессии.

Для удобства введём обозначения $V = V(x,t)$, $\bar{V} = (x,t^*)$. Внутренние точки сеточной области обозначим через $P_{in} = \{V_i = (x_i,t_i)\}_1^M$, точки пространства — $\bar{X} = \{\bar{x}_j | \bar{x}_j = j\bar{h}\}_1^{N_1}$, точки пространства на t^*-временном слое — $P_{ad} = \{\bar{V}_i = (x_i,t^*), x_i = i\tilde{h}\}_1^{N_2}$.

Будем искать неизвестные функции $H(x)$ и $v(x,t)$ в виде

$$\bar{v}(V) = A(V) + B(V)\left(\sum_{j=1}^{M}\alpha_j K(V,V_j) + b\right), \qquad \bar{H}(x) = \sum_{j=1}^{N_1}\bar{\alpha}_j k(x,\bar{x}_j) + \bar{b}, \tag{5}$$

где K и k — двумерное и одномерное ядра гауссового типа:

$$K(V,V_i)=\exp\left\{-\frac{1}{2\sigma_K^2}||V-V_i||_2^2\right\},\quad k(x,x_i)=\exp\left\{-\frac{1}{2\sigma_k^2}(x-x_i)^2\right\}.$$

Здесь σ_K и σ_k являются регулируемыми параметрами.

Функция $\bar{v}(V)$ должна удовлетворять граничным и начальным условиям. Для этого достаточно взять произвольные $A(V)$ и $B(V)$, удовлетворяющие условиям

$$A(x,t_0)=g(x),\quad A(l,t)=A(L,t)=0,\quad B(l,t)=B(L,t)=B(x,t_0)=0.$$

Например,

$$A(V)=g(t-t_0+x)-\frac{x-l}{L-l}g(t-t_0+L)+\frac{x-L}{L-l}g(t-t_0+l),\quad B(V)=(t-t_0)(x-l)(x-L).$$

Для построения приближённых функций $\bar{v}(V)$ и $\bar{H}(x)$ необходимо найти параметры регрессии $\alpha_i,\bar{\alpha}_i$ и $b,\bar{b}$. Общее количество параметров равно $M+N_1+2$. Для поиска параметров регрессии воспользуемся методом SVM, предложенный для задачи регрессии в работе [10]. Подставим приближенные функции $\bar{v}(V)$ и $\bar{H}(x)$ в главное уравнение:

$$A_t(V)+B_t(V)\left(\sum_{j=1}^{M}\alpha_jK(V,V_j)+b\right)+B(V)\sum_{j=1}^{M}\alpha_jK_t(V,V_j)\approx$$

$$\approx dA_{xx}(V)+dB_{xx}(V)\left(\sum_{j=1}^{M}\alpha_jK(V,V_j)+b\right)+2dB_x(V)\sum_{j=1}^{M}\alpha_jK_x(V,V_j)+dB(V)\sum_{j=1}^{M}\alpha_jK_{xx}(V,V_j)+$$

$$+S(V)\sum_{j=1}^{N_1}\bar{\alpha}_jk_x(x,\bar{x}_j)+\left(\sum_{j=1}^{N_1}\bar{\alpha}_jk(x,\bar{x}_j)+\bar{b}\right)S_x(V)+\tilde{F}_x(V).$$

Введём обозначения

$$l_B(V)=B_t(V)-dB_{xx}(V),\quad R(V)=A_t(V)-dA_{xx}(V)-\tilde{F}_x(V),$$

$$K_X(V,V_j)=B_{xx}(V)K(V,V_j)+2B_xK_x(V,V_j)+B(V)K_{xx}(V,V_j),$$

$$K_T(V,V_j)=B_t(V)K(V,V_j)+B(V)K_t(V,V_j),\quad \Phi(V,V_j)=K_T(V,V_j)-dK_X(V,V_j),$$

В новых обозначениях выражение с приближенными функциями имеет вид

$$\sum_{j=1}^{M}\alpha_j\Phi(V,V_j)+l_B(V)b+R(V)-S(V)\sum_{j=1}^{N_1}\bar{\alpha}_jk_x(x,\bar{x}_j)-S_x(V)\sum_{j=1}^{N_1}\bar{\alpha}_jk(x,\bar{x}_j)-S_x(V)\bar{b}\approx 0.$$

Введем вспомогательные переменные $e_i=v(\bar{V}_i;H)-w(x_i)$, характеризующие смещения, перепишем выражения в виде равенств, выполняющихся для внутренних точек. Аналогично с помощью $\bar{e}_i$ запишем равенства, полученные подстановкой функции $\bar{v}(V)$ в дополнительные условия для точек P_{ad}. Получим систему равенств:

$$\begin{cases}\sum\limits_{j=1}^{M}\alpha_j\Phi(V_i,V_j)+l_B(V_i)b+R(V_i)-S(V_i)\sum\limits_{j=1}^{N_1}\bar{\alpha}_jk_x(x_i,\bar{x}_j)-S_x(V_i)\sum\limits_{j=1}^{N_1}\bar{\alpha}_jk(x_i,\bar{x}_j)-S_x(V_i)\bar{b}+e_i=0,\\ \qquad\qquad V_i\in P_{in};\\ \sum\limits_{j=1}^{M}\alpha_jB(\bar{V}_l)K(\bar{V}_l,V_j)+A(\bar{V}_l)+B(\bar{V}_l)b-w(x_l)+\bar{e}_l=0,\quad \bar{V}_l\in P_{ad}.\end{cases}\tag{6}$$

Теорема. Параметры регрессии $\alpha_j,\bar{\alpha}_j$ и $b,\bar{b}$ могут быть найдены из решения линейной системы

$$\Lambda X=Y,$$

где $\Lambda=[m_{ij}]_{6\times 6}$ — известная матрица, Y — известный вектор, $X=[\alpha^T,\lambda^T,b,\bar{\alpha}^T,\mu^T,\bar{b}]$ — неизвестный вектор, содержащий параметры регрессии.

Доказательство. Задача минимизации (4) может быть сведена к оптимизационной задаче квадратичного программирования [11], имеющей единственное решение:

$$\min_{\alpha,\bar{\alpha},e,\bar{e}} J, \quad J = \frac{1}{2}(\alpha^T\alpha + \bar{\alpha}^T\bar{\alpha} + \gamma e^T e + \gamma\bar{e}^T\bar{e}),$$

при условиях (6), где $\alpha = [\alpha_1, \alpha_2, ..., \alpha_M]^T$, $\bar{\alpha} = [\bar{\alpha}_1, \bar{\alpha}_2, ..., \bar{\alpha}_{N_1}]^T$, $e = [e_1, e_2, ..., e_M]^T$, $\bar{e} = [\bar{e}_1, \bar{e}_2, ..., \bar{e}_{N_1}]^T$, $\gamma = \tau^{-1} \in R^+$. Вводя множители Лагранжа

$$\lambda = [\lambda_1, \lambda_2, ..., \lambda_M]^T: \ \ 0 \le \lambda_i \le \gamma/2, \quad i = 1, \dots, M,$$
$$\mu = [\mu_1, \mu_2, ..., \mu_{N_2}]^T: \ \ 0 \le \mu_l \le \gamma/2, \quad l = 1, \dots, N_2,$$

запишем функцию Лагранжа:

$$L = \frac{1}{2}(\alpha^T\alpha + \bar{\alpha}^T\bar{\alpha} + \gamma e^T e + \gamma\bar{e}^T\bar{e}) +$$
$$+\sum_{i=1}^{M}\lambda_i\left(\sum_{j=1}^{M}\alpha_j\Phi(V_i,V_j) + l_B(V_i)b + R(V_i) - S(V_i)\sum_{j=1}^{N_1}\bar{\alpha}_j k_x(x_i,\bar{x}_j) - S_x(V_i)\sum_{j=1}^{N_1}\bar{\alpha}_j k(x_i,\bar{x}_j) - S_x(V_i)\bar{b} + e_i\right) +$$
$$+\sum_{l=1}^{N_2}\mu_l\left(\sum_{j=1}^{M}\alpha_j B(\bar{V}_l)K(\bar{V}_l,V_j) + A(\bar{V}_l) + B(\bar{V}_l)b - w(\bar{t}_l) + \bar{e}_l\right).$$

Условия оптимальности Каруша-Куна-Таккера записываются в виде следующей системы:

$$\begin{cases}
\dfrac{\partial L}{\partial\alpha_i} = \alpha_i + \sum\limits_{j=1}^{M}\lambda_j\Phi(V_j,V_i) + \sum\limits_{j=1}^{N_2}\mu_j B(\bar{V}_j)K(\bar{V}_j,V_i) = 0, \quad \dfrac{\partial L}{\partial b} = \sum\limits_{j=1}^{M}\lambda_j l_B(V_j) + \sum\limits_{j=1}^{N_2}\mu_j B(\bar{V}_j) = 0,\\
\dfrac{\partial L}{\partial\lambda_i} = \sum\limits_{j=1}^{M}\alpha_j\Phi(V_i,V_j) + l_B(V_i)b + R(V_i) - S(V_i)\sum\limits_{j=1}^{N_1}\bar{\alpha}_j k_x(x_i,\bar{x}_j) - S_x(V_i)\sum\limits_{j=1}^{N_1}\bar{\alpha}_j k(x_i,\bar{x}_j) - S_x(V_i)\bar{b} + e_i = 0,\\
\dfrac{\partial L}{\partial\bar{\alpha}_i} = \bar{\alpha}_i - \sum\limits_{j=1}^{M}\lambda_j k_x(x_j,\bar{x}_i)S(V_j) - \sum\limits_{j=1}^{M}\lambda_j k(x_j,\bar{x}_i)S_x(V_j) = 0, \quad \dfrac{\partial L}{\partial\bar{b}} = -\sum\limits_{j=1}^{M}\lambda_j S_x(V_j) = 0,\\
\dfrac{\partial L}{\partial\mu_i} = \sum\limits_{j=1}^{M}\alpha_j B(\bar{V}_i)K(\bar{V}_i,V_j) + A(\bar{V}_i) + B(\bar{V}_i)b - w(\bar{t}_i) + \bar{e}_i = 0,\\
\dfrac{\partial L}{\partial e_i} = \gamma e_i + \lambda_i = 0, \quad \dfrac{\partial L}{\partial\bar{e}_i} = \gamma\bar{e}_i + \mu_i = 0.
\end{cases}$$

Исключив e_i и $\bar{e}_i$, получим линейную систему уравнений: $\Lambda X = Y$, где

$$X = ((\alpha^T)_{1\times M}, (\lambda^T)_{1\times M}, b, (\bar{\alpha}^T)_{1\times N_1}, (\mu^T)_{1\times N_2}, \bar{b})^T_{2M+N_1+N_2+2},$$
$$Y = (0_M, -R(V_i)_{1\times M}, 0, 0_{N_1}, (w(\bar{t}_i) - A(\bar{V}_i))_{1\times N_2}, 0)^T_{2M+N_1+N_2+2},$$

матрица $\Lambda = [m_{ij}]_{6\times 6}$ — симметрична, а ее блоки имеют вид:

$$\begin{array}{ccc}
m_{11} = I_M & m_{12} = (\Phi(V_j,V_i))_{M\times M} & m_{13} = (0)_{M\times 1}\\
m_{14} = (0)_{M\times N_1} & m_{15} = (B(\bar{V}_j)K(\bar{V}_j,V_i))_{M\times N_2} & m_{16} = (0)_{M\times 1}\\
m_{21} = m_{12}^T & m_{22} = -\dfrac{1}{\gamma}I_M & m_{23} = (l_B(V_i))_{M\times 1}\\
m_{24} = (-S(V_i)k_x(x_i,\bar{x}_j) - S_x(V_i)k(x_i,\bar{x}_j)_{M\times N_1} & m_{25} = (0)_{N_2\times 1} & m_{26} = (-S_x(V_i))_{M\times 1}\\
m_{31} = (0)_{1\times M} & m_{32} = m_{23}^T & m_{33} = (0)\\
m_{34} = (0)_{1\times N_1} & m_{35} = (B(\bar{V}_j))_{1\times N_2} & m_{36} = (0)\\
m_{41} = (0)_{N_1\times M} & m_{42} = m_{24}^T & m_{43} = (0)_{N_1\times 1}\\
m_{44} = I_{N_1} & m_{45} = (0)_{N_1\times N_2} & m_{46} = (0)_{N_1\times 1}\\
m_{51} = m_{15}^T & m_{52} = (0)_{N_2\times M} & m_{53} = m_{35}^T\\
m_{54} = (0)_{N_2\times N_1} & m_{55} = -\dfrac{1}{\gamma}I_{N_2} & m_{56} = (0)_{N_2\times 1}\\
m_{61} = (0)_{1\times M} & m_{62} = m_{26}^T & m_{63} = (0)\\
m_{64} = (0)_{1\times N_1} & m_{65} = (0)_{1\times N_2} & m_{66} = (0).
\end{array}$$

Отметим, что параметр регуляризации τ подбирается для каждой задачи индивидуально, что является вычислительно трудоемкой процедурой. □

Решая линейную систему алгебраических уравнений $\Lambda X = Y$, находим параметры регрессии $\alpha_j, \bar{\alpha}_j, b, \bar{b}$ и строим приближенное решение $\bar{H}(x)$, $\bar{v}(V)$ по формулам (5) задачи об источнике (3), после чего строим решение исходной задачи об источнике (1).

Замечание. Параметр $\bar{b}$ в выражении (5) может быть произвольным, если функция S не зависит от x. В самом деле, при переходе от задачи (2) к задаче (3) получаем $v_t = dv_{xx} + H'(x)S(t) + \bar{F}_x(x,t)$ и, следовательно, функция источника $H(x)$ может быть определена с точностью до константы. В этом случае матрица Λ в системе уравнений $\Lambda X = Y$ имеет вид $[m_{ij}]_{5\times 5}$ и не содержит блоков m_{6i} и m_{i6}, $i = 1, ..., 6$; векторы X, Y имеют размер $2M + N_1 + N_2 + 1$ и не содержат последних компонент. Для единственности решения задачи об источнике (3) необходимо задать значение $H(x)$ в какой-либо точке $x \in [l, L]$.

2.1 Алгоритм метода опорных векторов

1. Задаем значения границы области и коэффициентов задачи (3) l, L, t_0, T, d, а также значение точки измерения дополнительной информации t^*. Задаем функции $S(V)$, $\tilde{F}_x(V)$, $g(x)$, $w(x)$ и производные $S_x(V)$, $g_x(x)$, $g_{xx}(x)$.
2. Задаем значения параметров SVM σ_k, σ_K, γ.
3. Выбираем подходящие вспомогательные функции $A(V)$, $B(V)$ и вычисляем производные $A_x(V)$, $A_{xx}(V)$, $A_t(V)$, $B_x(V)$, $B_{xx}(V)$, $B_t(V)$.
4. Записываем функции ядер $K(V, V_i)$, $k(x, x_i)$, их производные $K_x(V, V_i)$, $K_{xx}(V, V_i)$, $K_t(V, V_i)$, $k_x(x, x_i)$ и функции $l_B(V)$, $R(V)$, $\Phi(V, V_j)$.
5. Задаем размеры сеточной области M, N_1, N_2. Генерируем точки расчетной сетки P_{in}, P_{ad}, $\bar{X}$.
6. Заполняем матрицу Λ и вектор Y заданными функциями.
7. Решаем линейную систему $\Lambda X = Y$ методом LU-разложения и определяем вектор X.
8. Используя найденные параметры регрессии, вычисляем приближенные функции $\bar{v}(V)$ и $\bar{H}(x)$.
9. Вычисляем функцию $U(x,t)$ интегрированием функции $\bar{v}(x,t)$ с использованием априорной информации о задаче (например, значение функции $U(x,t)$ в левом конце отрезка).
10. Выполняем обратную замену и переходим к функции $u(x,t)$ задачи (1).

3 Тестовые расчеты

Рассмотрим задачу определения источника вида (2) на области $D = (0,1) \times (1,5)$ с $d = 1$, функциями $S = S(x,t) = 2(1+\pi^2)e^{t-1}\cos(\frac{\pi}{4} - \frac{\pi x}{2})$ и $F(x,t) = 0$

$$\begin{cases} U_t = U_{xx} + 2(1+\pi^2)e^{t-1}\cos(\frac{\pi}{4} - \frac{\pi x}{2})H(x), & (x,t) \in D; \\ U(x,1) = \cos(\pi x), & 0 < x < 1; \\ U_x(0,t) = U_x(1,t) = 0, & t > 1; \\ U(x,t^*) = e^{t^*-1}\cos(\pi x), & t > 1, \end{cases} \tag{7}$$

точное решение которой имеет вид $U(x,t) = e^{t-1}\cos(\pi x)$, $H(x) = \sin(\frac{\pi}{4} - \frac{\pi x}{2})$. После введения новых функций (см. п. 1) задача перепишется в виде (3) следующим образом:

$$\begin{cases} v_t = v_{xx} + 2(1+\pi^2)e^{t-1}\cos(\frac{\pi}{4} - \frac{\pi x}{2})H'(x) + \pi(1+\pi^2)e^{t-1}\sin(\frac{\pi}{4} - \frac{\pi x}{2})H(x), & (x,t) \in D; \\ v(x,1) = -\pi\sin(\pi x), & 0 < x < 1; \\ v(0,t) = v(1,t) = 0, & t > 1; \\ v(x,t^*) = -\pi e^{t^*-1}\sin(\pi x), & t > 1. \end{cases}$$

В численных расчётах используем сеточную область размера $M = 40 \times 40$, $N_1 = 2500$, $N_2 = 2500$, значения параметров $\gamma = 10^{11}$, $\sigma_k = \sigma_K = 1$. Значения параметров были выбраны эмпирически, хотя они

согласуются с расчетами в работе [10]. В этом случае, размер матрицы Λ будет равен 8202×8202. Для однозначного определения функции $U(x,t)$ при интегрировании $\bar{v}(x,t)$ дополнительно задаются значения функции на левом конце отрезка: $U(0,t) = e^{t-1}$. Результаты восстановления функции источника $H(x)$ и функции $U(x,t)$ представлены на рис. 1. Для описания полученных результатов используем абсолютные ошибки $\varepsilon_{U_i}(x) = \bar{U}(x,t_i) - U(x,t_i)$ и $\varepsilon_H(x) = \bar{H}(x) - H(x)$.

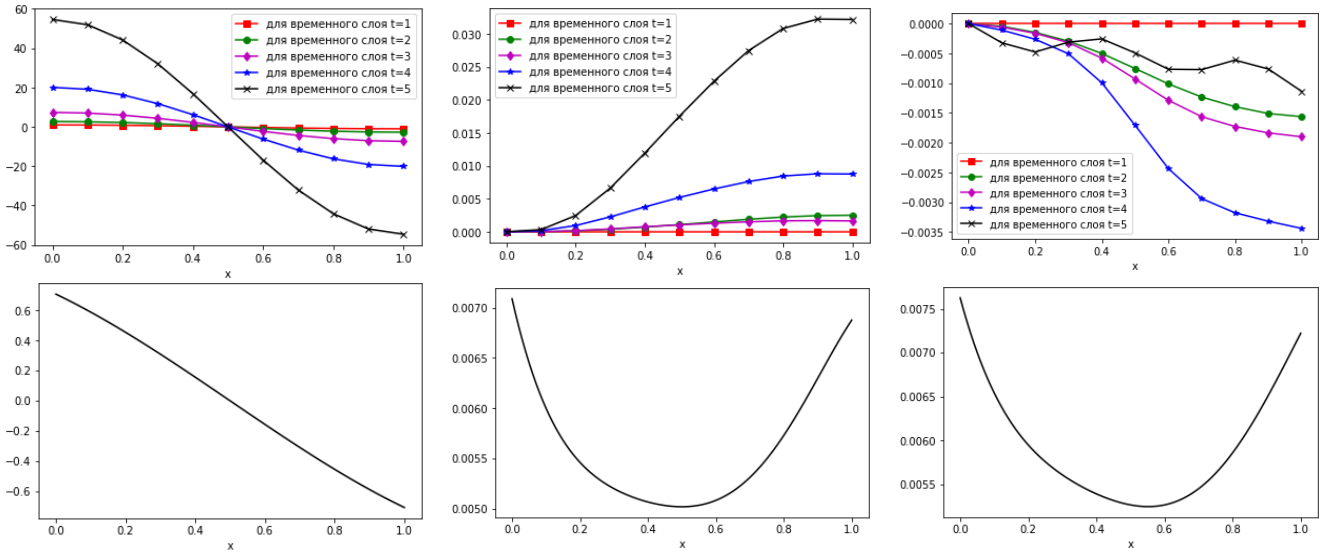


Рис. 1: Решение задачи об источнике для параболического уравнения (7) методом опорных векторов. Сверху расположены графики для функции $U(x,t)$ при некоторых фиксированных моментах времени, снизу — для $H(x)$. Слева направо: точное решение задачи (7), абсолютная ошибка ε при $t^* = 2$ и ε при $t^* = 5$.

Здесь максимальная абсолютная ошибка $\max\limits_{x\in[l,L]} \varepsilon(x)$ имеет порядок 10^{-2} для функции $U(x,t)$ и порядок 10^{-3} для функции $H(x)$. Вычислительная сложность алгоритма зависит от количества точек в расчетных областях. В таблице 1 приведены максимальные относительные ошибки при разных количествах точек расчетной области. Отметим, что чем больше точек в расчетной области, тем выше точность восстановления функций $U(x,t)$ и $H(x)$.

Таблица 1: Сравнение абсолютной ошибки $\max\limits_{x\in[l,L]} \varepsilon(x)$ для функций $U(x,t)$ и $H(x)$ при разных количествах расчетных точек алгоритма SVM

Количество расчетных точек	$\max\limits_{x\in[l,L]} \varepsilon_{U_5}(x)$			$\max\limits_{x\in[l,L]} \varepsilon_H(x)$		
$N_1 = N_2 \setminus M$	20×20	30×30	40×40	20×20	30×30	40×40
100	0.083109	0.135313	0.201724	0.007635	0.006735	0.006893
200	0.102774	0.089125	0.113807	0.007351	0.005969	0.004864
1000	0.119518	0.051144	0.044182	0.008042	0.009771	0.0067
2500	0.121188	0.0462093	0.032215	0.008178	0.008799	0.00709
5000	0.121899	0.043988	0.029626	0.007919	0.009344	0.008965

Для применения метода опорных векторов к решению задачи об источнике вида (2) напрямую, без использования замены $v = U_x$, необходимо, чтобы приближенная функция $\bar{U}(x,t)$ удовлетворяла граничным и начальным условиям. Для этого функции $A(V)$ и $B(V)$ удовлетворяют соотношениям

$$A(x,t_0) = \tilde{f}(x), \quad A_x(l,t) = A_x(L,t) = 0, \quad B_x(l,t) = B_x(L,t) = B(x,t_0) = B(l,t) = B(L,t) = 0. \tag{8}$$

Например,

$$A(V) = \tilde{f}(t - t_0 + x) - \frac{(x-l)^2}{2(L-l)}\tilde{f}'(t - t_0 + L) + \frac{(x-L)^2}{2(L-l)}\tilde{f}'(t - t_0 + l), \quad B(V) = (t - t_0)(x-l)^2(x-L)^2.$$

Отметим, что на границе области D функция $\bar{U}(x,t)$ определяется только функцией $A(V)$ в силу условий (8) и выражения для функции $\bar{U}(x,t)$. Ввиду этого разность между тестовым и полученным решениями для функции $u(x,t)$ может быть сколь угодно большой. В качестве регуляризации задачи об источнике необходимо задать значения функции $A(V)$ при $x = l$ и $x = L$, которые могут быть получены из априорной информации о задаче [9]. При этом выбор функции $A(V)$ усложняется.

Заключение

В работе приведен и проанализирован алгоритм метода опорных векторов решения задачи об источнике для параболического уравнения с граничными условиями типа Неймана по дополнительной информации о процессе в фиксированный момент времени. Основная идея метода опорных векторов заключается в представлении решения задачи об источнике, а также решения прямой задачи для параболического уравнения в виде разложения по базисным функциям — ядра гауссова типа. Ввиду того, что решение исследуемой задачи неединственно, задача сводится к задаче с граничными условиями типа Дирихле, для которой разработан алгоритм регуляризации. Основная идея регуляризации состоит в использовании априорной информации о решении задачи для параболического уравнения на границе области. Для рассмотренного примера ошибка имеет порядок 10^{-2} для функции $U(x,t)$ и порядок 10^{-3} для функции $H(x)$.

В дальнейшем, разработанный алгоритм планируется применить для решения задач определения источников распространения информации в онлайн социальных сетях, описываемой диффузионно-логистическим уравнением параболического типа [12].

Список литературы

[1] Cortes C., Vapnik V. Support-vector networks // Machine Learning. 1995. V. 20, no. 3. P. 273–297.

[2] Mehrkanoon S., Suykens J.A.K. Learning solutions to partial differential equations using LS-SVM // Neurocomputing. 2015. V. 159. P. 105–116.

[3] Yu J., Li F., Yu Sh., Wu Z. The approximate solutions to source inverse problem of 1-D convection-diffusion equation by LS-SVM // Inverse Probl Sci Eng. 2018. V. 26, iss. 5. P. 677–690.

[4] Han X., Wang J., Wu Z., Li G., Wu Y., Li J. Learning solutions to two dimensional electromagnetic equations using LS-SVM // Neurocomputing. 2018. V. 317. P. 15–27.

[5] Aarts L., Van der Veer P. Neural network method for solving partial differential equations // Neural Processing Letters. 2001. V. 14. P. 261–271.

[6] McFall K.S., Mahan J.R. Artificial neural network method for solution of boundary value problems with exact satisfaction of arbitrary boundary conditions // IEEE Trans. Neural Netw. 2009. V. 20, no. 8. P. 1221–1233.

[7] Dai G., Ma R., Wang H., Wang F., Xu K. Partial differential equations with Robin boundary condition in online social networks // Discrete Continuous Dyn Syst Ser B. 2015. V. 20, no. 6. P. 1609–1624.

[8] Kabanikhin S., Krivorotko O., Bektemessov M., Bektemessov Z., Zhang S. Differential evolution algorithm of solving an inverse problem for the spatial Solow mathematical model // arXiv:1904.10627. 2019.

[9] Kabanikhin S.I. Definitions and examples of inverse and ill-posed problems // J. Inverse Ill-Posed Probl. 2008. V. 16. P. 317–357.

[10] Wu Z., Ding C., Li G., Han X., Li J. Learning solutions to the source inverse problem of wave equations using LS-SVM // J. Inverse Ill-Posed Probl. 2019. DOI: 10.1515/jiip-2018-0066.

[11] Smola A., Schölkopf B. A Tutorial on Support Vector Regression // NeuroCOLT Technical Report NC-TR-98-030, Royal Holloway College, University of London, UK, 1998.

[12] Du B., Lian X., Cheng X. Partial differential equation modeling with Dirichlet boundary conditions on social networks // Boundary Value Problems. 2018. V. 2018. DOI: 10.1186/s13661-018-0964-4.

Кабанихин Сергей Игоревич — д.ф.-м.н., чл.-корр. РАН, профессор, гл.науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: kabanikhin@sscc.ru;
Криворотько Ольга Игоревна — к.ф.-м.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: olga.krivorotko@sscc.ru;
Ануфриенко Екатерина Андреевна — студент Новосибирского государственного университета;
e-mail: ekanufrienko@mail.ru.
Дата поступления — 30 июня 2019 г.

АЛГОРИТМ ВЫДЕЛЕНИЯ УГЛОВЫХ СТРУТУР НА ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ МАСШТАБИРУЕМОГО ИЕРАРХИЧЕСКОГО ДЕТЕКТОРА

И. Г. Казанцев[1], Б. О. Мухаметжанова[2], К. Т. Искаков[2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[2] *Евразийский национальный университет им. Л.Н. Гумилева, Нур-Султан, Казахстан*

УДК 517.562
DOI: 10.24411/9999-016A-2019-10032

Рассматриваются масштабируемые маски выделения углов на изображениях, применяемые при обработке скользящим по изображению окном. В отличие от известных алгоритмов, матрицы маски больших размеров конструируются добавлением строк и столбцов по периметру к матрице меньших масок, оставляя подматрицы неизменными и добавляя периферийные элементы по простому правилу.

Ключевые слова: обработка изображений, скользящее окно, детектор углов.

Введение

В работе рассматриваются новые маски выделения углов на изображениях для применения в традиционном методе скользящих фрагментов, или окон. Угловые точки являются важной локальной особенностью изображения и принадлежат к классу так называемых доминантных, характерных, сингулярных, или точек интереса и т.д. Они используются как опорные точки в работе со стереопарами, как признаки в распознавании лиц (например, уголки глаз), отпечатков пальцев и букв в текстах. Важные приложения включают также калибровку камер, отслеживание движущихся объектов в робототехнике и машинном зрении, согласование изображений и распознавание образов. Углы инвариантны к вращению и изменению условий освещения. Интерес к созданию помехоустойчивых и эффективных алгоритмов обнаружения углов существует на протяжении десятилетий и источниками создаваемых методов служат многие области науки, от цифровой обработки снимков и оптики до дифференциальной и интегральной геометрии. Недавние обзоры и статьи, посвященные отдельным методам, можно найти в библиографическом списке [1]–[15].

Данная работа основана на новой модели угла, обладающей пологими сторонами, в отличие от традиционного задания идеального угла с помощью обрывистых функций-ступенек. [15]. Значениям яркости изображения на границе предполагаемой угловой структуры придаются веса, промежуточные между весами области фона и собственно угла. В работе показаны примеры построения детекторов углов, рассмотрены их свойства и предложены алгоритмы выделения углов. Результаты иллюстрируются.

1 Модель

В зависимости от приложений, углом называют и собственно вершину угла, то есть отдельную точку, и менее локальный объект, включающий помимо вершины еще и лучи, распространяющиеся из нее, а также угловую структуру целиком, с определением длин каждой из сторон угла, то есть треугольник. В обработке изображений прямые, составляющие стороны угла, называются его границами, в которых визуально наблюдаются доминирующие изменения яркости, характеризующие различие одной области снимка (угла) от другой (фона). Один из распространенных подходов к поиску углов состоит в выделении границ и

Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2016-0003).

ISBN 978-5-901548-42-4

Таблица 1: Традиционные маски разхмера $3 \times 3,\ 5 \times 5,\ 7 \times 7$

A =

?			...			?
		-4	-4	**5**		
		-4	**5**	**5**		
		-4	-4	**5**		
?			...			?

B =

?			...			?
	-9	-9	-9	-9	**16**	
	-9	-9	-9	**16**	**16**	
	-9	-9	**16**	**16**	**16**	
	-9	-9	-9	**16**	**16**	
	-9	-9	-9	-9	**16**	
?			...			?

C =

-16			...			-16
-16			...			-16
		-16	-16	-16	-16	-16
		-16	**33**	**33**	**33**	**33**
		-16	**33**	**33**	**33**	**33**
		-16	**33**	**33**	**33**	**33**
-16	-16	-16	**33**	**33**	**33**	**33**

Таблица 2: Маски Кирша — первые шесть из восьми вращений

5	**5**	**5**
-3	**0**	-3
-3	-3	-3

,

-3	**5**	**5**
-3	**0**	**5**
-3	-3	-3

,

-3	-3	**5**
-3	**0**	**5**
-3	-3	**5**

,

-3	-3	-3
-3	**0**	**5**
-3	**5**	**5**

,

-3	-3	-3
-3	**0**	-3
5	**5**	**5**

,

-3	-3	-3
5	**0**	-3
5	**5**	-3

, ...

бинаризации, и последующих процедур обнаружения на бинарном аналоге изображения. В основе метода лежит исследование яркостей изображения в окрестности точки на равенство нулю второй производной и изменение знака в направлении нормали к границе. Известно множество градиентных масок, предложенных для этого метода. Отметим, что одномерный яркостной профиль изображения, нормальный к границе, может быть разнообразным, обычно выделяют такие виды: импульс, ступенька, пандус и крыша. Поскольку движение скользящего дифференцирующего окна и вычисление его свертки с локальным фрагментом изображений происходит автоматически, возникают проблемы, связанные с импульсным шумом, который проявляется в коротких ложных контурах длиной в единицы пикселов. Известны эффективные быстрые алгоритмы регуляризации дифференцирующей свертки, например, дополнительной сверткой с двумерной функцией Гаусса, параметры которой можно варьировать.

Вторая группа алгоритмов не производит выделение границ и бинаризацию, а работает непосредственно с полутоновым изображением, сканируя его элементы локальной окрестностью и вычисляя корреляцию фрагмента снимка с маской, программирующей модель угловой структуры. Предполагается, что внутренняя область угла приближенно представляет собой плато. Размер маски нечетный, при сканировании снимка ее центральный элемент помещается в центр исследуемого фрагмента изображения. Для каждого элемента изображения вычисляются величины корреляции маски при ее различных вращениях около центрального элемента. Максимальное абсолютное значение из них оставляется как мера наличия угла в точке. Как правило, маски сконструированы для выделения прямых углов. Практикой установлено, что маска прямых углов достаточно хорошо отслеживает углы несколько меньшие или большие 90 градусов.

Приведем примеры расчета масок традиционным методом. Рассмотрим традиционные маски A, B, C размерами 3×3, 5×5 и 7×7 элементов (Таб. 1) соответственно, предназначенные для выделения углов, раствор которых составляет 90 градусов. Например, значения 25 элементов маски 5×5 рассчитываются следующим образом. На угол отводится $3 \cdot 3 = 9$ положительных элементов, на отрицательные остается $25 - 9 = 16$ элементов. Чтобы маска была дифференцирующей, сумма ее элементов должна равняться нулю. Поэтому область угла заполняется девятью элементами со значениями $+16$, область фона – шестнадцатью элементами со значениями (-9). Имеем $9 \cdot (+16) + 16 \cdot (-9) = 0$. (Таб. 1, маска B).

Рассмотрим маску C размером 7×7 элементов для прямых углов (Таб. 1). Значения ее 49 элементов рассчитываются аналогично. На угол отводится $4 \cdot 4 = 16$ положительных элементов, на отрицательные остается $49 - 16 = 33$ элементов. Чтобы сумма ее элементов равнялась нулю, область угла заполняется шестнадцатью элементами со значениями $+33$, область фона – тридцатью тремя элементами со значениями (-16). Имеем $16 \cdot (+33) + 33 \cdot (-16) = 0$.

Среди подобных схем конструирования дифференцирующих масок получила недавно популярность маска Кирша, моделирующая ориентированные границы при анализе текстур в видеоданных (Таб. 2). В этих приложениях требуется высокое быстродействие ввиду огромного объема данных, поэтому используются маски размером 3×3.

Рассмотренные детекторы обладают рядом недостатков. Они имеют высокий коэффициент усиления шума и не обладают свойствами, необходимыми для организации иерархических вычислений. Например,

Таблица 3: Декомпозиция угла $\frac{\pi}{2}$: $W_3^{1,\pi/4} + W_3^{2,\pi/4} = W_3^{1,\pi/2}$

-1	⋮	-1	**3**	**7**	**7**	**3**
⋮	-1	-1	**3**	**7**	**3**	-1
⋮		-1	**3**	**3**	-1	-1
⋮	-1	-1	**0**	-1	-1	-1
⋮	…	-1	-1			
⋮	-1		…		-1	
-1			⋮			-1

+

-1	⋮	-1	-1	-1	-1	**3**
⋮	-1	-1	-1	-1	**3**	**7**
⋮		-1	-1	**3**	**7**	**7**
⋮	-1	-1	**0**	**3**	**3**	**3**
⋮	…	-1	-1			
⋮	-1		⋮		-1	
-1			⋮			-1

= 2 ·

-1	⋮	-1	**1**	**3**	**3**	**3**
⋮	-1	-1	**1**	**3**	**3**	**3**
⋮		-1	**1**	**3**	**3**	**3**
⋮	-1	-1	**0**	**1**	**1**	**1**
⋮	…	-1	-1	…	…	-1
⋮	-1		…		-1	
-1			⋮			-1

Таблица 4: Декомпозиция угла $\frac{3\pi}{4}$: $W_2^{1,\pi/4} + W_2^{2,\pi/4} + W_2^{3,\pi/4} = W_2^{1,3\pi/4}$

-1	-1	**3**	**7**	**3**
-1	-1	**3**	**3**	-1
-1	-1	**0**	-1	-1
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

+

-1	-1	-1	-1	**3**
-1	-1	-1	**3**	**7**
-1	-1	**0**	**3**	**3**
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

+

-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
-1	-1	**0**	**3**	**3**
-1	-1	-1	**3**	**7**
-1	-1	-1	-1	**3**

=

-3	-3	**1**	**5**	**5**
-3	-3	**1**	**5**	**5**
-3	-3	**0**	**5**	**5**
-3	-3	-3	**1**	**5**
-3	-3	-3	-3	**1**

данные сканирования маской 5×5 проблематично использовать в вычислениях с масками больших размеров. Это отмечено (Таб. 1) вопросительными знаками для матриц A и B. Маски Кирша (Таб. 2) 3×3 также требуют новых элементов при использовании масок больших размеров. Вычисления с масками последовательно увеличивающихся размерностей дают информацию о линейных и площадных параметрах и моменте перехода угла в область фона. Поэтому были предприняты попытки конструирования масок, обладающих свойствами иерархичности, или масштабируемости. Такие маски получаются, если предположить, что граница между угловой структурой и фоном проходит внутри одного какого-то пиксела, а не между двумя соседними пикселами по их общей границе, как в матрицах A, B, C.

Тогда, вычисляя доли пиксела, занимаемые фоном и углом, для каждого элемента маски, и приводя их к целым взаимно простым значениям, можно получить матрицы, представленные в Таб. 2 и Таб. 3. Применяется обозначение масок $W_n^{k,\alpha}$, где $k = 1, \ldots, 8$ — номер угла, α — величина угла, n — целое число, определяющее размер квадратной маски $N = 2n + 1$. Важное свойство матриц состоит в возможности разлагать маски углов $\pi/2$ и $3\pi/4$ в суммы углов с раствором $\pi/4$ (Таб. 2 и Таб. 3):

$$W_3^{1,\pi/4} + W_3^{2,\pi/4} = W_3^{1,\pi/2}$$
$$W_2^{1,\pi/4} + W_2^{2,\pi/4} + W_2^{3,\pi/4} = W_2^{1,3\pi/4}.$$

Это же справедливо и для различных сочетаний углов, например, имеющих общую вершину (Таб. 4). В используемой нами терминологии пикселы тела угла имеет вес $+7$, граница угла имеет вес $+3$, а фону присвоен вес -1. Приводимые примеры являются следствиями соотношений, доказательства которых будут приведены в готовящейся расширенной публикации.

Одним из свойств является то, что новые маски имеют значения на границе угла, равные полусумме фона и тела угла: $3 = (7 + (-1))/2$ (Таб. 3). Это придает такому ядру свертки свойства гладкости и стабильности в оценке значения свертываемого фрагмента изображения на зашумленных границах угла по сравнению с масками в Таб. 1.

Другое интересное свойство масок состоит в том, что периферийные, или обрамляющие элементы, производят также дифференцирующий эффект, как и всё ядро в целом. Периферийные элементы матрицы $W_2^{1,\pi/4}$ (табл. 4, справа, выделены жирным шрифтом), например, в сумме равны нулю: $3 + 7 + 7 + 3 + 20 \cdot (-1) = 0$. Это справедливо для всех подматриц. В задаче достижения под-пиксельного разрешения, или при обработке снимков высокого разрешения, где угловые структуры имеют размеры в десятки пикселов, использование иерархических масштабируемых масок становится особенно эффективным. Полученные детекторы легко трансформируются для полутоновой обработки заменой центрального нулевого элемента на единицу.

Таблица 5: Маска двух углов с общей вершиной $W_2^{1,\pi/4} + W_2^{5,\pi/4} = W_2^{(1\oplus 5),\pi/4}$

-1	-1	**3**	**7**	**3**
-1	-1	**3**	**3**	-1
-1	-1	**0**	-1	-1
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

+

-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
-1	-1	**0**	-1	-1
-1	**3**	**3**	-1	-1
3	**7**	**3**	-1	-1

= 2·

-1	-1	**1**	**3**	**1**
-1	-1	**1**	**1**	-1
-1	-1	**0**	-1	-1
-1	**1**	**1**	-1	-1
1	**3**	**1**	-1	-1

, $W_2^{1,\pi/4} =$

-1	**-1**	**3**	**7**	**3**
-1	-1	3	3	**-1**
-1	-1	0	-1	**-1**
-1	-1	-1	-1	**-1**
-1	**-1**	**-1**	**-1**	**-1**

1.1 Основные уравнения

Обозначим F — часть изображения, или фрагмент размером $N \times N$, на котором остановилось скользящее окно, или маска W того же размера. Мера сходства фрагмента и маски может быть получена из известного тождества

$$\|F - W\|^2 = \|F\|^2 - 2\langle F, W\rangle + \|W\|^2 \tag{1}$$

в выбранной норме гильбертова пространства. Из этого соотношения следует, что фрагмент F аппроксимируется идеальным углом W наилучшим образом поиском максимума скалярного произведения $\langle F, W\rangle$. Тогда можно сформулировать первый критерий обнаружения угла в виде

$$Q_1 = \max_{k,\alpha,n} \langle F, W_n^{k,\alpha}\rangle. \tag{2}$$

Двумерные массивы могут быть вытянуты в векторы для удобства использования обозначений линейной алгебры. Естественным способом при наших иерархических построениях выглядит вытягивание фрагмента в вектор по спирали, начиная с центрального элемента, дающего первый элемент вектора. Второй элемент вектора равен ближайшему элементу матрицы над центральным элементом, затем производится обход по часовой стрелке всех элементов, непосредственно обрамляющих центральный. Всего их восемь. Затем аналогично выбираются элементы следующих обрамлений. Результат такого лексикографического порядка перевода матрицы маски $W_2^{i,\pi/4}$ в векторы φ_i имеет вид:

$$\varphi_1 = (\underbrace{\mathbf{0}}\ \underbrace{\mathbf{3}\ \mathbf{3}\ -1\ -1\ -1\ -1\ -1\ -1}\ \underbrace{\mathbf{3}\ \mathbf{7}\ \mathbf{3}\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1}),$$

$$\varphi_2 = (\underbrace{\mathbf{0}}\ \underbrace{-1\ \mathbf{3}\ \mathbf{3}\ -1\ -1\ -1\ -1\ -1}\ \underbrace{-1\ -1\ \mathbf{3}\ \mathbf{7}\ \mathbf{3}\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1}),$$

$$\dots$$

$$\varphi_8 = (\underbrace{\mathbf{0}}\ \underbrace{-1\ -1\ -1\ -1\ -1\ -1\ \mathbf{3}\ \mathbf{3}}\ \underbrace{-1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ \mathbf{3}\ \mathbf{7}\ \mathbf{3}})$$

Тогда имеется восемь векторов идеальных углов $(\varphi_1, \dots, \varphi_8)$, которые принимаются за базисные и ищется проекция фрагмента F на линейную оболочку, составленную векторами $\{\varphi_i\}$ по методу Ритца. Искомая аппроксимация

$$F^* = \sum_{i=1}^{8} c_i\, \varphi_i \tag{3}$$

ищется скалярным умножением обеих частей равенства (3) на базисные функции $\{\varphi_k\}$, $k = 1, \dots, 8$, что дает систему линейных алгебраических уравнений

$$Gc = d \tag{4}$$

из восьми уравнений на коэффициенты $c = (c_i)$ с матрицей Грама

$$G(i,k) = \langle \varphi_i, \varphi_k\rangle \tag{5}$$

и вектором $d = \{d_k\}$ правой части

$$d_k = \langle F, \varphi_k\rangle. \tag{6}$$

Матрица G в случае $n = 1$, то есть для скользящего окна размера $(2n+1) \times (2n+1) = 3 \times 3$, имеет вид

$$G = 8 \begin{pmatrix} 3 & 1 & -1 & -1 & -1 & -1 & -1 & 1 \\ 1 & 3 & 1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & 3 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 3 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 3 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 3 & 1 & -1 \\ -1 & -1 & -1 & -1 & -1 & 1 & 3 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 & 1 & 3 \end{pmatrix} \tag{7}$$

и является циркулянтной и сингулярной. Применяя псевдообратную матрицу G^+ к вектору данных d, вычисляем минимальное нормальное решение

$$c = G^+ d. \tag{8}$$

Тогда критерий обнаружения угла имеет вид

$$Q_2 = \max_{i=1,\ldots,8} |c_i|. \tag{9}$$

2 Численный метод

При движении окна размером $N \times N$, $N = 2n - 1$ с центром в пикселе с координатами (i, j) по полю изображения f производится дискретная свертка фрагмента изображения с матрицей

$$u_n^r(i,j) = \sum_{p=-n}^{n} \sum_{q=-n}^{n} f(i-p, j-q)\, W_n^r(p,q). \tag{10}$$

Здесь индексы r и n обозначают версии маски, образованные вращением матрицы, а раствор угла фиксирован $(\pi/4)$. Для углов в 45 и 90 градусов имеется 8 дискретных вариантов вращения маски, оставляющих ее элементы целочисленными. Условимся маску, имеющую левую сторону угла ориентированной вертикально, называть основной и ей будет соответсвовать индекс $r = 1$. Например, первые слева маски в Таб. 2 и Таб. 3. Фиксируя неизменным размер n, вычисляем изображение

$$U_n^R(i,j) = \max_{r=1,\ldots,R} |u_n^r(i,j)|. \tag{11}$$

Здесь количество R вращательных версий маски угла равно 8 и для углов $\pi/2$ и $\pi/4$. Ожидается, что максимальный отклик происходит при ориентации маски, наилучшим образом соответствующей направлению угла, наблюдаемому во фрагменте.

Распространение угла вдоль своих сторон, то есть его размер, можно оценить, варьируя еще один параметр — размер маски n. Для этого выберем целое M в качестве оценки максимального размера угла, встречающегося на снимке, и вычисляем изображение

$$U_M^R(i,j) = \max_{n=1,\ldots,M} U_n^R(i,j). \tag{12}$$

Предполагается, что с увеличением испытуемого размера угла отклик растет до пределов распространения угла, и достигнув их, происходит насыщение отклика вследствие захвата маской областей фона. При исследовании визуальных свойств изображения $U_M^R(i,j)$ ожидается, что мы получаем контрастный вариант исходного снимка, с увеличенными яркостями в угловых точках. Эта схема алгоритма использована для обоих критериев Q_1 и Q_2.

Получить угловые точки теперь можно введением классифицирующего порога, оставляющего на изображении значимые, или доминантные точки. Известны многие методы поиска порога, в том числе и порогов динамических, локально адаптированных. Очевидно, что некоторая информация теряется при вычислении максимумов на этапах (10)–(12). Развиваемый нами иерархический масштабируемый подход предполагает сохранить эти этапы как признаки и исследовать данные в целом, не только покоординатным поиском максимумов. Алгоритм позволяет сортировать угловые структуры и затем анализировать по различным признакмк, например, по направленим углов, и строить так называнмые “розы углов”.

Масштабируемость проявляется в нескоьких аспектах. Здесь рассматриваются углы, раствор котрых кратен 45. Замечено, что матрицы углов в 90 и 135 могут быть сконструированы из углов в 45. Таким образом, достаточно вычислить отклики на 8 вращательных версий угла в 45, и затем суммированием их получить отклики на соответсвующие конфигурации. Эти конфигурации включают любые комбинации, например, сочетание нескольких углов с общей вершиной.

Масштабируемость размеров масок основывается на том, что n-я маска размера $(2n+1)\times(2n+1)$ содержит подматрицу размера $(2n-1)\times(2n-1)$ и структура опоясывающих элементов имеет простое описание — элементы повторяются вдоль рапространения сторон угла из ыершины. Поэтому вычисления скалярного произведения производятся рекурсивно, начиная с матриц 3×3, 5×5 и далее, последовательно вычисляя скалярные произведения периферийных строк и столбцов с соответствующими пикселами по периметру фрагмента.

Сумма периферийных по периметру элементов масок равна нулю. Это означает, что в добавление к свойству масок иметь дифференцирующий характер в целом (сумма элементов равна нулю), этим свойством обладает и периметр маски. При размерах маски, превышающих область угла, когда периметр маски достигает фона изображения с произвольными значениями, вклад этих областей в значения критерия изменится и будет наблюдаться разладка. Выбрать заранее размер маски соразмерно размеру искомых угловых структур сложно, и мы встречаемся с задачей обнаружения момента существенного изменения (скачка, насыщения и т.д.) критерия.

3 Результаты

На рис. 1 представлены результаты моделирования на примере изображения *house.tif*, полученнного из открытой базы изображений *http://imageprocessingplace.com/root_files_V3/image_databases.htm*. Средствами пакета МАТЛАБ изображение считано в память компьютера и насчитаны критерии Q_1 и Q_2 при размере маски 5×5. Визуально результаты применения двух критериев близки.

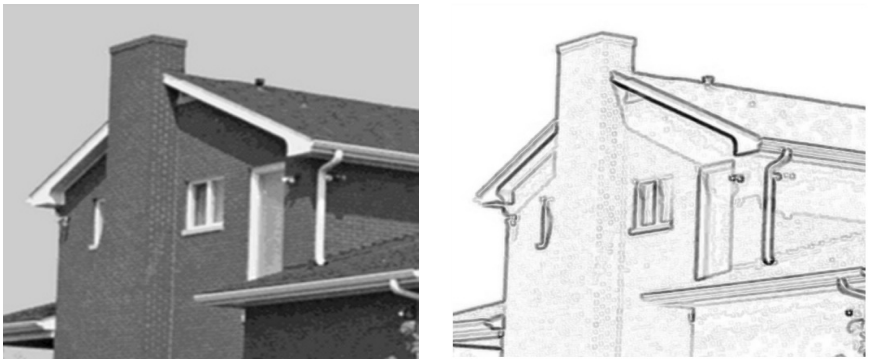


Рис. 1: Результаты моделирования. (Слева) Исходный снимок. (Справа) Визуализированы значения критерия Q_2 без применения порога. Доминантные точки, включая угловые структуры и прямолинейные границы, находятся в пикселах с более темной яркостью.

Заключение

В работе получены следующие результаты. Представлены матрицы масок для выделения углов, обладающие свойствами, удобными для конструирования иерархических масштабируемых детекторов. Структуры, потенциально выделяемые данным подходом, включают углы в 45, 90 и 135 градусов. Базовыми масками являются маски в 45 градусов, из которых возможно собирать детекторы углов с общей вершиной, детекторы линейных границ как полных углов в 180 градусов и другие. Эти конструкции не требуют повторного прохода массива изображения скользящим окном, поскольку комбинирование осуществляется на следующем этапе, на основе базисных процедур нижнего уровня.

Планируется провести сравнительный анализ и вычислительные эксперименты с предлагаемым и известными детекторами доминантных и угловых точек.

Список литературы

[1] Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006.

[2] Козловский А. Н. Детектор угловых точек на основе аппроксимации контуров объектов изображения // Информатика. 2010. V. 28, iss. 4. P. 36–47.

[3] Борисенко Д. И. Методы поиска угловых особенностей на изображениях // Молодой ученый. 2011. V. 1, iss. 5. P. 120–123.

[4] Shen F., Wang H. Corner detection based on modified Hough transform // Pattern Recognition Letters. 2002. V. 32, iss. 8. P. 1039–1049.

[5] Golightly I., Jones D. Corner Detection and Matching for Visual Tracking during Power Line Inspection // Image and Vision Computing. 2003. V. 21. P. 827–840.

[6] Gao X., Sattar F., Venkateswarlu R. Multiscale corner detection of gray level images based on loggabor wavelet transform // IEEE Transactions on Circuits and Systems for Video Technology. 2007. V. 17, iss. 7. P. 868–875.

[7] Dutta A., Kar A., Chatterji B. N. Corner Detection Algorithms for Digital Images in Last Three Decades // Institution of Electronics and Telecommunication Engineers Technical Review. 2008. V. 25, iss. 3. P. 123–133.

[8] Chen J., Zou L., Zhang J., Dou L. The Comparison and Application of Corner Detection Algorithms // Journal of Multimedia. 2009. V. 4, iss. 6. P. 435–441.

[9] Rosten E., Porte R., Drummond T. Faster and Better: A Machine Learning Approach to Corner Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010. V. 32, iss. 1. P. 105–119.

[10] Krig S. Computer Vision Metrics: Survey, Taxonomy, and Analysis. Berkely: Apress Berkely, 2014.

[11] Shekar B. H., Uma K. P. Kirsch directional derivatives based shot boundary detection: an e?cient and accurate method // Procedia Computer Science. 2015. V. 58. P. 565–571.

[12] Pietikainen M., Zhao G. Two decades of local binary patterns: A survey // Advances in Independent Component Analysis and Learning Machines. Academic Press, 2015. Chapter 9. P. 175–210.

[13] Peng W., Hongling X., Wenlin L., Wenlong S. Harris Scale Invariant Corner Detection Algorithm Based on the Significant Region // International Journal of Signal Processing, Image Processing and Pattern Recognition. 2016. V. 9, iss. 3. P. 413–420.

[14] Rivera A. R., Castillo J. R., Chae O. Local Directional Number Pattern for Face Analysis: Face and Expression Recognition // IEEE Transactions on Image Processing. 2013. V. 22, iss. 5. P. 1740–1752.

[15] Казанцев И. Г. Об одном детекторе угловых точек на изображениях // Труды XIV Междунар. научного конгресса "ИНТЕРЭКСПО ГЕО-Сибирь-2018 Т.1, "Дистанционные методы зондирования Земли и фотограмметрии, мониторинг окружающей среды, геоэкология". 2018. Новосибирск. P. 89–93.

[16] Sharpnack J. Learning Patterns for Detection with Multiscale Scan Statistics // Proceedings of Machine Learning Research. 2018. V. 75. P. 950-–969.

[17] Liu H., Tan S. Image Regularizations Based on the Sparsity of Corner Points // IEEE Transactions on Image Processing. 2019. V. 28, iss. 1. P. 72–87.

Казанцев Иван Гаврилович — д.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: kig@ooi.sscc.ru;
Мухаметжанова Бигуль Олжабаевна — докторант Евразийского национального университета им. Л.Н. Гумилева, физико-технический факультет, кафедра системного анализа и управления (Нур-Султан), Казахстан
e-mail: grek71@mail.ru.
Искаков Казизат Такуадинович — д.т.н., профессор факультета информационных технологий Евразийского национального университета им. Л.Н. Гумилева (Нур-Султан), Казахстан
e-mail: kazizat@mail.ru.
Дата поступления — 30 апреля 2019 г.

НЕКЛАССИЧЕСКИЕ ЗАДАЧИ В МОДЕЛЯХ ГЛОБАЛЬНОЙ ЭЛЕКТРИЧЕСКОЙ ЦЕПИ

А. В. Калинин[1], А. А. Тюхтина[1], С. Р. Лаврова[1]

[1] *Нижегородский государственный университет им. Н.И. Лобачевского, 603950, Нижний Новгород*

УДК 517.958
DOI: 10.24411/9999-016A-2019-10033

Обсуждаются различные постановки неклассических математических задач, возникающие при описании глобальной электрической цепи в атмосфере Земли. Приводятся результаты о корректности рассматриваемых постановок. Выделен класс задач об определении потенциальных электрических полей в нерелятивистском электрическом приближении для системы уравнений Максвелла с граничными условиями, учитывающими специфику взаимодействия атмосферы и ионосферы. Приводится и обосновывается новая постановка задач определения квазистационарных электромагнитных полей, охватывающая нерелятивистские электрическое и магнитное приближения. Устанавливается корректность этой задачи. Обсуждается связь между различными постановками задач в зависимости от физических параметров.

Ключевые слова: атмосферное электричество, глобальная электрическая цепь, квазистационарное приближение для системы уравнений Максвелла.

Введение

Под глобальной электрической цепью (ГЭЦ) понимают распределенный токовый контур, образованный высокопроводящими слоями верхнего слоя океана и земной коры и атмосферой, проводимость которой ничтожно мала в пограничном приземном слое, но экспоненциально растет с высотой [1, 2]. В последние годы существенно возрос интерес к вопросам математического, физического и численного моделирования глобальной электрической цепи, что отражено в современной литературе [1]– [22].

Наиболее важным компонентом ГЭЦ является так называемый квазистационарный ток, который, согласно гипотезе, выдвинутой Вильсоном [23, 24], поддерживается за счет постоянного разделения зарядов в грозовых и других электрифицированных облаках. Большинство существующих моделей ГЭЦ направлены на поиск распределения квазистационарной плотности тока и электрического поля в атмосфере с учетом генераторов ГЭЦ, соответствующих электрифицированным облакам. При этом используется нерелятивистское электрическое приближение для системы уравнения Максвелла [25], предполагающее потенциальность электрического поля и подходящее для описания медленно протекающих процессов в средах с малой проводимостью [4, 5, 19, 26, 27]. Однако реальная атмосфера Земли является существенно неоднородной средой, включающей в себя как области с малой проводимостью, так и области с достаточно высокой проводимостью. Использование нерелятивистских магнитного или электрического приближений не даёт возможности описать атмосферу в целом.

Для описания достаточно медленных процессов в неоднородных средах, включающих в себя области с высокой и малой проводимостью, в настоящей работе предлагается новое квазистационарное приближение для системы уравнений Максвелла, обобщающее классические нерелятивистское магнитное и нерелятивистское электрическое приближения. В работе приводятся постановки начально-краевых задач для системы уравнений Максвелла в различных квазистационарных приближениях. Рассматриваемые задачи относятся к неклассическим задачам математической физики [28]– [31].

ISBN 978-5-901548-42-4

1 Нерелятивистское электрическое приближение для квазистационарных процессов в атмосфере

Для описания электромагнитных полей используется система уравнений Максвелла, которая имеет в гауссовой системе единиц вид [25, 32, 33]

$$\operatorname{rot}\mathbf{H}(\mathbf{x},t)=\frac{4\pi}{c}\mathbf{J}(\mathbf{x},t)+\frac{1}{c}\frac{\partial\mathbf{D}(\mathbf{x},t)}{\partial t}, \tag{1}$$

$$\operatorname{rot}\mathbf{E}(\mathbf{x},t)=-\frac{1}{c}\frac{\partial\mathbf{B}(\mathbf{x},t)}{\partial t}, \tag{2}$$

$$\operatorname{div}\mathbf{B}(\mathbf{x},t)=0, \tag{3}$$

$$\operatorname{div}\mathbf{D}(\mathbf{x},t)=4\pi\rho(\mathbf{x},t). \tag{4}$$

Здесь $(\mathbf{x},t)\in\Omega\times(0,T)$, $\Omega\subset\mathbb{R}^3$, $T>0$, $\mathbf{H}$ — напряженность магнитного поля, $\mathbf{J}$ — объёмная плотность тока, $\mathbf{D}$ — электрическая индукция, $\mathbf{E}$ — напряженность электрического поля, $\mathbf{B}$ — магнитная индукция, ρ — плотность электрических зарядов, $\mathbf{J}^{\text{ext}}$ — объёмная плотность сторонних токов. В линейных средах справедливы материальные соотношения

$$\mathbf{D}(\mathbf{x},t)=\varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x},t),\qquad \mathbf{B}(\mathbf{x},t)=\mu(\mathbf{x})\mathbf{H}(\mathbf{x},t), \tag{5}$$

$$\mathbf{J}(\mathbf{x},t)=\sigma(\mathbf{x},t)\mathbf{E}(\mathbf{x},t)+\mathbf{J}^{\text{ext}}(\mathrm{x},t). \tag{6}$$

Предполагается, что область Ω диффеоморфна шаровому слою в $\mathbb{R}^3$ с липшиц-непрерывной границей Γ, состоящей из двух компонент связности Γ_1 и Γ_2, дифферeоморфных сфере в $\mathbb{R}^3$ и обозначающих, соответственно, поверхность Земли и поверхность, ограничивающую верхние слои атмосферы. В почти каждой точке $\mathbf{x}\in\Gamma$ определен единичный вектор внешней нормали $\nu(\mathbf{x})$.

При моделировании электромагнитных процессов в атмосфере Земли предполагается, что диэлектрическая проницаемость ε и магнитная проницаемость μ атмосферы постоянны и равны 1.

Поскольку проводимость Земли значительно выше проводимости нижних слоев атмосферы и проводимость растет по экспоненциальному закону с ростом высоты [1], можно полагать, что границы области Ω являются идеальными проводниками. Это соответствует заданию однородных граничных условий для касательной компоненты напряженности электрического поля:

$$\mathbf{E}(\mathbf{x},t)\times\nu(\mathbf{x})=0,\ (\mathbf{x},t)\in\Gamma\times(0,T). \tag{7}$$

При моделировании задач глобальной электрической цепи обычно рассматривается нерелятивистское электрическое приближение для системы уравнений Максвелла [2, 26, 27], заключающееся в пренебрежении слагаемым $\partial/\partial t\mathbf{B}$ в уравнении (2), что приводит к условию потенциальности напряженности электрического поля

$$\operatorname{rot}\mathbf{E}=0. \tag{8}$$

С учётом этого условия система сводится к уравнениям

$$\frac{\partial}{\partial t}\operatorname{div}\mathbf{E}+4\pi\operatorname{div}\sigma\mathbf{E}=-4\pi\operatorname{div}\mathbf{J}^{\text{ext}}, \tag{9}$$

$$\int_{\Gamma_1}\left(\frac{\partial}{\partial t}\mathbf{E}+4\pi\sigma\mathbf{E}\right)\cdot\nu d\gamma=-4\pi\int_{\Gamma_1}\mathbf{J}^{\text{ext}}\cdot\nu d\gamma. \tag{10}$$

Начальные условия для напряженности электрического поля $\mathbf{E}$ имеют вид

$$\mathbf{E}(\mathbf{x},0)=\mathbf{e}(\mathbf{x}),\ \mathbf{x}\in\Omega. \tag{11}$$

Математические задачи для квазистационарных электромагнитных полей в рамках нерелятивистского электрического приближения получили достаточно полное исследование в работах [2, 9, 10, 22]. Некоторые из полученных результатов будут сформулированы далее в теоремах 1.1–1.3.

Пусть $\mathbf{J}^{\text{ext}}\colon\Omega\times(0,T)\to\mathbb{R}^3$, $\mathbf{e}\colon\Omega\to\mathbb{R}^3$ — суммируемые с квадратом функции, σ — 3×3 матрица, элементами которой являются измеримые на $\Omega\times(0,T)$ функции, удовлетворяющая условиям

$$\sigma_1|\xi|^2\le(\sigma(\mathbf{x},t),\xi)\le\sigma_2|\xi|^2,\ (\mathbf{x},t)\in\Omega\times(0,T),\ \xi\in\mathbb{R}^3,$$

σ_1, σ_2 — заданные положительные числа.

Определим гильбертовы пространства вектор-функций с соответствующими скалярными произведениями [34, 35]:

$$H(\mathrm{div};\Omega)=\{\mathbf{u}\in\{L_2(\Omega)\}^3:\ \mathrm{div}\,\mathbf{u}\in L_2(\Omega)\},\ K(\mathrm{div};\Omega)=\{\mathbf{u}\in\{L_2(\Omega)\}^3:\ \mathrm{div}\,\mathbf{u}=0\},$$

$$(\mathbf{u},\mathbf{v})_{\mathrm{div}}=(\mathbf{u},\mathbf{v})_{2,\Omega}+(\mathrm{div}\,\mathbf{u},\mathrm{div}\,\mathbf{v})_{2,\Omega},$$

$$H(\mathrm{rot};\Omega)=\{\mathbf{u}\in\{L_2(\Omega)\}^3:\ \mathrm{rot}\,\mathbf{u}\in\{L_2(\Omega)\}^3\},\ K(\mathrm{rot};\Omega)=\{\mathbf{u}\in\{L_2(\Omega)\}^3:\ \mathrm{rot}\,\mathbf{u}=0\},$$

$$(\mathbf{u},\mathbf{v})_{\mathrm{rot}}=(\mathbf{u},\mathbf{v})_{2,\Omega}+(\mathrm{rot}\,\mathbf{u},\mathrm{rot}\,\mathbf{v})_{2,\Omega}.$$

Через $H_0(\mathrm{rot};\Omega)$, $H_0(\mathrm{div};\Omega)$ обозначается замыкание множества пробных вектор-функций $\{\mathcal{D}(\Omega)\}^3$ соответственно в $H(\mathrm{rot};\Omega)$ и $H(\mathrm{div};\Omega)$, $K_0(\mathrm{rot};\Omega)=K(\mathrm{rot};\Omega)\cap H_0(\mathrm{rot};\Omega)$. Для каждой функции из $H(\mathrm{div};\Omega)$ определена в смысле теории следов её нормальная составляющая на границе области, а для каждой функции из $H(\mathrm{rot};\Omega)$ — касательная составляющая. При этом $\mathbf{u}\cdot\nu|_\Gamma=0$ для $\mathbf{u}\in H_0(\mathrm{div};\Omega)$, $\mathbf{u}\times\nu|_\Gamma=0$ для $\mathbf{u}\in H_0(\mathrm{rot};\Omega)$. Ортогональным дополнением к $K_0(\mathrm{rot};\Omega)$ в $\{L_2(\Omega)\}^3$ является пространство [35]

$$K(\Omega)=\{\mathbf{u}\in K(\mathrm{div};\Omega):\langle\mathbf{u}\cdot\nu,1\rangle_{\Gamma_i}=0,i=1,2\}.$$

Задача определения вектор-функции $\mathbf{E}$, удовлетворяющей уравнениям (9), (10) и условиям (7), (11) допускает следующую постановку: найти функцию $\mathbf{E}\in H^1(0,T,K_0(\mathrm{rot};\Omega))$ такую, что выполнено (11) и для всех $\mathbf{u}\in K_0(\mathrm{rot};\Omega)$

$$\frac{1}{c}\frac{d}{dt}(\mathbf{E},\mathbf{u})_{2,\Omega}+\frac{4\pi}{c}(\sigma\mathbf{E},\mathbf{u})_{2,\Omega}=-\frac{4\pi}{c}(\mathbf{J}^{\mathrm{ext}},\mathbf{u})_{2,\Omega}. \tag{12}$$

Поля ($\mathrm{rot}\,\mathbf{H}$, $\mathbf{J}$, $\mathbf{D}$, ρ) могут быть найдены из равенств (1), (3)-(6).

Теорема 1.1. *[9,22] Для любых $\mathbf{e}\in K_0(\mathrm{rot};\Omega)$, $\mathbf{J}^{\mathrm{ext}}\in L_2(0,T,\{L_2(\Omega)\}^3)$ существует единственное решение задачи (12), (11).*

Поскольку область Ω пространственно односвязна, уравнение (8) позволяет ввести в рассмотрение электрический потенциал $\varphi(\mathbf{x},t)$ по формуле $\mathbf{E}=-\,\mathrm{grad}\,\varphi$. Уравнения (9), (10) принимают вид

$$\frac{\partial}{\partial t}\Delta\varphi+4\pi\,\mathrm{div}(\sigma\,\mathrm{grad}\,\varphi)=4\pi\,\mathrm{div}\,\mathbf{J}^{\mathrm{ext}}, \tag{13}$$

$$\langle(\mathrm{grad}\,\frac{\partial\varphi}{\partial t}+4\pi\sigma\,\mathrm{grad}\,\varphi)\cdot\nu,1\rangle_{\Gamma_1}=4\pi\langle\mathbf{J}^{\mathrm{ext}}\cdot\nu,1\rangle_{\Gamma_1}. \tag{14}$$

Условия (7) означают, что в каждый момент времени $t\in(0,T)$ функция $\varphi(\cdot,t)$ постоянна на компонентах связности границы. Не ограничивая общности, можно считать, что потенциал равен нулю на поверхности Земли:

$$\varphi(\mathbf{x},t)=0,\ (\mathbf{x},t)\in\Gamma_1\times(0,T),\ \exists U(t):\varphi(\mathbf{x},t)=U(t),\ (\mathbf{x},t)\in\Gamma_2\times(0,T). \tag{15}$$

Начальные условия имеют вид

$$\varphi(\mathbf{x},0)=\varphi_0(\mathrm{x}),\ \mathbf{x}\in\Omega, \tag{16}$$

где $\mathrm{grad}\,\varphi_0=-\mathbf{e}$, $\varphi_0(\mathbf{x})=0$, $\mathbf{x}\in\Gamma_1$, $\varphi_0(\mathbf{x})=\mathrm{const}$, $\mathbf{x}\in\Gamma_2$, U — неизвестная функция, зависящая только от t, удовлетворяющая условию $U(0)=\varphi_0|_{\Gamma_2}$.

В теории глобальной электрической цепи $U(t)$ принято называть ионосферным потенциалом [1]. Показано, что одновременное измерение ионосферного потенциала на границе атмосферы и ионосферы Γ_2 даёт близкие результаты [3], что является экспериментальным обоснованием рассматриваемых в работе постановок задач.

В задаче (13)–(16) функция $U(t)$ — неизвестная функция, подлежащая определению. Возможность аналитического определения ионосферного потенциала исходя из параметров задачи при некоторых упрощающих предположениях исследовалась, в частности, в работе [2].

Уравнение (13) не разрешено относительно производной по времени и не относится к классическим задачам математической физики. Уравнения такого вида называются уравнениями Соболева (уравнениями соболевского типа) или псевдопараболическими уравнениями [28]- [31].

Введём гильбертово пространство

$$V_D(\Omega)=\{u\in H^1(\Omega):u(\mathbf{x})=0,\mathbf{x}\in\Gamma_1,\ u(\mathbf{x})=\mathrm{const},\mathbf{x}\in\Gamma_2\},$$

$(u,v)_{V_D} = (u,v)_{H^1(\Omega)}$, $K_0(\mathrm{rot};\Omega) = \{\mathrm{grad}\, u, u \in V_D(\Omega)\}$.

Задача (13)-(16) допускает следующую постановку: найти функцию $\varphi \in H^1(0,T,V_D(\Omega))$, удовлетворяющую условию (16) и такую, что при всех $\psi \in V_D(\Omega)$

$$\frac{1}{c}\frac{d}{dt}(\mathrm{grad}\,\varphi, \mathrm{grad}\,\psi)_{2,\Omega} + \frac{4\pi}{c}(\sigma\,\mathrm{grad}\,\varphi, \mathrm{grad}\,\psi)_{2,\Omega} = \frac{4\pi}{c}(\mathbf{J}^{\mathrm{ext}}, \mathrm{grad}\,\psi)_{2,\Omega}. \tag{17}$$

Теорема 1.2. *[10] Для любых $\varphi_0 \in V_D(\Omega)$, $\mathbf{J}^{\mathrm{ext}} \in L_2(0,T,\{L_2(\Omega)\}^3)$ существует единственное решение задачи (17), (16).*

В работе [22] рассмотрен достаточно широкий спектр задач для уравнений (13), (14) с различными типами граничных условий, естественно возникающих при моделировании взаимодействия атмосферы и ионосферы: с условиями Дирихле, Неймана, смешанными условиями, граничными условиями в магнито-сопряженных точках. Приведём результаты, касающиеся последнего типа задач.

В моделях глобальной электрической цепи на верхней границе атмосферы иногда используется нестандарное граничное условие, соответствующее предположению, что в ионосфере заряженные частицы движутся вдоль силовых линий магнитного поля Земли, и эти силовые линии являются эквипотенциалями электрического поля [26, 27]. Если внешняя граница атмосферы симметрична относительно геомагнитного экватора, то каждая пара симметричных точек на Γ_2 лежит на одной и той же линии геомагнитного поля. Такие точки называются магнито-сопряженными, в них должен соблюдаться баланс электрических токов и равенство электрических потенциалов.

Пусть поверхность Γ_2 симметрична относительно некоторой плоскости, пересекающей Ω, и существует отображение $\delta : \mathbb{R}^3 \to \mathbb{R}^3$ такое, что $\delta(\Gamma_2) = \Gamma_2$. Очевидно, δ индуцирует линейный непрерывный оператор $\delta_* : H^{1/2}(\Gamma_2) \to H^{1/2}(\Gamma_2)$, $\delta_* u = u \circ \delta$, и двойственный к нему оператор $\delta^* : H^{-1/2}(\Gamma_2) \to H^{-1/2}(\Gamma_2)$, $\langle \delta^* f, u\rangle_{\Gamma_2} = \langle f, u \circ \delta\rangle_{\Gamma_2}$.

Пусть $\mathbf{J}^{\mathrm{ext}} \in L_2(0,T,H(\mathrm{div};\Omega))$, функции $\varphi_0 \in H^1(\Omega)$, $\varphi_1 \in H^1(0,T,H^{1/2}(\Gamma_1))$, $\varphi_2 \in H^1(0,T,H^{1/2}(\Gamma_2))$ и $j_2 \in L_2(0,T,H^{-1/2}(\Gamma_2))$ удовлетворяют условиям

$$\varphi_0|_{\Gamma_1} = \varphi_1|_{t=0},\ (\mathrm{id} - \delta_*)\varphi_0|_{\Gamma_2} = \varphi_2|_{t=0},\ \delta_*\varphi_2 = -\varphi_2,\ \delta^* j_2 = j_2,\ \langle j_2, 1\rangle_{\Gamma_2}\langle j_1, 1\rangle_{\Gamma_1} = 0,\ \langle j_2, 1\rangle_{\Gamma_2} = 0. \tag{18}$$

Рассматривается задача определения функции $\varphi \in H^1(0,T,H^1(\Omega))$ такой, что справедливы равенства (13), (14), начальное условие (16) и граничные условия

$$\varphi|_{\Gamma_1} = \varphi_1,\ (\mathrm{id} - \delta_*)\varphi|_{\Gamma_2} = \varphi_2, \tag{19}$$

$$(\mathrm{id} + \delta^*)(-\frac{1}{4\pi}\,\mathrm{grad}\,\frac{\partial\varphi}{\partial t} - \sigma\,\mathrm{grad}\,\varphi + \mathbf{J}^{\mathrm{ext}}) \cdot \nu|_{\Gamma_2} = j_2. \tag{20}$$

Вводится гильбертово пространство функций

$$V_{MC}(\Omega) = \{u \in H^1(\Omega) : u|_{\Gamma_1} = 0, \delta_* u|_{\Gamma_2} = u|_{\Gamma_2}\}.$$

Рассматриваемая задача допускает следующую вариационную постановку: найти такую удовлетворяющую условиям (15), (19) функцию $\varphi \in H^1(0,T,H^1(\Omega))$, что для всех $\psi \in V_{MC}(\Omega)$

$$\frac{d}{dt}(\mathrm{grad}\,\varphi, \mathrm{grad}\,\psi)_{2,\Omega} + 4\pi(\sigma\,\mathrm{grad}\,\varphi, \mathrm{grad}\,\psi)_{2,\Omega} = 4\pi(\mathbf{J}^{\mathrm{ext}}, \mathrm{grad}\,\psi)_{2,\Omega} - 2\pi\langle j_2, \psi\rangle_{\Gamma_2}. \tag{21}$$

Теорема 1.3. *[22] Для всех $\mathbf{J}^{\mathrm{ext}} \in L_2(0,T,\{L_2(\Omega)\})$, $\varphi_0 \in H^1(\Omega)$, $\varphi_1 \in L_2(0,T,H^{1/2}(\Gamma_1))$, $\varphi_2 \in L_2(0,T,H^{1/2}(\Gamma_2))$ и $j_2 \in L_2(0,T,H^{-1/2}(\Gamma_2))$, удовлетворяющих (18), существует единственное решение $\varphi \in H^1(0,T,H^1(\Omega))$ задачи (21), (16), (19). Если $\mathbf{J}^{\mathrm{ext}} \in L_2(0,T,H(\mathrm{div};\Omega))$, φ — решение задачи (13), (14), (16), (19), (20).*

2 Общая формулировка квазистационарного приближения для неоднородных сред

Экспоненциальный характер зависимости проводимости атмосферы от высоты и характерные временные и пространственные масштабы, возникающие при изучении задач атмосферного электричества, не позволяют ограничиться нерелятивистским электрическим или нерелятивистским магнитным приближением и

требуют для описания электромагнитных явлений построения более общей модели, не связанной с предположениями о характере проводимости. В данном пункте приводится формулировка такого квазистационарного приближения, охватывающего указанные классические нерелятивистские приближения. Его особенностью является сохранение в системе уравнений Максвелла потенциальной части тока смещения.

Вектор напряженности электрического поля однозначно представим в виде суперпозиции вихревой и потенциальной части [34, 35]:

$$\mathbf{E} = \mathcal{E} - \operatorname{grad}\varphi,\ \operatorname{div}\mathcal{E} = 0,\ \langle \mathcal{E}\cdot\nu, 1\rangle_{\Gamma_i} = 0,\ i = 1,2,\ \operatorname{grad}\varphi\times\nu|_{\Gamma} = 0.$$

Предлагаемое в работе для изучения существенно неоднородных сред квазистационарное приближение для системы уравнений Максвелла имеет вид

$$\operatorname{rot}\mathbf{H}(\mathbf{x},t) = \frac{4\pi}{c}\sigma(\mathbf{x})\mathbf{E}(\mathbf{x},t) + \frac{4\pi}{c}\mathbf{J}^{\mathrm{ext}}(\mathbf{x},t) - \frac{1}{c}\frac{\partial}{\partial t}\operatorname{grad}\varphi(\mathbf{x},t), \tag{22}$$

$$\operatorname{rot}\mathbf{E}(\mathbf{x},t) = -\frac{1}{c}\frac{\partial}{\partial t}\mathbf{H}(\mathbf{x},t). \tag{23}$$

Система рассматривается при граничных условиях

$$\mathbf{E}(\mathbf{x},t)\times\nu(\mathbf{x}) = 0,\ (\mathbf{x},t)\in\Gamma\times(0,T) \tag{24}$$

и начальных условиях

$$\mathbf{H}(\mathbf{x},0) = \mathbf{h}(\mathbf{x}),\ \operatorname{grad}\varphi(\mathbf{x},0) = \operatorname{grad}\varphi_0(\mathbf{x}),\ \mathbf{x}\in\Omega. \tag{25}$$

Введём гильбертовы пространства $V(\Omega) = H(\mathrm{rot};\Omega)\times K_0(\mathrm{rot};\Omega)$, $L(\Omega) = \{L_2(\Omega)\}^3\times\{L_2(\Omega)\}^3$, $U(\Omega) = H_0(\mathrm{rot};\Omega)\cap K(\Omega)$, $(\mathbf{u},\mathbf{v})_U = (\mathbf{u},\mathbf{v})_{\mathrm{rot},\Omega}$.

Задача (22)–(25) допускает следующую постановку: найти $\Psi = \{\mathbf{H}, \operatorname{grad}\varphi\}\in L_2(0,T,V(\Omega))$ и $\mathcal{E}\in L_2(0,T,U(\Omega))$ такие, что для всех $\Phi = \{\mathbf{u}, \operatorname{grad}\psi\}\in V(\Omega)$, $\mathbf{v}\in U(\Omega)$

$$\frac{1}{c}\frac{d}{dt}(\Psi,\Phi)_L + (\mathcal{E},\operatorname{rot}\mathbf{u})_{2,\Omega} - \frac{4\pi}{c}(\sigma\mathcal{E},\operatorname{grad}\psi)_{2,\Omega} + \frac{4\pi}{c}(\sigma\operatorname{grad}\varphi,\operatorname{grad}\psi)_{2,\Omega} = \frac{4\pi}{c}(\mathbf{J}^{\mathrm{ext}},\operatorname{grad}\psi)_{2,\Omega}, \tag{26}$$

$$(\sigma\mathcal{E},\mathbf{v})_{2,\Omega} - (\sigma\operatorname{grad}\varphi,\mathbf{v})_{2,\Omega} - \frac{c}{4\pi}(\mathbf{H},\operatorname{rot}\mathbf{v})_{2,\Omega} = -(\mathbf{J}^{\mathrm{ext}},\mathbf{v})_{2,\Omega} \tag{27}$$

и

$$\Psi(0) = \Psi_0 = \{\mathbf{h}, \operatorname{grad}\varphi_0\}. \tag{28}$$

Теорема 2.1. *Для любых $\Psi_0\in V(\Omega)$, $\mathbf{J}^{\mathrm{ext}}\in H^1(0,T,\{L_2(\Omega)\}^3)$ существует единственное решение задачи (26), (27), (28). При этом $\Psi\in L_\infty(0,T,L(\Omega))$, $\partial/\partial t\Psi\in L_2(0,T,L(\Omega))$ и справедливы соотношения (22), (23).*

Квазистационарные приближения для системы уравнений Максвелла обусловлены медленностью протекания процессов, то есть выполнением условия $\beta^2 = (\frac{\Delta x}{c\Delta t})^2 \ll 1$, где Δx — характерный пространственный масштаб, Δt — характерный временной масштаб.

Показано, что при малых значениях параметра β^2 решение задачи (22), (23), (27), (28) будет близко к решению соответствующей начально-краевой задачи для полной нестационарной системы уравнений Максвелла (1)–(6) при выполнении определенных условий согласования начальных данных.

Заметим, что на основании рассмотренного приближения могут быть получены нерелятивистское магнитное приближение (при условии высокой проводимости) и нерелятивистское электрическое приближение (если характерная проводимость мала).

Список литературы

[1] Мареев Е. А. Достижения и перспективы исследований глобальной электрической цепи // Успехи физических наук. 2010. Т. 180. N 5. С. 527–534.

[2] Калинин А. В., Слюняев Н. Н., Мареев Е. А., Жидков А. А. Стационарные и нестационарные модели глобальной электрической цепи: корректность, аналитические соотношения, численная реализация // Известия РАН. Физика атмосферы и океана. 2014. Т. 50, N 3. С. 314–322.

[3] Markson R. The global circuit intensity: its measurement and variation over the last 50 years // Bull. Am. Met. Soc. 2007. V. 88. N 2. P. 223–241.

[4] Rycroft M. J., Harrison R. G., Nicoll K. A., Mareev E. A. An overview of Earth's global electric circuit and atmospheric conductivity // Space Sci. Rev. 2008. V. 137. N 1–4. P. 83–105.

[5] Анисимов С. В., Мареев Е. А. Геофизические исследования глобальной электрической цепи // Физика Земли. 2008. N 10. С. 8–18.

[6] Mareev E. A., Yashunin S. A., Davydenko S. S. et al. On the role of transient currents in the global electric circuit // Geophys. Res. Lett. 2008. V. 35. N 15. L15810.

[7] Tinsley B. A. The global atmospheric electric circuit and its effects on cloud microphysics // Rep. Progr. Phys. 2008. V. 71. N 6. 066801.

[8] Williams E. R. The global electrical circuit: a review // Atmos. Res. 2009. V. 91. N 2–4. P. 140–152.

[9] Жидков А. А., Калинин А. В. Корректность одной математической задачи атмосферного электричества // Вестник Нижегородского госуниверситета. 2009. N 4. С. 122–129.

[10] Жидков А. А., Калинин А. В. Некоторые вопросы математического и численного модлирования глобальной электрической цепи в атмосфере // Вестник Нижегородского госуниверситета. 2009. Т 6. N 1. С. 150–158.

[11] Liu C., Williams E. R., Zipser E. J., Burns G. Diurnal variation of global thunderstorms and electrified shower clouds and their contribution to the global electrical circuit // J. Atmos. Sci. 2010. V. 67. N 2. P. 309–323.

[12] Zhou L., Tinsley B. A. Global circuit model with clouds // J. Atmos. Sci. 2010. V. 67. N 4. P. 1143–1156.

[13] Денисенко В. В., Помозов Е. В. Расчет глобальных электрических полей в земной атмосфере // Вычислительные технологии. 2010. Т. 15. N 5. С. 34–50.

[14] Денисенко В. В., Помозов Е. В. Проникновение электрического поля из приземного слоя атмосферы в ионосферу // Солнечно-земная физика. 2010. N 16. С. 70–75.

[15] Rycroft M. J., Harrison R. G. Electromagnetic atmosphere plasma coupling: the global atmospheric electric circuit // Space Sci. Rev. 2011. V. 168. N 1–4. P. 363–384.

[16] Mach D. M., Blakeslee R. J., Bateman M. G. Global electric circuit implications of combined aircraft storm electric current measurements and satellite-based diurnal lightning statistics // J. Geophys. Res. 2011. V. 116. N D5. D05201.

[17] Mallios S. A., Pasko V. P. Charge transfer to the ionosphere and to the ground during thunderstorms // J. Geophys. Res. 2012. V. 117. N A8. A08303.

[18] Williams E., Mareev E. Recent progress on the global electrical circuit // Atmos. Res. 2014. V. 135-136. P. 208–227.

[19] Davydenko S. S., Mareev E. A. Comments on the article by S.A. Mallios and V.P. Pasko "Charge transfer to the ionosphere and to the ground during thunderstorms"// J. Geophys. Res. 2014. V. 119. N 3. P. 2359–2362.

[20] Jansky J., Pasko V. P. Charge balance and ionospheric potential dynamics in timedependent global electric circuit model // J. Geophys. Res. Space Physics. 2014. V. 229. N. 12. P. 10184–10203.

[21] Bayona V., Flyer N., Lucas G. M., Baumgaertner A. J. G. A 3-D RBF-FD solver for modeling the atmospheric global electric circuit with topography (GEC-RBFFD v1.0) // Geosci. Model Dev. 2015. V. 8. N 10. P. 3007–3020.

[22] Kalinin A. V., Slyunyaev N. N. Initial-boundary value problems for the equations of the global atmospheric electric circuit // J. Math. Anal. Appl. 2017. V. 450. N 1. P. 112–136.

[23] Wilson T. R. Investigations on lightning discharges and on the electric field of thunderstorms // Phil. Trans. Roy. Soc. Lond. A. 1921. V. 221. P. 73–115.

[24] Wilson T. R. The electric field of a thundercloud and some of its effects // Proc. Phys. Soc. London. 1924. V. 37. P. 32D–37D.

[25] Толмачев В. В., Головин А. М., Потапов В. С. Термодинамика и электродинамика сплошной среды. М.: Изд-во МГУ, 1988.

[26] Roble R. G., Hays P. B. A quasi-static model of global atmospheric electricity: 2. Electrical coupling between the upper and lower atmosphere // J. Geophys. Res. 1979. V. 84. N A12. P. 7247–7256.

[27] Ogawa T. Fair-weather electricity // J. Geophys. Res. 1985. V. 90 N D4. P. 5951–5960.

[28] Соболев С. Л. Об одной новой задаче математической физики // Изв. АН СССР. Сер. Мат. 1954. Т. 18. N 1. С. 3–50.

[29] Демиденко Г. В., Успенский С. В. Уравнения и системы, не разрешенные относительно старшей производной. Новосибирск: Научная книга, 1998.

[30] Свешников А. Г., Альшин А. Б., Корпусов М. О., Плетнер Ю. Д. Линейные и нелинейные уравнения соболевского типа. М.: Физматлит, 2007.

[31] Showalter R. E., Ting T. W., Pseudoparabolic partial differential equations // SIAM J. Math. Anal. 1970. V. 1. N 1. P. 1-26.

[32] Ландау Л. Д., Лифшиц Е. М. Теоретическая физика. Том 8. Электродинамика сплошных сред. М.: Наука. Гл. ред. физ.-мат. лит., 1982.

[33] Тамм И. Е. Основы теории электричества. М.: Наука, 1989.

[34] Темам Р. Уравнения Навье-Стокса. Теория и численный анализ. М.: Мир, 1981.

[35] Girault V., Raviart P. Finite element methods for Navier–Stokes equations. N.Y.: Springler-Verlag, 1986.

Калинин Алексей Вячеславович — к.ф.-м.н., зав. кафедрой,
Институт информационных технологий, математики и механики,
Нижегородский государственный университет им. Н. И. Лобачевского;
e-mail: avk@mm.unn.ru;
Тюхтина Алла Александровна — к.ф.-м.н., доцент,
Институт информационных технологий, математики и механики,
Нижегородский государственный университет им. Н. И. Лобачевского;
e-mail: kalintmt@yandex.ru;
Лаврова Светлана Романовна — аспирант,
Институт информационных технологий, математики и механики,
Нижегородский государственный университет им. Н. И. Лобачевского;
e-mail: motivation-cv-lav@yandex.ru.
Дата поступления — 31 мая 2019 г.

MATHEMATICAL AND NUMERICAL MODELING OF OIL POLLUTION WASTE PROCESSING

G. B. Kalmenova, G. T. Balakaeva

Al-Farabi Kazakh National University, 050040, Almaty

UDC 510.58
DOI: 10.24411/9999-016A-2019-10034

В работе рассмотрена переработка нефтешлама для решения проблем по защите окружающей среды и улучшению экологической обстановки при разработке месторождений и транспортировке нефти. Неблагоприятное воздействие нефтесодержащих отходов на окружающую среду делает вопрос переработки нефтяных остатков весьма актуальным. Один из наиболее эффективный способ очистки нефтешламов — термическая обработка. В настоящей статье рассматривается задача термической переработки нефтешлама. Процессом термической переработки нефтесодержащих отходов необходимо управлять. В связи с этим потребовалось создание математической модели обработки нефтешлама. Математическая модель процесса описывается уравнениями тепломассообмена и включает систему нестационарных уравнений в частных производных второго порядка. Решение проводится по неявной разностной схеме методом переменных направлений до выполнения условия сходимости. В расчетах поставлены достаточные условия корректности и устойчивости прогонки. В этой исследовательской работе использовалось математическое и численное моделирование, и решение проводится с использованием современного языка программирования, такого как C/C ++, и результат дается с визуализацией

Ключевые слова: oil slime, thermal processing, mathematical model, method of alternating directions, environment.

Introduction

The anthropogenic impact on the environment causes its irreparable harm, and one of the major sources of environmental pollution is the oil and oil refining industry. Kazakhstan today is one of the largest countries in the extraction and processing of petroleum and petroleum products. Oil slime is a complex physico-chemical mixture, which consists of petroleum products, mechanical impurities (clay, metal oxides, sand) and water. The ratio of its constituent elements can be very different. The qualitative characteristics of oil slime at enterprises fit within the following limits:

- organic matter from 10 percent to 25 percent of the weight;
- mechanical impurities 5-30 percent of the weight;
- water from 50 percent to 70 percent of the weight.

Shipments are occurring during the technological process of oil refining, and also during the transportation and extraction of oil. These residues, which contain oil, are dangerous for the external environment. In the most general form, all oil slime can be divided into three main groups in accordance with the conditions of their formation: soil, bottom and reservoir type. The first are formed as a result of spills of petroleum products on the soil during manufacturing operations, or in emergency situations. Bottom sludges are formed due to oil spills settling to the bottom of reservoirs, and tank type sludge - during storage and transportation of petroleum products in containers of different designs [1].

By chemical and physical indicators of oil slime are extremely different from each other. The components of water, mechanical impurities and petroleum products in oil lime have an extremely wide range of ratios. The

ISBN 978-5-901548-42-4

physical properties of oil slime are also far from constant. Density can be either 830 kg / m3 or 1700 kg / m3, and the pour point varies from -3 to +50 degrees Celsius. Such a large range of physico-chemical properties of oil slime are obtained due to completely different environments of their appearance [2]. Recycling of oil products and oil slime is a problem that is relevant for modern Kazakhstan. A rational solution to this problem will have a positive impact on the ecology and the economy of our country.

Every year in our country during the processing or transportation of oil, as a result of natural spills and accidents, about 400 thousand tons of oil waste are generated, and the resources located in earthen barns are estimated at 4.5 million tons. It should be understood that the presence of such barns increases the risk of death of animals, pollution of groundwater and air. Soil pollution with oil, in addition to its direct impact, can lead to excessive accumulation of heavy metals in them - zinc, copper, lead, which adversely affects both the ecology of the region and the quality of life of people. In Kazakhstan, the urgency of this urgent task was highlighted for the first time in the Environmental Code of the Republic of Kazakhstan dated January 9, 2007. The direction of "green"growth and a low-carbon economy as a tool for sustainable development is laid down in the Strategy-2020, in the international initiatives of our country. The head of state constantly emphasizes that the wealth of the subsoil must be used with minimal damage to the environment [3].

At the moment as a result of disposal of slime is already getting a lot of useful products: commercial oil, fuel for heating plants, some building materials. There are also technology and special equipment for processing oil slime with recovery of residual oil and solid waste in materials for road construction. Huge amounts of raw materials provided in the recycling of oil slime, make it possible to produce large quantities of asphaltic concrete - durable road surface friction with improved performance and durability. Total can be divided into 3 main areas of application of slime:

- involvement in the boiler fuel;
- obtaining fuel components and preventive lubricants;
- production of building materials.

Nowadays, as we know the most common methods of disposal of oil slime can be classified into:

- mechanical;
- physical and chemical;
- thermal;
- biochemical;
- combined methods based on a combination of the above methods.

1 Thermal methods

The adverse impact of oil slime on the environment makes the issue of processing oil residue is very important. One of the most effective way to clean oil slime is heat treatment.

Both in foreign and domestic practice, the method of thermal disposal of oil slime is the most common. This method allows to neutralize the following types of oily waste:

- oily sludges resulting from wastewater treatment and liquid oil waste from treatment plants;
- sludge formed during the cleaning of tanks and process equipment;
- contaminated soil;
- oil slimes, which are complex multicomponent disperse systems formed as a result of the pistoning of the conduits or formed over time in barns;
- products from blowing dust collectors, oil separators and separators, characterized by a fairly homogeneous composition and high content of hydrocarbons, as well as used compressor and industrial oils.

Thermal methods of oil slime processing are based on the processes of thermal decomposition of petroleum products. As a result of the complete thermal decomposition of petroleum products, final degradation products are formed — CO2 and H2O [4].

The structure of oil slime is a physicochemical system that includes oil products, water and mineral additives (clay, sand, metal oxides, etc.). One of the reasons for the formation of reservoir oil slime is the interaction of oil products with moisture, oxygen, mechanical impurities and the material of the walls of the reservoir. The result of such interactions is the partial oxidation of the feed oil to form resin-like joints and corrosion of the tank walls. Entry into reservoirs with oil products of moisture and mechanical contaminants promotes the formation of water-oil emulsions and mineral dispersions. All oil slime differ in their physic-chemica0 l characteristics, which is due to the different composition of the raw materials, the environmental conditions. As a result of various studies, oil slime reservoir type impurities: hydrocarbons are from 5 to 90 percent, water from 1 to 70 percent, solid impurities from 0.8 to 65 percent have a wide range of ratios of oil, water, and mechanical. The method of processing depends directly on the amount of slime that is in oil products. Oil sludges have extremely diverse compositions and represent complex systems consisting of petroleum products, water, and a mineral portion (sand, clay, silt). The ratio of these components fluctuates over a very broad range. The organic materials on the average comprise from 10 to 56 percent; water, 30 to 85 percent; solids, 1-46 percent. Oil wastes can be arbitrarily divided according to the way they form into soil, resid, and tank types and can be divided according to phase state into liquid and solid oil sludges. By liquid sludges, we mean oil wastes in which the crude oil content is greater than 50-90 percent, and accordingly solid oil sludges are oil wastes in which the crude oil content does not exceed 50 percent, i.e., this is oil-contaminated soil [5].

Thermal systems are divided into two categories namely, thermal desorption and incineration. In thermal desorption, the oily wastes are normally heated in an inert atmosphere and the volatilized hydrocarbons and water are recovered and recycled into the main oil/water separation process. By comparison, the hydrocarbons are destroyed by heating to very high temperatures in the presence of air in thermal incineration processes. Thermal treatment processes are being coupled with current methods such as landfill, land farming, and land spreading for final residue disposal. Thermal treatment is the most efficient method for destruction of organics. It also significantly reduces the volume of inorganics such as metals and salts, and reduces their mobility so that the residue can be effectively disposed of. Thermal processes can handle oily drill cuttings, contaminated soils (resulting from spills, reclamation of contaminated waste sites), tank bottoms and other process wastes. In other words, thermal technology is best suited to process oil (preferably brine free) contaminated solids. Another technological method of thermal processing of oil slime is the pyrolysis process, carried out at 500 ° -550 ° C, in which combustible gases and a solid residue are obtained. This process is recommended for processing solid oil sludge with low humidity (no more than 1-3 percent). It is most economically acceptable, as it allows the organic part of the waste not to be converted into toxic products of combustion, but used as an additional fuel for waste incineration.

Pyrolysis is more environmentally friendly than incineration, since it allows the organic part of the waste not to be converted into toxic products of combustion, but used as an additional fuel for incineration of waste or to condense to produce by-products [6].

The goal of High-Temperature Reprocessing is to use heat for recovery of hydrocarbons from influent oily emulsions and sludges. Temperature influences the physical nature of hydrocarbons and water; hence, temperature affects the separation of oil-in-water emulsions. The fundamental principle of thermally enhanced separation of emulsions is to create conditions where four main benefits of heat are used:

- reduction of oil viscosity;
- increased molecular movement of droplets, which aids coalescence through increased collision frequency of the dispersed-phase droplets;
- heat deactivation of the emulsifier, such as the dissolution of paraffins;
- and increased difference in density between oil and water [7].

The method of heat treatment for the purpose of sludge dewatering is becoming more widely used abroad.

2 Mathematical model of the problem

Numerical heat transfer is a broad term denoting the procedures for the solution, on a computer, of a set of algebraic equations that approximate the differential (and, occasionally, integral) equations describing conduction,

convection and/or radiation heat transfer. The usual objective in any heat transfer calculation is the determination of the rate of heat transfer to or from some surface or object. In conduction problems, this requires finding the temperature gradient in the material at its surface. In convection problems, the temperature gradient in a fluid flowing over a surface is needed to find the heat flux at that surface. In both cases, the determination of the complete temperature distribution in the region of interest is needed as a first step, and in convection one must also find the velocity distribution.

The equations describing heat transfer are complex, having some or all of the following characteristics: they are nonlinear; they comprise algebraic, partial differential and/or integral equations; they constitute a coupled system; the properties of the substances involved are usually functions of temperature and may be functions of pressure; the solution region is usually not a simple square, circle or box; and it may (in problems involving solidification, melting, etc.) change in size and shape in a manner not known in advance. Thus numerical methods, leading to exact, closed form solutions, are almost always not available [8].

When we develop a mathematical model of thermal processing of oil slime, consider the heating of the sludge layer. We consider a non-stationary, two-dimensional problem with constant flow rates in a rectangular area. The mathematical model of the process is described by heat and mass transfer equations and includes a system of nonstationary second-order partial differential equations. The solution is carried out according to an implicit difference scheme using the alternating directions method until the convergence condition is satisfied. In the calculations, sufficient conditions for the correctness and stability of the sweep were set, such as the condition of the diagonal dominance of the matrix and the condition in which the modulus of the average driving coefficient is greater than the sum of the modules of the other driving coefficients.

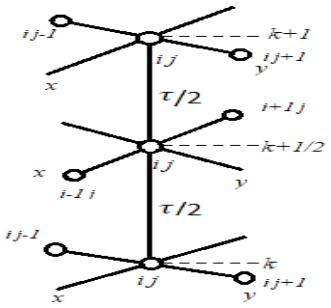


Рис. 1: Pattern diagrams of the method of alternating directions

There are various methods for solving boundary value problems for the heat equation, including the variable direction method. In the scheme of the method of alternating directions (MAD), as in all splitting methods, the time step is divided into the number of independent spatial variables (in the two-dimensional case, into two). At each fractional time layer, one of the spatial differential operators is approximated implicitly (scalar sweeps are performed in the corresponding coordinate direction), and the others explicitly. At the next fractional step, the next in order differential operator is approximated implicitly, and the others explicitly, etc [9, 10].

The implicit scheme of variable directions (longitudinal-transverse scheme) expresses quite clearly this algorithmic idea. This scheme is often called the scheme of the Pisman–Reckford (by the name of the authors who first proposed it) [11]. In the picture 1 can be shown that in the two-dimensional case the MAD scheme is absolutely stable. The advantages of the method of variable directions include high accuracy, since the method has the second order of accuracy in time. The disadvantages include conditional stability when the number of spatial variables is more than two. In addition, the MAD is conditionally stable in problems with mixed derivatives already in the two-dimensional case.

Graphical result of this task shown in below picture:

Conclusion

The amount of waste from the refining industry is increasing every year. The harmful effects of sludge waste affects environmental factors that pose a threat to human life and health, or a threat to the life or health of future generations. Therefore, the main objective of the task was the development of heat treatment of oil slime

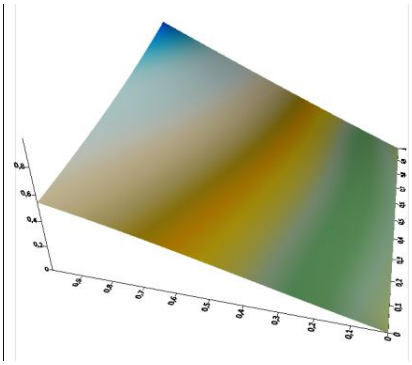


Рис. 2: The graphical result

by numerical methods to protect the environment. The goals were achieved. The resulting mathematical model of thermal processing of oil slime characterizes the temperature change in the stream over time and makes it possible to manage.

Список литературы

[1] A.F. Ahmetov, A.R. Gaysina, I.A. Mustafin. Metody utilizacii nefteshlamov razlichnogo proishojdenii// Prikladnye i academicheskiye issledovanye. — 2011. —9.— 3. — 98–101

[2] https://vtorothody.ru/utilizatsiya/nefteshlamov.html

[3] S. Temirgalyev. Prevrawaya othody v dohod // Kazakhstanskaya Pravda— 2015.— 30 jan.— 15p

[4] https://neftok.ru/pererabotka/utilizatsiya-nefteproduktov.html

[5] P. R. Khaidarov and S. K. Kudratova, Molodoi Uchenyi, No. 11, 125-127 (2014). I. R. Khairudinov, Crude Oil Waste and Sludge Treatment Methods, Khimiya, Moscow (1989). 425 pp.

[6] http://ru-ecology.info/term/48585/

[7] Naghi Jadidi , Behrooz Roozbehani , Akram Saadat. The Most Recent Researches in Oily Sludge Remediation Process // American Journal of Oil and Chemical Technologies: Volume 2. Issue 10. —Octoober 2014

[8] http://www.thermopedia.com/content/991/

[9] Samarskii A.A., Gulin A.V. Chislennye metody matematicheskoi fiziki. — 2—e izd. -M.: Nauchnyi mir. — 2003.—316 p.

[10] Shan Zhao. A Matched Alternating Direction Implicit (ADI) Method for Solving the Heat Equation with Interfaces// Journal of Scientific Computing. — April,2015. —Volume 63—Issue 1 —pp 118-137

[11] Samarskii A.A. The Theory of Difference Schemes — USA, Marcel Dekker, Inc.—2001—788p

Balakayeva Gulnar Tultaevna, doctor of phys.- mathematical sciences, professor of the Department of Computer Science, Faculty of Information Technology, Al-Farabi Kazakh National University ; e-mail: gulnardtsa@gmail.com;

Kalmenova Gaukhar Bolatbekovna, Phd student of the Department of Computer Science, Faculty of Information Technology, Al-Farabi Kazakh National University ; e-mail: kalmenova.g.b@gmail.com;

Received — June 1, 2019

ЧИСЛЕННОЕ РЕШЕНИЕ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ ТРЕХМЕРНЫХ ЗАДАЧ ДИРИХЛЕ ДЛЯ УРАВНЕНИЯ ГЕЛЬМГОЛЬЦА С ИСПОЛЬЗОВАНИЕМ МОЗАИЧНО-СКЕЛЕТОННОГО МЕТОДА

А. А. Каширин, С. И. Смагин, М. Ю. Тимофеенко

Вычислительный центр ДВО РАН, 680000, Хабаровск

УДК 519.642
DOI: 10.24411/9999-016A-2019-10035

Рассматриваются трехмерные задачи Дирихле для уравнения Гельмгольца. Они сводятся к граничным слабо сингулярным интегральным уравнениям Фредгольма первого рода. Эти уравнения приближаются системами линейных алгебраических уравнений, которые затем решаются обобщенным методом минимальной невязки. Вычислительная сложность решения снижается за счет использования мозаично-скелетонного метода. Приводятся результаты вычислительных экспериментов.

Ключевые слова: задача Дирихле, уравнение Гельмгольца, интегральное уравнение, мозаично-скелетонный метод.

Введение

Задачи Дирихле для уравнения Гельмгольца являются классическими задачами математической физики. Их аналитические решения могут быть найдены лишь в случае простейших областей, допускающих разделение переменных. Вследствие этого, основным инструментом исследования таких задач является компьютерное моделирование.

В данной работе задачи Дирихле формулируются в виде слабо сингулярных граничных интегральных уравнений Фредгольма первого рода. Дискретные аналоги указанных уравнений могут быть построены различными способами. Мы используем для этого согласованный с шагом сетки на границе области специальный метод осреднения слабо сингулярных ядер интегральных операторов. Он позволяет вычислять коэффициенты систем линейных алгебраических уравнений (СЛАУ), аппроксимирующих интегральные уравнения, по весьма простым формулам [1].

Матрицы получаемых СЛАУ являются плотно заполненными. Поэтому вычислительная сложность решения этих систем прямыми методами имеет оценку $O(n^3)$, где n — порядок системы. Однако спектральные свойства матриц таковы, что использование для поиска приближённых решений обобщённого метода минимальных невязок (GMRES) [2] позволяет понизить эту сложность до $O(mn^2)$. Здесь m — число итераций, которое значительно меньше n и слабо от него зависит. Отметим, что такая оценка обусловлена многократным применением в GMRES процедуры матрично-векторного умножения.

В последние десятилетия для СЛАУ с плотными матрицами развиваются различные «быстрые» методы решения. Называя метод быстрым, обычно имеют в виду, что его сложность составляет $o(n^2)$, при $n \to \infty$. В последнее время этот термин используют по отношению к алгоритмам почти линейной по n сложности, то есть сложности $O(n \log^a n)$, с некоторым $a > 0$. Для приближённых методов чаще всего отдельно

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-01-00682) и Программы фундаментальных исследований ДВО РАН (код проекта 18-5-100).

ISBN 978-5-901548-42-4

выделяют зависимость сложности от требуемой точности ε, обычно также логарифмическую, указывая $O(n\log^a n\log^b \varepsilon^{-1})$, $b>0$.

В настоящей работе для приближённого решения СЛАУ нами был использован один из таких методов — мозаично-скелетонный метод [3]– [5]. При его реализации основная матрица СЛАУ полностью не вычисляется и не сохраняется, а в процедуре матрично-векторного умножения используется некоторое её приближение. При определённых предположениях относительно свойств ядра интегрального оператора для построения такого приближения достаточно вычислить относительно небольшое число элементов исходной матрицы.

Мы приводим постановку задачи, описание численного метода и результаты численных экспериментов, которые позволяют судить об эффективности применяемого подхода.

1 Постановка исходной задачи

1.1 Обобщённая постановка задачи Дирихле

Рассмотрим трёхмерное евклидово пространство $\mathbb{R}^3$ с ортогональной системой координат $ox_1x_2x_3$. Пусть в этом пространстве имеется произвольная замкнутая липшицева поверхность Γ, разделяющая его на внутреннюю область Ω_i и внешнюю область $\Omega_e=\mathbb{R}^3\backslash\overline{\Omega}_i$.

Сформулируем исходную задачу [6].

Задача 1.1 (внешняя задача Дирихле). Найти функцию $u_e(x)\in H^1(\Omega_e)$, удовлетворяющую уравнению Гельмгольца

$$\Delta u_e+k^2u_e=0,\quad x\in\Omega_e, \tag{1}$$

граничному условию

$$\gamma u_e=f,\quad x\in\Gamma, \tag{2}$$

и условию излучения на бесконечности

$$\partial u_e/\partial|x|-iku_e=o\left(|x|^{-1}\right),\quad |x|\to\infty. \tag{3}$$

Здесь $\Delta=\nabla^2$ — оператор Лапласа, $\nabla=(\partial/\partial x_1,\partial/\partial x_2,\partial/\partial x_3)$, k — волновое число, $\mathrm{Im}(k)\geq 0$, γu_e — след на Γ функции u_e, $f\in H^{1/2}(\Gamma)$ — известная функция. Определения используемых здесь и далее функциональных пространств имеются в работе [6].

Замечание. Если $\mathrm{Im}(k)=0$, то $u_e\in H^1_{\mathrm{loc}}(\Omega_e)$.

Справедлива следующая теорема [6].

Теорема. Для любой функции $f\in H^{1/2}(\Gamma)$ существует единственное решение внешней задачи 1.1 из пространства $H^1(\Omega_e)$.

Здесь и далее $\langle\cdot,\cdot\rangle_\Gamma$ — отношение двойственности на $H^{1/2}(\Gamma)\times H^{-1/2}(\Gamma)$, обобщающее скалярное произведение в $H^0(\Gamma)$, $Nu\in H^{-1/2}(\Gamma)$ — нормальная производная u, понимаемая в смысле распределений [6].

1.2 Интегральная формулировка задачи 1.1

Решение задачи 1.1 будем искать в виде потенциала простого слоя

$$u_e(x)=(Sq)(x)\equiv\langle G(x,\cdot),q\rangle_\Gamma,\quad x\in\Omega_e, \tag{4}$$
$$G(x,y)=\exp(ik|x-y|)/(4\pi|x-y|).$$

Ядром интегрального оператора (4) является фундаментальное решение уравнения Гельмгольца, поэтому u_e удовлетворяет уравнению (1) и условию излучения (3). Эта функция будет решением задачи 1.1, если подобрать плотность q так, чтобы u_e удовлетворяла граничному условию (2). Таким образом, задача 1.1 сводится к граничному тождеству

$$\langle Sq,\mu\rangle_\Gamma=\langle f,\mu\rangle_\Gamma\ \forall\mu\in H^{-1/2}(\Gamma). \tag{5}$$

Теорема [4]. Пусть $\mathrm{Im}(k)>0$ или k^2 не является собственным значением задачи

$$\Delta u+k^2u=0,\quad x\in\Omega_i,\qquad \gamma u=0,\quad x\in\Gamma.$$

Тогда уравнение (5) корректно разрешимо в пространстве $H^{-1/2}(\Gamma)$ и формулы (4) дают решения задачи 1.1.

2 Дискретизация интегрального уравнения

Построим покрытие поверхности Γ системой $\{\Gamma_m\}_{m=1}^M$ окрестностей узловых точек $x'_m \in \Gamma$, лежащих внутри сфер радиусов h_m с центрами в x'_m, и обозначим через $\{\varphi_m\}$ подчинённое ему разбиение единицы. В качестве φ_m будем использовать функции

$$\varphi_m(x) = \varphi'_m(x)\left(\sum_{k=1}^{M}\varphi'_k(x)\right)^{-1}, \qquad \varphi'_m(x) = \begin{cases} \left(1 - r_m^2/h_m^2\right)^3, & r_m < h_m, \\ 0, & r_m \geq h_m, \end{cases}$$

где $x \in \Gamma$, $r_m = |x - x'_m|$, $\varphi_m \in C^1(\Gamma)$ при $\Gamma \in C^{r+\beta}$, $r + \beta > 1$.

Приближённые решения уравнения (5) будем искать на сетке $\{x_m\}$,

$$x_m = \frac{1}{\bar{\varphi}_m}\int\limits_\Gamma x\varphi_m d\Gamma, \quad \bar{\varphi}_m = \int\limits_\Gamma \varphi_m d\Gamma,$$

узлами которой являются центры тяжестей функций φ_m. Будем предполагать, что для всех $m = 1, 2, \ldots, M$ выполняются неравенства

$$0 < h' \leq |x_m - x_n|, \;\; m \neq n, \;\; n = 1, 2, \ldots, M,$$

$$h' \leq h_m \leq h, \;\; h/h' \leq q_0 < \infty,$$

где h, h' — положительные числа, зависящие от M, q_0 не зависит от M.

Вместо заданной на Γ неизвестной функции q будем искать обобщённую функцию $q\delta_\Gamma$, действующую по правилу

$$(q\delta_\Gamma, \eta)_{\mathbb{R}^3} = \langle q, \eta\rangle_\Gamma \qquad \forall \eta \in H^1(\mathbb{R}^3).$$

Приближать эту функцию будем выражением

$$q(x)\,\delta_\Gamma(x) \approx \sum_{n=1}^{M} q_n\bar{\varphi}_n\psi_n(x),$$

где q_n — неизвестные коэффициенты, $\psi_n(x) = (\pi\sigma_n^2)^{-3/2}\exp\left(-(x - x_n)^2/\sigma_n^2\right)$, $\sigma_n^2 = 0.5\bar{\varphi}_n$.

Тогда, используя результаты [1], получаем

$$\langle Sq, \varphi_m\rangle_\Gamma \approx \sum_{n=1}^{M} S_{mn}q_n, \quad m = 1, 2, ..., M, \tag{6}$$

$$S_{mn} = \frac{\beta_{mn}}{2r_{mn}}\exp\left(-\gamma_{mn}^2\right)\left(w(z_1) - w(z_2)\right), \quad m \neq n, \quad \beta_{mn} = \frac{\bar{\varphi}_m\bar{\varphi}_n}{4\pi\left(1 - \mu_{mn}^2 + 0.5\mu_{mn}^4\right)}, \tag{7}$$

$$S_{mm} = \beta_{mm}\left(2\pi^{1/2}\left(\bar{\varphi}_m\right)^{-1}\sigma_{mm}\left(1 - 2\mu_{mm}^2\Big/3\right) + \frac{2}{\pi^{1/2}\sigma_{mm}} + i\kappa w\left(\mu_{mm}\right)\right),$$

$$z_1 = \mu_{mn} - i\gamma_{mn}, \quad z_2 = \mu_{mn} + i\gamma_{mn}, \quad i^2 = -1, \quad \mu_{mn} = 0.5k\sigma_{mn},$$

$$w(z) = -\frac{2i}{\pi^{1/2}}\exp\left(-z^2\right)\int\limits_z^\infty \exp\left(t^2\right)dt.$$

Решая СЛАУ

$$\sum_{n=1}^{M} S_{mn}q_n = \bar{\varphi}_m f_m, \quad m = 1, 2, ..., M,$$

где

$$f_m = \frac{1}{\bar{\varphi}_m}\int\limits_\Gamma f\varphi_m d\Gamma,$$

найдём приближённые значения коэффициентов q_n. После этого решения исходных задач могут быть найдены по формулам

$$u_e(x) = \sum_{n=1}^{M} \int_{\Gamma_n} G(x,y) q_n \varphi_n(y) d\Gamma_y, \quad x \in \Omega_e,$$

в «ближней» зоне, и по формулам

$$u_e(x) = \sum_{n=1}^{M} G(x,x_n) q_n \bar{\varphi}_n, \quad x \in \Omega_e,$$

в «дальней» зоне.

3 Мозаично-скелетонный метод

Мозаично-скелетонный метод является одним из методов приближённого быстрого матрично-векторного умножения. Он состоит из трёх этапов: построение дерева кластеров, нахождение списка блоков и поиск мозаично-скелетонной аппроксимации. Для описания мозаично-скелетонного метода нам понадобятся следующие определения [3].

Пусть A_k — некоторая подматрица (блок) матрицы $A^{n\times m}$, а $\Pi(A_k)$ — матрица размера $n \times m$, полученная из A_k путём дополнения до A нулями.

Конечное множество блоков $\{A_k\}$ будем называть покрытием A, если

$$A = \sum_k \Pi(A_k),$$

и мозаичным разбиением A, если, дополнительно, $\bigcap\limits_k A_k = \varnothing$.

Мозаичным рангом матрицы $A \in \mathbb{C}^{n\times m}$, соответствующим покрытию $\{A_k\}$, будем называть число

$$\mathrm{mr}(A) = \sum_k \mathrm{mem}(A_k)/(n+m),$$

где сумма берётся по всем блокам покрытия $A_k \in \mathbb{C}^{n_k \times m_k}$, и $\mathrm{mem}\,(A_k) = \min\{n_k m_k,\ (n_k + m_k)\,\mathrm{rank}\ A_k\}$.

Отметим, что мозаичный ранг характеризует как вычислительную сложность матрично-векторного умножения, так и объём памяти, необходимый для хранения покрытия матрицы A.

Будем называть скелетоном матрицу вида uv^*, где u и v — вектор-столбцы, v^* обозначает вектор-строку, эрмитово-сопряжённую к вектору v. Из последнего определения следует, что $\mathrm{rank}\,(uv^*) = 1$.

На первом этапе мозаично-скелетонного метода все точки дискретизации объединяются в кластер нулевого уровня (корень дерева). Например, как в работе [7], кластером нулевого уровня может быть куб со стороной α. Создание дерева кластеров происходит следующим образом: каждое ребро куба делится пополам, при этом получается 8 подкубов. Точки, принадлежащие исходному кубу, распределяются по подкубам, причём некоторые подкубы могут оказаться пустыми. Далее на каждом следующем шаге таким же образом делятся подкубы, и в итоге получается иерархическое разбиение исходного куба. Определим k-й уровень разбиения L_k как множество кубов с длиной стороны $a_k = \alpha 2^{-k}$; таких кубов будет, очевидно, 2^{3k}. Центр куба k-го уровня, содержащего заданную точку x (или y), будем обозначать символом ξ_k (или η_k), а сам куб — символом $Q_k(\xi_k)$ (или $Q_k(\eta_k)$). Для всякого куба k-го уровня с центром η определим области (или «зоны») [7]

$$\begin{aligned} N_k(\eta) &= \{Q_k(\xi) \in L_k : \quad ||\xi - \eta||_\infty \leqslant \varkappa a_k\}, \\ F_k(\eta) &= \{Q_k(\xi) \in L_k : \quad 2(\varkappa+1)a_k \geqslant ||\xi - \eta||_\infty \geqslant (\varkappa+1)a_k\}, \end{aligned} \tag{8}$$

где $\varkappa$ — некоторое натуральное число. Зоны N_k будем называть «ближними», F_k — «дальними».

На втором этапе мозаично-скелетонного метода на основе дерева кластеров создаётся список блоков. Каждый блок определяется взаимодействием пары кластеров. Если для пары кластеров выполняется первое из условий (8), то блок попадает в ближнюю зону, если выполняется второе условие — в дальнюю зону.

Наконец, на третьем этапе метода осуществляется поиск мозаично-скелетонных аппроксимаций для блоков дальней зоны, блоки ближней зоны при этом рассчитываются по формулам (6)–(7). Малоранговая аппроксимация может быть построена разными способами. В данной работе применяется неполная крестовая аппроксимация [4].

Критерий остановки алгоритма из [4] сформулирован в виде

$$||A_k - A^{(p)}||_F \leqslant \varepsilon ||A^{(p)}||_F, \tag{9}$$

где A_k — блок дальней зоны размерности $n_k \times m_k$, $A^{(p)}$ — приближённая матрица, найденная по методу неполной крестовой аппроксимации, ε — относительная погрешность аппроксимации, $||\cdot||_F$ — норма Фробениуса.

Однако на практике используется более выгодный с вычислительной точки зрения критерий [4]

$$(l-p)||u_p||_F||v_p||_F/|A_{i_p j_p}| \leqslant \varepsilon \sum_{\alpha=1}^{p} \frac{||u_\alpha||_F||v_\alpha||_F}{|A_{i_\alpha j_\alpha}|}, \tag{10}$$

здесь u_p — столбец матрицы $A^{(p)}$, v_p — строка матрицы $A^{(p)}$, $A_{i_p j_p}$ — элемент, стоящий на пересечении u_p и v_p, $l = \min(n_k, m_k)$.

Рекуррентная формула (10) требует всего $O(m_k + n_k)$ операций на каждом шаге и не слишком ухудшает оценку точности мозаично-скелетонной аппроксимации (9).

4 Численные эксперименты

Программное обеспечение для численного решения трёхмерных задач Дирихле для уравнения Гельмгольца создано на языке Фортран 90 и представляет собой консольное приложение, предназначенное для работы на многопроцессорных вычислительных системах. В качестве компилятора выступает Intel Fortran Compiler, также используются библиотека Intel Math Kernel Library и реализация стандарта OpenMP для данного компилятора. Для выполнения расчётов были использованы вычислительные ресурсы ЦКП "Центр данных ДВО РАН" [8].

Для демонстрации возможностей предлагаемого подхода численно решались задачи, имеющие известные аналитические решения. Количество точек дискретизации M изменялось от 1032 до 128496, уровень дерева кластеров до 16033 узловых точек включительно был равен 4, далее — 5. Блоки дальней зоны аппроксимировались с точностью $\varepsilon = 1.6 \cdot 10^{-6}$, блоки ближней зоны вычислялись по формулам (6)–(7). Точность GMRES составляла 10^{-8}, $\varkappa = 1$.

Пример. Рассматриваются внешние задачи Дирихле для эллипсоида с полуосями $(0.75, 1, 0.5)$ и центром в точке $(0, 0, 0)$, граничные условия имеют вид

$$u_e(x) = \exp(ik\,|x - x_0|)/|x - x_0|,$$

волновые числа $k_1 = 1, k_2 = 14, k_3 = 30$, $x_0 = (0.375, 0.5, 0.25)$.

На рисунке 1 приведены погрешности решений внешних задач Дирихле, вычисленные в норме пространства $H_h^0(\Omega_e)$. Видно, что при удвоении количества точек погрешности уменьшаются в два раза, т.е. метод имеет второй порядок точности относительно $h^2 \sim M^{-1}$.

Вычислим теперь мозаичный ранг и фактор сжатия для рассматриваемых задач. Их значения приведены в таблицах 1 и 2. Фактор сжатия задаётся следующей формулой:

$$I = \frac{\mathrm{mem}(A^{(p)})}{\mathrm{mem}(A)} 100,$$

где $\mathrm{mem}(A^{(p)})$ — объём памяти, необходимой для хранения матрицы в малоранговом формате, $\mathrm{mem}(A)$ — объём памяти для хранения исходной матрицы. Видно, что с ростом числа узловых точек мозаичный ранг растёт как $O(\ln^3(M))$, а фактор сжатия, начиная с некоторого M, уменьшается как $O(\ln^3(M)/M)$. Это значит, что эффективность применения мозаично-скелетонного метода с ростом M существенно возрастает.

Заключение

В работе изучены возможности применения мозаично-скелетонного метода и его параллельного варианта для численного решения трёхмерных задач Дирихле для уравнения Гельмгольца. Эти задачи сформулированы в виде эквивалентных им граничных интегральных уравнений Фредгольма первого рода. Данный метод реализован на этапе численного решения СЛАУ, аппроксимирующих интегральные уравнения. Он сокращает требования к ресурсам компьютера и позволяет значительно снизить общее время решения рассматриваемых задач без потери точности.

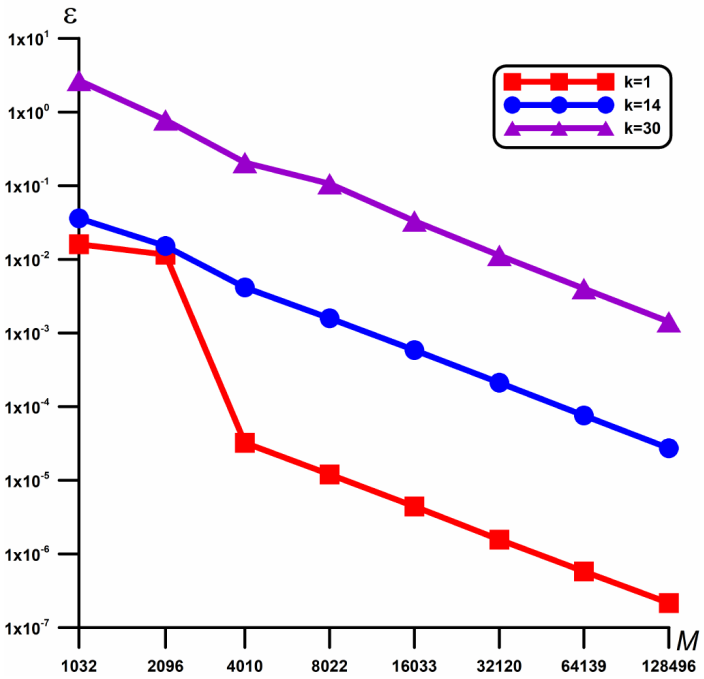


Рис. 1: Погрешности решений задач 1.1

Таблица 1: Мозаичный ранг для аппроксимаций интегральных операторов.

$k\backslash M$	1032	2096	4010	8022	16033	32120	64139	128496
1	482.4	739.4	988.1	1259.9	1560.2	1921.6	2360.9	2889.0
14	516.0	920.7	1286.5	1633.6	1982.5	2378.8	2845.2	3399.4
30	516.0	983.4	1387.9	1765.8	2119.8	2496.8	2933.4	3428.5

Таблица 2: Фактор сжатия для аппроксимаций интегральных операторов.

$k\backslash M$	1032	2096	4010	8022	16033	32120	64139	128496
1	93.5	70.6	49.3	31.4	19.5	12.0	7.4	4.5
14	100.0	87.9	64.2	40.7	24.7	14.8	8.9	5.3
30	100.0	93.8	69.2	44.0	26.4	15.6	9.2	5.3

Список литературы

[1] Каширин А. А., Смагин С. И. О численном решении задач Дирихле для уравнения Гельмгольца методом потенциалов // ЖВМиМФ. 2012. Т. 52, № 8. С. 1492–1505.

[2] Saad Y., Schultz M. GMRES: a general minimal residual algorithm for solving nonsymmetric linear systems // SIAM J. Sci Stat. Comput. 1986. V. 7, iss. 3. P. 856–869.

[3] Tyrtyshnikov E. E. Incomplete cross approximations in the mosaic-skeleton method // Computing. 2000. V. 64, iss. 4. P. 367–380.

[4] Каширин А. А., Смагин С. И., Талтыкина М. Ю. Применение мозаично-скелетонного метода при численном решении трехмерных задач Дирихле для уравнения Гельмгольца в интегральной форме // ЖВМиМФ. 2016. Т. 56, № 4. С. 625–638.

[5] Каширин А. А., Талтыкина М. Ю. О существовании мозаично-скелетонных аппроксимаций дискретных аналогов интегральных операторов // ЖВМиМФ. 2017. Т. 57, № 9. С. 1421–1432.

[6] McLean W. Strongly elliptic systems and boundary integral equations. Cambridge: Cambridge University Press, 2000.

[7] Горейнов С.А. Мозаично-скелетонные аппроксимации матриц, порожденных асимптотически гладкими и осцилляционными ядрами // Матричные методы и вычисления. М.: ИВМ РАН, 1999. С. 42–76.

[8] Сорокин А.А., Макогонов С.В., Королев С.П. Информационная инфраструктура для коллективной работы ученых Дальнего Востока России // Научно-техническая информация. Серия 1: Организация и методика информационной работы. 2017. № 12. С. 14-16.

Каширин Алексей Алексеевич — к.ф.-м.н.,
ст. науч. сотр. Вычислительного центра ДВО РАН;
e-mail: elomer@mail.ru;
Смагин Сергей Иванович — чл.-корр. РАН,
гл. науч. сотр. Вычислительного центра ДВО РАН;
e-mail: smagin@ccfebras.ru;
Тимофеенко Мария Юрьевна — мл. науч. сотр.
Вычислительного центра ДВО РАН;
e-mail: maria.timofeenko@ya.ru.
Дата поступления — 30 апреля 2019 г.

ОЦЕНКА СИСТЕМАТИЧЕСКОЙ ОШИБКИ ДАННЫХ СПУТНИКОВОГО ЗОНДИРОВАНИЯ О ТЕМПЕРАТУРЕ ПОВЕРХНОСТИ ЗЕМЛИ

Н. В. Киланова[1], Д. Л. Чубаров[1], Е. Г. Климова[1]

[1] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

УДК 551.509.313
DOI: 10.24411/9999-016A-2019-10036

Данные спутникового зондирования используются в настоящее время при решении большого круга задач, возникающих при моделировании процессов в окружающей среде. Важным преимуществом этих данных является непрерывное распределение по времени и высокое пространственное разрешение. В то же время, спутниковые данные обладают рядом недостатков. Это, в частности, отсутствие информации при наличии облачности. Кроме того, спутниковые данные имеют систематическую ошибку (bias), что затрудняет проведение исследований с использованием этих данных.

В докладе рассматривается алгоритм оценки систематической ошибки данных MODIS о температуре поверхности Земли. Информация о температуре поверхности Земли, в том числе климатическая, очень важна для проведения исследований в различных областях (моделирование, агрометеорология, почвоведение и т.д.). Алгоритм основан на 1-мерном фильтре Калмана и позволяет по временному ряду наблюдений оценивать систематическую ошибку спутниковых данных. В алгоритме используются данные измерительной станции (Зотино). В докладе приводятся результаты численных экспериментов с различными вариантами алгоритма по оценке среднемесячной систематической ошибки.

Ключевые слова: спутниковые данные, систематическая ошибка, фильтр Калмана.

Введение

Данные спутникового зондирования используются в настоящее время при решении большого круга задач, возникающих при моделировании процессов в окружающей среде. Важным преимуществом этих данных является непрерывное распределение по времени и высокое пространственное разрешение. В то же время, спутниковые данные обладают рядом недостатков. Это, в частности, отсутствие информации при наличии облачности. Кроме того, спутниковые данные имеют систематическую ошибку (bias), что затрудняет проведение исследований с использованием этих данных.

Температуру поверхности Земли является важным параметром для моделей климата, а также для приложений в различных отраслях экономики, в первую очередь — в сельском хозяйстве. Спутниковые измерения позволяют измерить значение температуры в любой точке поверхности один или несколько раз в сутки с пространственным разрешением от нескольких километров до нескольких сотен метров. На точность определения температуры влияет точность калибровки спутникового прибора, состояние атмосферы на пути излучения от участка поверхности до спутникового прибора, а также особенности подстилающей поверхности. Для валидации спутниковых измерений необходимы сопоставимые наземные измерения, которые бы позволили исключить часть помех.

Цель работы состоит в оценке точности восстановления температуры поверхности Земли по данным дистанционного зондирования. Полученные оценки могут быть использованы в дальнейшем для уточнения динамики температуры поверхности Земли на больших территориях с высоким пространственным разрешением. В докладе рассматривается алгоритм оценки систематической ошибки данных MODIS о температуре поверхности Земли. Алгоритм основан на 1-мерном фильтре Калмана и позволяет по временному ряду наблюдений оценивать систематическую ошибку спутниковых данных. В алгоритме используются данные

ISBN 978-5-901548-42-4

измерительной станции (Зотино). В докладе приводятся результаты численных экспериментов с различными вариантами алгоритма по оценке среднемесячной систематической ошибки.

1 Получение временного ряда данных LST по данным MODIS и измерительной станции

Для оценки точности спутниковых данных требуется наличие данных измерений, обладающих известной погрешностью и не имеющих систематической ошибки. Один из способов получения сопоставимых со спутниковыми данными наземных измерений состоит в использовании измерений интенсивности уходящего длинноволнового излучения, выполняемых в рамках глобальной сети мониторинга потоков на нижней границе атмосферы FLUXNET [8]. Создание сети FLUXNET началось в 1995 году при поддержке NASA как объединение нескольких региональных сетей. К 2003-му году в сеть входило 216 башен, оснащенных измерительным оборудованием, в том числе, 8 башен, расположенных на территории Сибири, оборудованных за счёт средств проектов рамочной программы научных исследований Европейского союза. Использование этих данных позволило получить оценки среднемесячной температуры поверхности Земли по спутниковым данным и по данным наземных наблюдений [2].

Приборы, установленные на башнях FLUXNET, измеряют величину интенсивности потока уходящего длинноволнового излучения L_{OUT} в единицах Вт/м^2. Для дальнейшей обработки значения L_{OUT} пересчитываются в температуры по формуле Стефана-Больцмана

$$T = \left(\frac{L_{OUT}}{\varepsilon\sigma}\right)^{\frac{1}{4}},$$

где σ — постоянная Стефана-Больцмана, а ε — излучательная способность, которая для поверхности Земли варьируется в интервале 0.85–1.0. Измеренные значения L_{OUT} осредняются по времени в пределах 30-минутных интервалов. На станции Зотино в Красноярском крае измерения велись с 2002-го до демонтажа оборудования в сентябре 2004-го года. В рамках настоящего исследования данные, соответствующие второму уровню обработки, были получены с разрешения EFDC — европейской подсети проекта FLUXNET.

Данные спутникового зондирования были загружены с использованием инструмента Google Earth Engine [6]. В работе использованы значения температуры поверхности Земли соответствующие продукту 11A1 коллекции 6 в номенклатуре информационных продуктов MODIS [9], представляющие собой значения температуры приведенные к ячейкам регулярной сетки в синусоидальной проекции с разрешением 1 км. Данные получены спектрорадиометрами MODIS, установленными на космических аппаратах Terra и Aqua и отнесены к ночным или дневным в соответствии с локальным солнечным временем. Загруженные данные были осреднены по площадке содержащей 5x5 ячеек с центром в ячейке с координатой 60° 48' 2.88" с.ш., 89° 21' 2.88" в.д., что соответствует координатам измерительной станции.

Для оценки систематической ошибки спутниковых данных LST были сформировпаны два временных ряда. Спутниковые данные MODIS и данные станции Зотино за 2004 год были перенесены на регулярную временную шкалу с интервалами продолжительностью 6 часов с центрами в 03, 09, 15 и 21 час местного времени и приведены к общим единицам измерения температур.

2 Методика оценки систематической ошибки

Известно, что спутниковые данные обладают систематической ошибкой. Если представить данные наблюдений в виде

$$y_0 = x_t + \varepsilon_0,$$

где x_t — «истинное» значение измеряемой переменной, ε_0 — случайная ошибка измерений, то систематической ошибкой считается $b = \overline{\varepsilon_0}$, черта сверху обозначает теоретико-вероятностное осреднение.

Для оценки систематической ошибки спутниковых данных о температуре поверхности Земли (LST) MODIS будем использовать данные измерений станции Зотино. Рассмотрим разность $d = y_{sat} - y_b$, здесь y_b и y_{sat} — данные на измерительной станции и спутниковые данные, проинтерполированные в точку наблюдений на станции. Эту разность можно проедставить в виде $d \approx \varepsilon_{sat} - \varepsilon_b$, ε_{sat} и ε_b — ошибки наблюдений спутниковых данных и данных на станции, при этом будем считать, что $\overline{\varepsilon_b} = 0$. В этом случае может быть

осуществлена оценка систематической ошибки спутниковых данных (bias) по временному ряду наблюдений с помощью алгоритма, рассмотренного в работах [3,5]. Оценка производится отдельно для дневных и ночных наблюдений.

Алгоритм оценки систематической ошибки представляет собой 1-мерный фильтр Калмана [7]. Пусть $y^n = y^n_{sat} - y^n_b$ считаются данными наблюдений о "bias". Зададим начальное значение "bias" $b^0 = 0$. Алгоритм оценки работает циклически по времени $n = 1, \ldots, N$:

$$b^{n+1} = b^n + k_n(y^n - b^n),$$

$$k_n = \frac{\sigma^2_{f,n}}{\sigma^2_{f,n} + \sigma^2_{o,n}},$$

$$\sigma^2_{a,n} = (1 - k_n)\sigma^2_{f,n},$$

$$\sigma^2_{f,n+1} = \sigma^2_{a,n} + q^2_n.$$

Для работы алгоритма требуется:

- задание $\sigma^2_{f,0}$ (в начальный момент времени задается равным 4);
- задание среднеквадратических ошибок наблюдений $\sigma^2_{o,n}$ (в начальный момент времени задается равным 1);
- задание параметра q^2_n (его можно задать равным $\sigma^2_{o,n}$ или оценивать дополнительно по данным).

По оцененному значению b^n можно производить уточнение спутниковых данных на n-м шаге по времени.

Оценка параметров алгоритма может быть произведена также по временному ряду наблюдений. Значение $\sigma^2_{o,n}$ — среднеквадратической ошибки наблюдений и q^2_n — ошибки «модели» могут быть оценены по приведенным в [4] формулам:

$$\overline{\left(y^0_n - x^f_n\right)\left(y^0_n - x^a_n\right)^T} = R_n,$$

$$\overline{\left(y^0_n - x^f_n\right)\left(y^0_n - x^f_n\right)^T} = P^f_n + Q_n + R_n.$$

В этих формулах x^f_n и x^a_n — шаги прогноза и анализа фильтра Калмана (в точках наблюдений) [7], R_n — матрица ковариаций ошибок наблюдений, P_n — матрица ковариаций ошибок прогноза, Q_n — матрица ковариаций шумов модели, n — номер шага по времени. В рассматриваемом нами 1-мерном случае R_n соответствует $\sigma^2_{o,n}$, $P_n - \sigma^2_{f,n}$, $Q_n - q^2_n$.

Обозначим $r_i = y_i - b_i$, $r1_i = y_i - b_{i+1}$ (i — номер шага про времени), тогда можно вычислить оценки

$$O1 = \frac{1}{N-1}\sum_{i=1}^{N}\left(r_i - \frac{1}{N}\sum_{k=1}^{N} r_k\right)^2,$$
$$O2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(r_i - \frac{1}{N}\sum_{k=1}^{N} r_k\right)\left(r1_i - \frac{1}{N}\sum_{k=1}^{N} r1_k\right)$$

В приведенных формулах вероятностное осреднение заменено на осреднение по времени, $O1$ — это оценка $\sigma^2_{f,n} + \sigma^2_{o,n} + q^2_n$, $O2$ — оценка $\sigma^2_{o,n}$. В статьях [3, 5] при осреднении по времени рекомендуется взять $N = 7$. Вместо осреднения по времени можно также рассмотреть среднее по близлежащим станциям, получая среднюю для этих станций оценку.

3 Численные эксперименты

Для проведения численных экспериментов были взяты данные станции высотной мачты (Зотино) за 2004 год. Данные представляют собой плотный массив без пропусков, значения температуры поверхности Земли приведены с интервалом 30 минут начиная с 00 часов 1 января 2004г. Эти данные были условно разделены на дневные и ночные значения температур. К ночным значениям были отнесены данные с 21 часа 00

минут одних суток до 8 часов 30 минут следующих суток. "Ночная"температура — это среднее арифметическое значение температур из указанного промежутка времени. К дневным значениям были отнесены данные с 9 часов 00 минут до 20 часов 30 минут. "Дневная"температура получена как среднее арифметическое значение температур из указанного промежутка. В результате каждые сутки имеется одно значение ночной температуры и одно значение дневной температуры. Данные спутникового зондирования MODIS представляют собой массив с пропусками значений температуры поверхности Земли (имеющиеся данные приведены в 00, 06, 12, 18 часов суток). Аналогично данным станции высотной мачты имеющиеся данные спутникового зондирования были условно разделены на дневные и ночные значения температур. В результате такого разделения каждые сутки имеется одно значение ночной температуры (если в 00 либо 06 часов данных суток имелись данные наблюдений) и одно значение дневной температуры (если в 12 либо 18 часов этих суток имелись данные наблюдений). Для вычисления систематической ошибки спутниковых данных были взяты данные июля — один из месяцев 2004 года с наибольшим количеством данных наблюдения MODIS. Поскольку в дальнейшем предполагается проводить оценку среднемесячных значений LST, важной является оценка среднемесячного значения систематической ошибки. Как отмечается в [1], для того, чтобы осредненная по времени оценка отражала климатическое значение, следует удалить из временного ряда синоптическую составляющую. Для этого используется метод скользящего среднего с интервалом осреднения порядка 7 суток. Для данных станции высотной мачты и данных MODIS по методу скользящего среднего были найдены сглаженные значения температур для каждых суток этого месяца. Результаты применения этого метода для данных станции высотной мачты приведены на рис. 1. Результаты применения метода скользящего среднего для данных наблюдения MODIS представлены на рис. 2.

а б

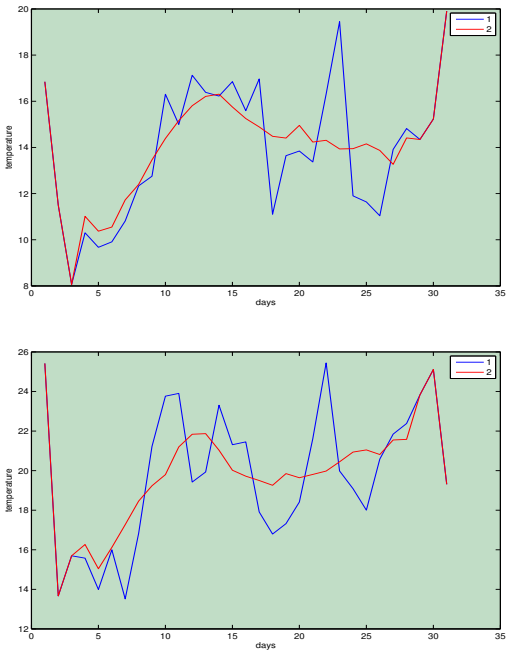


Рис. 1: Осредненная температура поверхности Земли (данные на станции Зотино) для ночных значений (а) и дневных значений (б) без сглаживания (1) и со сглаживанием (2).

Как сказано выше, данные спутникового зондирования имеют пропуски, что можно увидеть на рис. 2. Даты тех суток, в которые данные отсутствуют, отображены на оси абсцисс. Метод скользящего среднего для данных с пропусками характеризуется тем, что осреднение происходит для имеющихся значений из интервала сглаживания 7 суток. В итоге в каждый интервал может попасть разное количество данных (от 0 до 7). Результат сглаживания для имеющихся ночных температур приведен на графике 2 рис. 2(а), для имеющихся дневных температур — на графике 2 рис. 2(б).

Оценка систематической ошибки спутниковых данных проводилась как для сглаженных по методу скользящего среднего значений температуры, так и для не сглаженных значений. На рис.3 приведен график оценки систематической ошибки данных MODIS для июля 2004 года. Систематическая ошибка вычисля-

а б

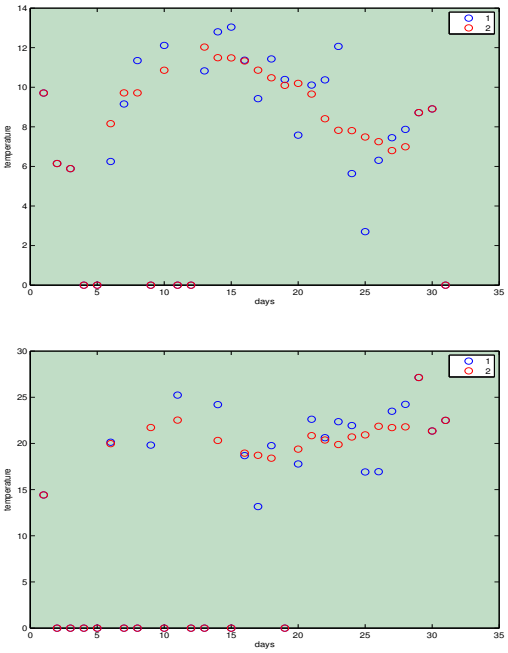


Рис. 2: Осредненная температура поверхности Земли (спутниковые данные) для ночных значений (а) и дневных значений (б) без сглаживания (1) и со сглаживанием (2).

лась отдельно для ночных и для дневных средних температур. В численных экспериментах использовался начальный набор параметров ($\sigma^2_{f,0} = 4$, $\sigma^2_{o,n} = 1$, $q^2_n = \sigma^2_{o,n}$). В дальнейшем предполагается провести численные эксперименты с оценкой параметров по формулам, приведенным в пункте 2.

Из рис. 1 и рис. 2 видно, что процедура сглаживания меняет характер поведения значений температуры. Из рис. 3 можно увидеть, что для ночных температур (а) и дневных температур (б) характер поведения систематической ошибки разный. Для дневных значений она близка к нулю. Тогда как для ночных значений имеется выраженный отрицательный тренд. В результате проведенных численных экспериментов можно сделать вывод, что характер систематической ошибки может быть различным для данных, полученных в дневное время и в ночное время суток. Эту особенность предполагается в дальнейшем учитывать для корректировки данных спутниковых наблюдений.

Заключение

В работе рассматривается алгоритм оценки систематической ошибки данных MODIS о температуре поверхности Земли. Алгоритм основан на 1-мерном фильтре Калмана и позволяет по временному ряду наблюдений оценивать систематическую ошибку спутниковых данных. В алгоритме используются данные измерительной станции (Зотино). Приводятся результаты численных экспериментов с различными вариантами алгоритма по оценке среднемесячной систематической ошибки. Показано отличие значений систематической ошибки для данных в дневное и ночное время а также влияние синоптической составляющей на значение систематической ошибки. Результаты могут быть в дальнейшем использованы для уточнения границ ареалов растений и изучения динамики их изменения в связи с изменениями климата.

Благодарности

В настоящей работе использованы данные измерений сети TCOS-Siberia, входящей в сеть FLUXNET. Были использованы данные, собранные в рамках проекта Ru-Zot (руководитель проекта — Коринна Ребманн), предварительно обработанные в рамках European Fluxes Database Cluster.

а б

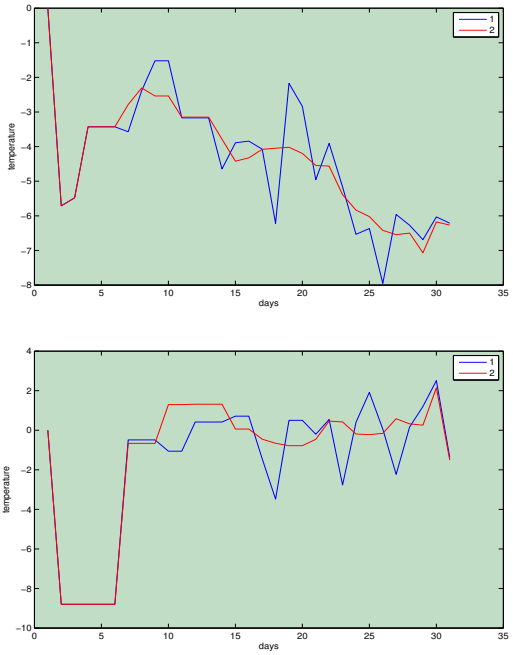


Рис. 3: Систематическая ошибка данных спутникового наблюдения поверхности земли для ночных значений (а) и дневных значений (б) без сглаживания (1) и со сглаживанием (2).

Список литературы

[1] Груза Г.В., Рейтенбах Р.Г. Статистика и анализ гидрометеорологических данных. 1982. Ленинград. Гидрометеоиздат. 216 с.

[2] Chen X., Su Z., Ma Y., Cleverly J., Liddell M. An accurate estimate of monthly mean land surface temperatures from MODIS clear-sky retrievals. Journal of Hydrometeorology // 2017. V. 18, P. 2827–2847.

[3] Cheng, W. Y. Y., and W. J. Steenburgh Strengths and weaknesses of MOS, run-ning-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States // Wea. Forecasting. 2007. V. 22, P. 1304–1318.

[4] Dezroziers G., Berre L., Chapnic B., Poli P. Diagnosis of observation, background and analysis-error statistics in observational space. Quarterly Journal of the Royal Meteorological Society // 2005. V. 131, P. 3385–3396.

[5] Galanis, G., and M. Anadranistakis A one-dimensional Kalman filter for the correction of near surface temperature forecasts // Meteorol. Appl. 2002. V. 9, P. 437–441.

[6] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone // Remote Sensing of Environment, Volume 202, 2017, Pages 18–27.

[7] Jazwinski, A.H. Stochastic processes and filtering theory. Academic Press, New York, 1970.

[8] Olson, R.J., Holladay, S.K., Cook, R.B., Falge, E., Baldocchi, D. FLUXNET: Database of Fluxes, Site Characteristics, and Flux-Community Information // Oak Ridge National Laboratory Tech. Report ORNL/TM-2003/204, 2004.

[9] Wan, Z., Hook, S., Hulley, G. MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006 // NASA EOSDIS Land Processes DAAC, 2015, doi://10.5067/MODIS/MOD11A1.006

Киланова Наталья Владимировна — к.ф.-м.н., науч. сотр. Института вычислительных технологий СО РАН;
e-mail: kilanova@ict.sbras.ru
Чубаров Дмитрий Леонидович — науч. сотр. Института вычислительных технологий СО РАН;
e-mail: dchubarov@ict.nsc.ru
Климова Екатерина Георгиевна — д.ф.-м.н., ст. науч. сотр. Института вычислительных технологий СО РАН;
e-mail: klimova@ict.nsc.ru
Дата поступления — 30 апреля 2019 г.

КОМБИНАТОРНАЯ СТРУКТУРА МОДЕЛЕЙ КОЛЬЦЕВЫХ ГЕННЫХ СЕТЕЙ

Н. Е. Кириллова[1]

[1] *Институт математики им. С. Л. Соболева СО РАН, 630090, Новосибирск*

УДК 514.745.82
DOI: 10.24411/9999-016A-2019-10037

Разработан алгоритм построения инвариантных областей в фазовых портретах нелинейных динамических систем, моделирующих функционирование кольцевых генных сетей. Актуальность этих построений связана с тем, что поиск периодических траекторий таких динамических систем можно сузить до рассмотрения указанных инвариантных областей.

Ключевые слова: кольцевые генные сети, существование циклов, стационарная точка, инвариантная область, фазовое пространство.

Введение

В недавних публикациях [1,4,6,8] рассматривались вопросы существования периодических траекторий нелинейных динамических систем, моделирующих функционирование широкого класса генных сетей. Принципы построения таких математических моделей описаны в [9]. В настоящей работе проведена дискретизация фазовых портретов указанных динамических систем, с помощью которой удается исследовать периодические траектории этих систем. Наличие таких осциллирующих режимов функционирования генных сетей является важной характеристикой происходящих в них биохимических процессов.

1 Модель

Мы рассматриваем модель генной сети в виде динамической системы размерности n:

$$\frac{dx_i}{dt} = f_i(x_{i-1}) - k_i x_i, \quad i = \overline{1,n}. \tag{1}$$

Здесь функции $f_i(x_{i-1})$ — гладкие функции положительного аргумента, которые либо строго монотонно убывают (и моделируют отрицательные обратные связи), либо строго монотонно возрастают (и моделируют положительные обратные связи). Пусть $A_i = \dfrac{\max f_i(x_{i-1})}{k_i}$.

Рассмотрим случай, когда количество убывающих функций f_i в правых частях уравнений системы (1) нечётно. В дальнейшем будем обозначать через L_i убывающие функции f_i, а монотонно возрастающие функции f_i — через Γ_i. Построим в положительном октанте $\mathbb{R}^n_+$ параллелепипед Q^n следующим образом:

$$Q^n := \prod_{i=1}^{i=n} [0, A_i].$$

Следующие две леммы для ряда систем вида (1) были установлены в [1,6].

Лемма 1. *Q^n является инвариантной областью системы* (1).

Лемма 2. *Система* (1) *имеет единственную стационарную точку S_0.*

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 18-01-00057).

ISBN 978-5-901548-42-4

Для описания поведения траекторий системы (1) разобьем инвариантный параллелепипед Q^n гиперплоскостями, проходящими через точку S_0, на 2^n более мелких параллелепипедов, которые мы будем называть блоками и нумеровать бинарными мультииндексами:

$$\mathcal{E} = \{\varepsilon_1 \dots \varepsilon_n\} = \left\{\mathbf{X} \in Q^n \mid x_1 \gtrless_{\varepsilon_1} x_1^0;\ x_2 \gtrless_{\varepsilon_2} x_2^0;\ \dots;\ x_n \gtrless_{\varepsilon_n} x_n^0 \right\}, \tag{2}$$

где $\mathbf{X} = (x_1, \dots, x_n)$, $\varepsilon_1, \dots, \varepsilon_n \in \{0, 1\}$, и отношения порядка задаются следующим образом: символ $\gtrless_0$ означает $\leq$, а символ $\gtrless_1$ означает $\geq$.

Мы говорим, что *валентность блока* $\mathcal{E} = \{\varepsilon_1, \dots, \varepsilon_n\}$ равна ℓ, если число его соседних блоков $\mathcal{E}_j$, в которые переходят траектории точек $\mathcal{E}$, равно ℓ.

Построим диаграмму переходов траектории системы (1) по блокам валентности 1.

2 Алгоритм построения блоков диаграммы

Для начала нужно построить блок валентности 1. Это означает, что из него переход будет осуществляться только в один соседний блок, при таком переходе в наборе его бинарных индексов изменится всего один индекс.

1 шаг. Пусть первый индекс начального блока равен нулю, то есть набор его индексов выглядит следующим образом: $\mathcal{E}_1 = \{0\,\varepsilon_2 \dots \varepsilon_n\}$. Чтобы соблюсти условие одновалентности, построим начальный блок так, чтобы изменилась, например, только первая пара индексов $\{0\,\varepsilon_2\}$, а остальные индексы остались неизменными. Для этого и для всех последующих рассуждений воспользуемся леммой:

Лемма 3 (Казанцев [7]). *Возможные переходы из блока в блок для систем, описываемых схемой* (1)*, можно построить, пользуясь правилом: для убывающих функций любая пара чисел* $\{00\}$ *переходит в* $\{01\}$*, а пара* $\{11\}$ *переходит в* $\{10\}$*; для возрастающих функций наоборот* $-$ $\{01\} \to \{00\}, \{10\} \to \{11\}$.

2 шаг. Далее возможны две ситуации:

а) если второе уравнение системы (1) содержит убывающую функцию, то для того, чтобы первая пара индексов поменялась, нужно, чтобы $\varepsilon_2 = 0$;

б) если второе уравнение системы (1) содержит возрастающую функцию, то для того, чтобы первая пара индексов поменялась, нужно, чтобы $\varepsilon_2 = 1$.

3 шаг. Чтобы блок оставался одновалентным, то есть чтобы из него больше не было никаких переходов, нужно заполнить оставшиеся индексы таким образом, чтобы ни один из них при переходе в следующий блок не изменился.

Если на втором шаге срабатывает условие а) и мы получаем, что индекс ε_2 равен 0, тогда:

а) если третье уравнение системы (1) содержит убывающую функцию, то для того, чтобы первая пара индексов не менялась, нужно, чтобы $\varepsilon_3 = 1$;

б) если третье уравнение системы (1) содержит возрастающую функцию, то для того, чтобы первая пара индексов не менялась, нужно, чтобы $\varepsilon_3 = 0$.

В случае, если при выполнении второго шага выполнился пункт б) и $\varepsilon_2 = 1$, то:

а) если третье уравнение системы (1) содержит убывающую функцию, то для того, чтобы первая пара индексов не менялась, нужно, чтобы $\varepsilon_3 = 0$;

б) если третье уравнение системы (1) содержит возрастающую функцию, то для того, чтобы первая пара индексов не менялась, нужно, чтобы $\varepsilon_3 = 1$.

Итак, рассмотрим все возможные варианты построения блоков для первых четырех шагов алгоритма.

1 шаг. Задаем $\mathcal{E}_1 = \{0\,\varepsilon_2\ \dots\}$.

2 шаг. $\{0\,\varepsilon_2\ \dots\}$, меняем индекс в первой паре:

$\{f_1\, L_2\ \dots\} \Rightarrow \{0\,0\,\varepsilon_3\ \dots\}$;

$\{f_1\, \Gamma_2\ \dots\} \Rightarrow \{0\,1\,\varepsilon_3\ \dots\}$.

3 шаг. $\{0\,0\,\varepsilon_3\ \dots\}$, не меняем индекс во второй паре:

$\{f_1\, L_2\, L_3\ \dots\} \Rightarrow \{0\,0\,1\,\varepsilon_4\ \dots\}$;

$\{f_1\, L_2\, \Gamma_3\ \dots\} \Rightarrow \{0\,0\,0\,\varepsilon_4\ \dots\}$;

4 шаг. $\{0\,0\,1\,\varepsilon_4\ \dots\}$, не меняем индекс в третьей паре:

$\{f_1\, L_2\, L_3\, L_4\ \dots\} \Rightarrow \{0\,0\,1\,0\,\varepsilon_5\ \dots\}$;

$\{f_1\, L_2\, L_3\, \Gamma_4\ \dots\} \Rightarrow \{0\,0\,1\,1\,\varepsilon_5\ \dots\}$;

$\{0\,1\,0\,\varepsilon_4\ \dots\}$, не меняем индекс в третьей паре:

$\{f_1\, \Gamma_2\, L_3\, L_4\ \dots\} \Rightarrow \{0\,1\,1\,0\,\varepsilon_5\ \dots\}$;

$\{0\,1\,\varepsilon_3\ \dots\}$, не меняем индекс во второй паре:

$\{f_1\, \Gamma_2\, L_3\ \dots\} \Rightarrow \{0\,1\,0\,\varepsilon_4\ \dots\}$;

$\{f_1\, \Gamma_2\, \Gamma_3\ \dots\} \Rightarrow \{0\,1\,1\,\varepsilon_4\ \dots\}$.

$\{0\,0\,0\,\varepsilon_4\ \dots\}$, не меняем индекс в третьей паре:

$\{f_1\, L_2\, \Gamma_3\, L_4\ \dots\} \Rightarrow \{0\,0\,0\,1\,\varepsilon_5\ \dots\}$;

$\{f_1\, L_2\, \Gamma_3\, \Gamma_4\ \dots\} \Rightarrow \{0\,0\,0\,0\,\varepsilon_5\ \dots\}$;

$\{0\,1\,1\,\varepsilon_4\ \dots\}$, не меняем индекс в третьей паре:

$\{f_1\, \Gamma_2\, \Gamma_3\, L_4\ \dots\} \Rightarrow \{0\,1\,0\,0\,\varepsilon_5\ \dots\}$;

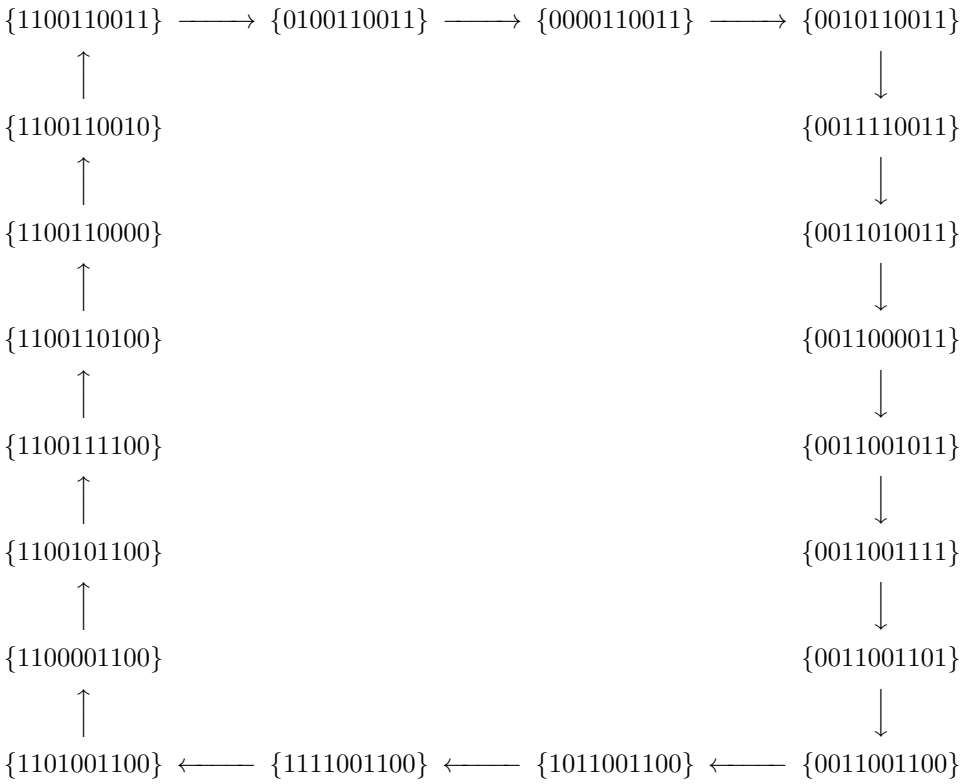
$\{f_1\,\Gamma_2\,L_3\,\Gamma_4\,\ldots\} \Rightarrow \{0\,1\,1\,0\,\varepsilon_5\,\ldots\}$; $\{f_1\,\Gamma_2\,\Gamma_3\,\Gamma_4\,\ldots\} \Rightarrow \{0\,1\,1\,1\,\varepsilon_5\,\ldots\}$.

Дальнейшие шаги выполняем аналогично предыдущим до тех пор, пока не будет найден последний блок, из которого возможен переход только в блок $\mathcal{E}_1$, и система не замкнется.

3 Пример построения диаграмм для систем вида (1)

В работе [6] рассматривалась десятимерная модель кольцевой генной сети, у которой правые части уравнений задаются функциями: $f_{2j-1} = L_{2j-1}, f_{2j} = \Gamma_{2j}, j = \overline{1,10}$.

С помощью описанного выше алгоритма строится диаграмма



переходов траекторий этой системы по блокам валентности 1. Пользуясь леммой 3, заметим, что переход из блока $\mathcal{E}_1 = \{1100110011\}$ можно осуществить только один, по первой паре индексов: пара мультииндексов $\{11\}$ соответствует убывающей функции, а значит, переходит в пару $\{01\}$. Таким образом, делаем вывод, что траектории точек блока $\mathcal{E}_1$ переходит в блок $\mathcal{E}_2 = \{0100110011\}$, из которого, согласно лемме 3, можно совершить только один переход, по второй паре индексов: $\{10\}$ переходит в $\{00\}$, соответственно, третий блок $\mathcal{E}_3$ будет состоять из следующего набора мультииндексов: $\{0000110011\}$. И так далее, продолжая аналогичные рассуждения, мы замыкаем цепочку блоков последним блоком $\{1100110010\}$, откуда траектории переходят в блок $\mathcal{E}_1$.

В работе [6] подобные построения были проведены и для восемнадцатимерной системы вида (1), для которой с помощью дискретизации (2) фазового портрета была установлена неединственность цикла.

Аналогичная диаграмма для несимметричной динамической системы

$$\frac{dx_1}{dt} = -k_1x_1 + f_1(x_9);\ \ \frac{dx_2}{dt} = \mu_2x_1 - k_2x_2;\ \ \frac{dx_3}{dt} = \mu_3x_2 - k_3x_3;$$

$$\frac{dx_4}{dt} = \mu_4x_3 - k_4x_4;\ \ \frac{dx_5}{dt} = -k_5x_5 + f_5(x_4);\ \ \frac{dx_6}{dt} = \mu_6x_5 - k_6x_6;$$

$$\frac{dx_7}{dt} = \mu_7x_6 - k_7x_7;\ \ \frac{dx_8}{dt} = -k_8x_8 + f_8(x_7);\ \ \frac{dx_9}{dt} = \mu_9x_8 - k_9x_9;$$

была построена в [4]. Во всех случаях построение диаграмм из блоков валентности один проводилось с помощью описанного выше алгоритма. Подобные динамические системы рассмотрены и в [2, 3].

Замечание. *При переходе траекторий согласно стрелкам построенных диаграмм валентности блоков не увеличиваются, и поэтому для каждой системы вида* (1) *объединение блоков валентности один является инвариантной областью этой системы.*

Заключение

Разработанный алгоритм построения диаграмм переходов траекторий систем по блокам валентности один применялся и при описании комбинаторных структур фазовых портретов более сложных динамических систем, см. [5].

Список литературы

[1] Акиньшин А.А., Голубятников В.П. Геометрические характеристики циклов в некоторых симметричных динамических системах // Вестник НГУ. Серия "Математика, механика, информатика". 2012. Т. 12, № 12, С. 3–12.

[2] Аюпова Н.Б., Голубятников В.П., Казанцев М.В. О существовании цикла в одной несимметричной модели молекулярного репрессилятора // Сибирский журнал вычислительной математики. 2017. Т. 20, № 2, С. 121–129.

[3] Banks H.T., Mahaffy J.M. Stability of cyclic gene models for systems involving repression // J. theor. Biol. 1978. V 74, P. 323–334.

[4] Bukharina T.A., Golubyatnikov V.P., Kirillova N.E., et al. Mathematical and numerical models of two asymmetric gene networks // SEMR. 2018. V. 15, P. 1271–1283.

[5] Голубятников В.П., Иванов В.В. Циклы в нечетномерных моделях кольцевых генных сетей // Сибирский журнал индустриальной математики. 2018. Т. 21, № 4, С. 28–38.

[6] Голубятников В.П., Кириллова Н.Е. О циклах в моделях функционирования кольцевых генных сетей // Сиб. Жур. Чистой и Прикл. Математики. 2018. Т. 18, № 1, С. 54–63.

[7] Казанцев М.В. О некоторых свойствах графов доменов динамических систем // Сибирский журнал индустриальной математики. 2015. Т. 18, № 4, С. 42–49.

[8] Колесов А.Ю., Розов Н.Х., Садовничий В.А. Периодические решения типа бегущих волн в кольцевых генных сетях // Известия РАН, серия математическая. 2016. Т. 80, № 3. С. 67 – 94.

[9] Murray J.D. Mathematical biology: I. An introduction, third edition. Springer, 2007.

Кириллова Наталья Евгеньевна — аспирантка
Института математики им. С. Л. Соболева СО РАН;
e-mail: kne@math.nsc.ru.
Дата поступления — 30 апреля 2019 г.

ПРИМЕНЕНИ АНСАМБЛЕВОГО ПИ-АЛГОРИТМА В ЗАДАЧЕ ОБРАТНОГО МОДЕЛИРОВАНИЯ

Е. Г. Климова

Институт вычислительных технологий СО РАН, 630090, Новосибирск

УДК 551.509.313
DOI: 10.24411/9999-016A-2019-10038

Решение задачи обратного моделирования для заданной модели процесса и набора данных наблюдений циклически по времени является задачей усвоения данных. В проблему обратного моделирования включена также задача оценки параметров модели. Использование динамико-стохастического подхода к решению задачи усвоения данных привело к развитию ансамблевых методов: ансамблевого фильтра Калмана, ансамблевого сглаживания Калмана. В задаче усвоения данных при моделировании процессов в окружающей среде оцениваемый вектор и вектор данных наблюдений имеют высокую размерность, что требует разработки эффективных алгоритмов. Ансамблевый пи-алгоритм разработан как эффективная реализация стохастического ансамблевого фильтра Калмана. Он может быть использован в задачах оценки состояния окружающей среды, а также параметров модели. Одной из проблем ансамблевых фильтров является возможная расходимость алгоритма со временем. В докладе рассматриваются подходы к практической реализации ансамблевого фильтра Калмана и ансамблевого сглаживания Калмана с использованием пи-алгоритма, связанные с адаптивной коррекцией поведения алгоритма по времени. Приводятся результаты численных экспериментов с 1-мерной тестовой моделью.

Ключевые слова: усвоение данных, ансамблевые алроритмы.

Введение

Решение задачи обратного моделирования для заданной модели процесса и набора данных наблюдений циклически по времени является задачей усвоения данных. В проблему обратного моделирования включена также задача оценки параметров модели. Для ее решения часто применяется байесовский подход, при этом данные наблюдений и оцениваемые параметры считаются случайными величинами. В случае, если рассматриваемые случайные величины являются гауссовскими, а модели прогноза и наблюдений линейными, такая постановка задачи эквивалентна вариационной постановке задачи усвоения данных (4DVAR) [7]. Для аппроксимации рассматриваемых в алгоритме ковариационных матриц может быть использован ансамбль (выборка) прогнозов и наблюдений [4]. В случае, если оптимальная оценка ищется в конце заданного временного интервала, то задача сводится к ансамблевому фильтру Калмана. Если рассматривается оптимальная оценка на заданном временном интервале, такая задача называется задачей ансамблевого сглаживания [4]. В задаче усвоения данных при моделировании процессов в окружающей среде оцениваемый вектор и вектор данных наблюдений имеют высокую размерность, что требует разработки эффективных алгоритмов.

Одной из проблем ансамблевых фильтров является возможная расходимость алгоритма со временем. В обзоре [8] перечисляются основные подходы к улучшению сходимости ансамблевого фильтра Калмана. Одним из популярных способов регулирования сходимости ансамблевого фильтра Калмана является использование так называемого “inflation factor” (увеличивающий множитель). В работе [8] отмечаются такие способы борьбы с расходимостью ансамблевого фильтра Калмана, как увеличивающий множитель (multiplicative inflation) и увеличивающее дополнительное возмущение (additive inflation). В нелинейном случае в настоящее время применяются итерационные алгоритмы сглаживания, представляющие собой вариационную задачу с применением ансамблей для вычисления ковариаций и линеаризации операторов прогноза и наблюдений [3,5]. В строго нелинейном негауссовском случае используется метод частиц, который основан на байесовском подходе [3]. В данных подходах также рассматриваются способы управления сходимостью алгоритмов. Ансамблевый пи-алгоритм эффективен при использовании при реализации стохастического ансамблевого

ISBN 978-5-901548-42-4

фильтра Калмана и стохастического ансамблевого сглаживания Калмана. Он может быть использован в задачах оценки состояния окружающей среды, а также параметров модели. В настоящее время ансамблевый фильтр Калмана является базой для развития метода частиц, применяемого в случае негауссовских случайных величин. Для реализации этого метода также требуется развитие эффективных алгоритмов.

В докладе рассматриваются подходы к практической реализации ансамблевых алгоритмов, связанные с адаптивной коррекцией поведения алгоритма по времени. Приводятся результаты численных экспериментов с 1-мерной тестовой моделью.

1 Задача обратного моделирования

Задачей обратного моделирования называют оценку неизвестного вектора параметров большой размерности $x \in R^K$ по вектору данных наблюдений $y \in R^L$ [13]. Если предположить, что

$$y = M(x) + \delta,$$

где M — известный оператор прогноза–наблюдений, δ — случайный шум с заданной функцией распределения, то оценка вектора состояния x по данным наблюдений y производится на основе байесовского подхода. Задачу последовательного оценивания по времени неизвестного величины по временному ряду данных наблюдений называют задачей усвоения данных [13].

Предположим, что изменение по времени оцениваемой величины x_k описывается моделью

$$x_{k+1} = f_{k+1,k}(x_k) + \eta_k,$$

где k — номер шага по времени. Кроме того, известны данные наблюдений y_k:

$$y_k = h_k(x_k) + \varepsilon_k.$$

В этих формулах η_k, ε_k — случайные ошибки прогноза и наблюдений, соответственно, $f_{k+1,k}$ — оператор модели, h_k — оператор трансформации прогнозируемой переменной в наблюдаемую.

Байесовский подход состоит в применении теоремы Байеса для получения оптимальной оценки по данным наблюдений и прогнозу:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}.$$

Известны различные варианты оценки состояния по данным и прогнозу:
$p(x_l|y_{k,1}),\ k > l$ — прогноз,
$p(x_k|y_{k,1})$ — фильтрация,
$p(x_{k,0}|y_{k,1})$ — сглаживание,
где $x_{k,0} = \{x_k, x_{k-1}, \cdots, x_0\}$, $y_{k,1} = \{y_k, \cdots, y_1\}$. Обозначения и формулировки взяты из обзора [3].

В линейном гауссовском случае решением задачи фильтрации является фильтр Калмана, решением задачи сглаживания — сглаживание Калмана. В работе [4] предложено использовать метод Монте-Карло для решения задач фильтрации и сглаживания, так называемые ансамблевые алгоритмы фильтрации и сглаживания. В ансамблевом фильтре Калмана в нелинейном случае прогноз ковариационной матрицы осуществляется с помощью нелинейной модели, при этом нарушается условие гауссовости. Также в этом случае оценка на шаге анализа является приближением оценки минимальной дисперсии (linear variance minimizing). Для решения задачи сглаживания в нелинейном случае в настоящее время рассматриваются итерационные алгоритмы сглаживания с применением ансамблей. Как отмечается в [5], итерационные методы дают хорошее приближение для слабо нелинейных моделей. В нелинейном негауссовском случае используется метод частиц, который также основан на байесовсом подходе.

2 Применение ансамблевого подхода к задаче оптимальной фильтрации

Стохастический ансамблевый фильтр Калмана [4] состоит из ансамбля прогнозов $\{x_k^{f,n}, n = 1, \cdots, N\}$

$$x_k^{f,n} = f(x_{k-1}^{a,n}) + \eta_{k-1}^n \tag{1}$$

и ансамбля анализов $\{x_k^{a,n}, n = 1, \cdots, N\}$

$$x_k^{a,n} = x_k^{f,n} + K_k[y_k^n + \varepsilon_k^n - h(x_k^{f,n})]. \tag{2}$$

В этих формулах N — количество элементов ансамбля. Ансамбли (1) и (2) задают выборку значений «истины», при этом среднее по выборке значение будет являться оптимальной оценкой, а отклонение от среднего – ансамблем ошибок анализа и прогноза, соответственно. Для осуществления ансамблевого варианта алгоритма фильтра Калмана требуется задание ансамбля ошибок наблюдений $\{\varepsilon_k^n, n = 1, \cdots, N\}$ а также ансамбля ошибок прогноза $\{dx_k^{f,n} = x_k^{f,n} - \overline{x_k^{f,n}},\ n = 1, \cdots, N\}$, где $\overline{x_k^{f,n}} \cong \frac{1}{N}\sum_{n=1}^N x_k^{f,n}$ и ансамбля шумов модели $\{\eta_{k-1}^n, n = 1, \cdots, N\}$: $E[\eta_{k-1}^n \left(\eta_{k-1}^n\right)^T] = Q_{k-1}$. Матрица K_k имеет вид

$$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k)^{-1},$$

где P_k^f и R_k — матрицы, оцениваемые по ансамблю

$$P_k^f \triangleq \frac{1}{N-1}\sum_{n=1}^N dx_k^{f,n}\left(dx_k^{f,n}\right)^T,\ R_k \triangleq \frac{1}{N-1}\sum_{n=1}^N \varepsilon_k^n \left(\varepsilon_k^n\right)^T,$$

H_k линеаризованный оператор $h(x_k^{f,n})$ относительно $\overline{x_k^{f,n}}$:

$$h(x_k^{f,n}) \cong h(\overline{x_k^{f,n}}) + H_k dx_k^{f,n}.$$

Детерминированный ансамблевый фильтр Калмана (шаг анализа) состоит из уравнения для среднего значения

$$\overline{x_k^{a,n}} = \overline{x_k^{f,n}} + K_k[y_k^n - \overline{h(x_k^{f,n})}]$$

и оценки ансамбля ошибок анализа такой, чтобы соответствующая ковариационная матрица удовлетворяла уравнению фильтра Калмана $P_k^a = (I - K_k H_k)P_k^f$ [8].

Одним из преимуществ стохастического фильтра Калмана является то, что отклонение элемента ансамбля от среднего (ensemble spread) имеет вид, аналогичный ошибке оценки — отклонению среднего значения от “истины” (ensemble skill) [8, 12].

3 Проблемы реализации ансамблевых алгоритмов

Усвоение данных для современных задач моделирования процессов в окружающей среде имеет большую размерность как прогнозируемого вектора. так и вектора наблюдений. Всвязи с этим актуальной является разработка эффективных алгоритмов реализации процедуры усвоения данных. Одним из способов создания экономичного алгоритма является использование методов трансформации ансамбля возмущений. Ансамблевый пи-алгоритм является эффективной реализацией стохастического ансамблевого фильтра Калмана [1,10]. Ансамблевый пи-алгоритм является стохастическим фильтром, в котором ансамбль ошибок анализа получается с помощью матрицы трансформаций, не зависящей от узла сетки. Это делает возможным реализацию алгоритма локально, при этом для реализации ансамблевого пи-алгоритма требуются операции с матрицами порядка размерности ансамбля. Как видно из формул для оценки параметра в ансамблевом алгоритме, матрица трансформаций одна и та же для оценки моделируемой переменной и параметра модели. Обоснование метода и детали его реализации содержатся в [1, 10].

При использовании ансамблей в задачах усвоения данных возникает ряд проблем, связанных с небольшим размером выборки (ансамбля). В частности, это проблема высоких значений корреляций ошибок прогноза на больших расстояниях. В этом случае принять применять так называемую локализацию, а именно, поэлементное умножение ковариационных матриц на функцию, убывающую с расстоянием [8].

Малый размер выборки также является одной из причин расходимости алгоритма. Расодимость бывает двух типов. В первом случае происходит рост среднеквадратиеской ошибки. Во втором случае быстро убывает след матрицы ковариаций ошибок оценивания и в результате весовая матрица при данных наблюдений

становится близкой к нулевой. Для предотвращения расходимости второго типа в алгоритмах ансамблевого фильтра Калмана (EnKF), ансамблевого сглаживания Калмана (EnKS) используется «увеличивающий множитель» (inflation factor), при этом либо элементы ковариационной матрицы умножаются на некоторый множитель (multiplicative inflation), либо к ансамблю возмущений добавляется дополнительное возмущение (additive inflation). Следует отметить, что эти приемы используются и в ряде итерационных алгоритмов сглаживания с применением ансамблей, а также в методе частиц с применением ансамблей [8].

Следует отметить, что применение мультипликативного множителя в линейном случае эквивалентно увеличению весового множителя при более близко расположенных к заданному моменту времени данных [9].

В целях улучшения сходимости алгоритмов по времени популярным является адаптивное уточнение следа матрицы ковариаций ошибок оценивания. В основе этого подхода лежит свойство вектора невязок r [9]:

$$r = y_0 - H(x_b) \approx \varepsilon_0 - H(\varepsilon_b),$$

где y_0 — вектор данных наблюдений, x_b — вектор прогностических переменных (первого приближения), H — оператор наблюдений, ε_0 , ε_b — векторы случайных ошибок прогноза и наблюдений, соответственно. Вектор невязок обладает свойством

$$\overline{rr^T} = \overline{\varepsilon_0\varepsilon_0^T} + H\overline{\varepsilon_b\varepsilon_b^T}H^T = R + HP_bH^T,$$

где R, P_b — матрицы ковариаций ошибок наблюдений и первого приближения, соответственно.

Для оценки увеличивающего множителя используется формула

$$\rho = \frac{\overline{r^Tr} - tr(R)}{tr(HP_bH^T)},$$

при этом принято предполагать, что $\overline{r^Tr} \approx r^Tr$, tr означает след матрицы.

В локальных алгоритмах возможно проводить адаптивную оценку увеличивающего множителя в заданной подобласти. В частности, рассматриваемый в данной работе ансамблевый пи-алгоритм является локальным, т.е. оценка по данным производится в отдельном узле сетки независимо.

4 Пример управления сходимостью ансамблевого алгоритма с помощью уеличивающего множителя

Рассмотрим подход к управлению сходимостью ансамблевого фильтра Калмана на примере ансамблевого пи-алгоритма для 1-мерной модели переноса и диффузии пассивной примеси с неизвестным параметром (эмиссией примеси). Для решения уравнения переноса и диффузии применялся полу-лагранжев метод, при этом бралась неявная схема по времени и схема центральных разностей по пространству. Для решения конечно-разностного аналога уравнения диффузии использовался метод циклической прогонки. Уравнение решалось на отрезке $(0, 1)$, при этом рассматривались периодические граничные условия. Задавалось 240 узлов сетки, $u = 1$, $k^2 = 0,6 \times 10^{-3}$.

Рассмотрим конечно-разностный аналог уравнения переноса и диффузии в виде

$$\phi_{k+1} = A_k\phi_k + g_k,$$

где A_k — линейный оператор, k — номер шага по времени. Заданные начальные значения ϕ_0^t, g_0^t считались «истинными». Для получения начальных данных для прогноза по модели ϕ_0^d к «истинным» начальным данным добавлялось возмущение $\phi_0^d = \phi_0^t + \delta$, $\delta = \mathrm{N}(0, \mathrm{s}_0)$. Через $N(a,b)$ обозначена случайная величина, распределенная по нормальному закону с математическим ожиданием, равным a и дисперсией, равной b.

Для организации численных экспериментов задавались: ансамбль начальных полей $\phi_0^n = \phi_0^d + \delta^{\mathrm{n}}$, $\delta^{\mathrm{n}} \sim \mathrm{N}(0, \mathrm{s}_0)$, $\mathrm{n} = 1, \cdots, \mathrm{N_{ens}}$; $g_0^n = g_0^d + \delta_g^n$, $\delta_g^n \sim \mathrm{N}(0, \mathrm{dg}_0)$, $\mathrm{n} = 1, \cdots, \mathrm{N_{ens}}$, наблюдения $y_0 = \phi_t(0) + \delta_0; \delta_0 \sim \mathrm{N}(0, \varepsilon_0)$; ансамбль наблюдений с возмущениями $y_0^{\mathrm{n}} = y_0 + \delta_0^{\mathrm{n}}$, $\delta_0^{\mathrm{n}} \sim \mathrm{N}(0, \varepsilon_0)$, $\mathrm{n} = 1, \cdots, \mathrm{N_{ens}}$. Через $\mathrm{N_{ens}}$ обозначено число элементов ансамбля. Данные наблюдений считались известными во всей области интегрирования. Прогноз осуществлялся в течение N_t=240 шагов по времени, усвоение данных проводилось на каждом шаге по времени, при этом производилась оценка значений ϕ и g. Численные эксперименты проводились для значений $\mathrm{s}_0 = \varepsilon_0 = 0.01$, $\mathrm{N_{ens}} = 20$ и 40.

Во всех численных экспериментах рассматривался вариант $R = \varepsilon_0^2 I$. При анализе в узле сетки l брались данные наблюдений из интервала $(l - id, l + id)$. При этом при анализе в узле сетки l вместо матрицы R

бралась матрица $\tilde{R} = R \circ e^{-0.5(\rho_{il}/bc)^2}$, где ρ_{il} — расстояние между узлом сетки и наблюдением, " $\circ$ " — знак поэлементного умножения. В проводимых экспериментах брались значения $id = 5$, $bc = 5\Delta x$ (Δx — шаг сетки). Такой алгоритм носит название локализация. Его принято использовать в ансамблевых методах для подавления ложных ковариаций на больших расстояниях, возникающих из-за малого размера выборки (ансамбля).

В численных экспериментах "истинное" значение параметра задавалось кусочно-постоянным в виде дискретного аналога функции $\hat{g}(x,t)$, где:

$$\hat{g}(x,t) = \begin{cases} g_0(x)\ ,\ 0 \le t < 120 \\ 0.8 * g_0(x),\ t \ge 120 \end{cases},$$

где

$$g_0(x) = \begin{cases} 0.1\ ,\ 0.375 \le x \le 0.625 \\ 0,\ 0.375 > x\ \wedge\ x > 0.625 \end{cases}.$$

Для предотвращения расходимости алгоритма со временем использовался так называемый «увеличивающий множитель» $infl$ ($inflationfactor$) [8]. То есть, на шаге прогноза ансамбль возмущений домножался на константу: вместо P_f рассматривалась матрица, все элементы которой умножены на $infl$.

В численных экспериментах использовалась ансамблевый пи-алгоритм — стохастический ансамблевый фильтр Калмана. Были проведены численные эксперименты с различным заданием параметра $infl$. В 1-м эксперименте $infl = 1$. Во 2-м эксперименте $infl = 1.3$. В 3-м эксперименте $infl$ рассчитывался в каждом узле сетки отдельно, что возможно при реализации пи-алгоритма, поскольку он является локальным и коррекция производится в каждом узле сетки (или группе узлов) независимо. В эксперименте 4 $infl$ рассчитывается как среднее значение по подобласти узла сетки, для которого производится усвоение. В эксперименте 5 вычисляется среднее по времени значение $infl$, то есть на каждом шаге по времени берется вместо $infl$ значение

$$infl1 = \beta * infl + (1 - \beta) * infl_0,$$

где $infl_0$ — значение с предыдущего шага по времени, $\beta = 0.8$. Результаты экспериментов приведены на рис. 1 и рис. 2. На этих рисунках представлено поведение среднеквадратической ошибки переменных ϕ и g, соответственно при различных способах задания $infl$.

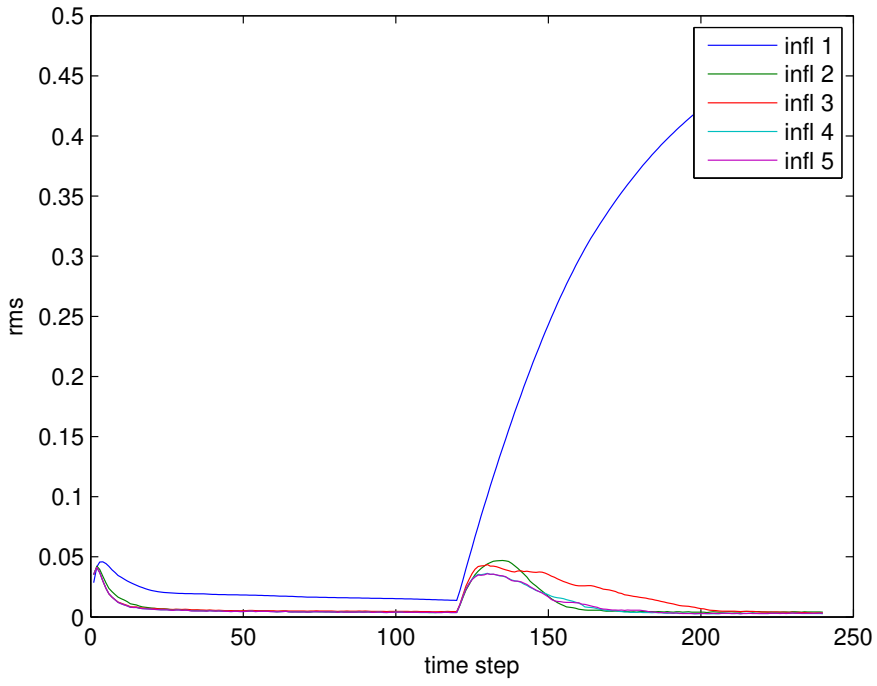


Рис. 1: Среднеквадратическая ошибка оценки прогнозируемой переменной

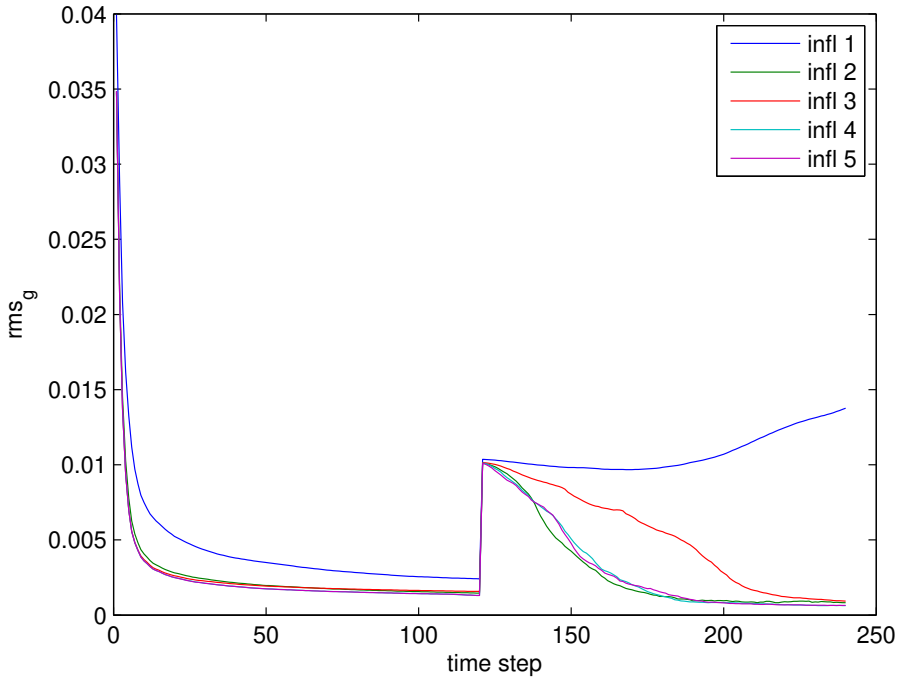


Рис. 2: Среднеквадратическая ошибка оценки параметра

Как видно из рисунков, расходимость алгоритма усвоения для заданной модели происходит после изменения "истинного" значения неизвестного параметра. Использование параметра $infl$ позволяет предотвратить развитие расходимости процесса усвоения, при этом адаптивное задание $infl$ позволяет получить более точную оценку как прогнозируемой переменной так и неизвестного параметра.

Заключение

В докладе рассмотрены современные постановки задачи усвоения данных, основанные на ансамблевом подходе. Для повышения эффективности этих алгоритмов предлагается использовать алгоритм с трансформацией ковариационных матриц, в частности, предложенный в [1, 10] ансамблевый пи-алгоритм. Рассмотрены способы управления сходимостью ансамблевых алгоритмов.

Список литературы

[1] Климова Е.Г. Стохастический ансамблевый фильтр Калмана с трансформацией ансамбля возмущений // Сибирский журнал вычислительной математики /РАН. Сибирское отделение — Новосибирск. 2019. — Т. 22, №1. С. 27–40.

[2] Справочник по теории автоматического управления под ред. А.А. Красовского. Москва: Наука, 1987. 711 с.

[3] Carrassi A., Bocquet M., Bertino L., Evensen G. Data assimilation in the geosciences: An overview of methods, issuers and perspectives // Wiley interdisciplinary reviews: Climate Change. 2018. V. 131, Issue 5, e535, doi: 10.1002/wcc535.

[4] Evensen, G. Data assimilation. The ensemble Kalman filter, Spriger-Verlag, Berlin Heideberg, 2009.

[5] Evensen G. Analysis of iterative ensemble smoother for solving inverse problems. Computational Geosciences. 2018. 22. P.885-908. doi.org/10.1007/s10596-018-9731-y.

[6] Evensen, G., P.J. van Leeuwen An ensemble Kalman smoother for nonlinear dynamics // Monthly Weather Review. 2000. V. 128. P. 1852-1867.

[7] Fisher M., Leutbecher M., Kelly G.A. On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation// Quarterly Journal of the Royal Meteorological Society. 2005. V. 131, P. 3235–3246.

[8] Houtekamer, H.L. Zhang, F. Review of the ensemble Kalman filter for atmospheric data assimilation // Monthly Weather Review. 2016. V. 144, P. 4489–4532.

[9] Jazwinski, A.H. Stochastic processes and filtering theory. Academic Press, New York, 1970.

[10] Klimova E. A suboptimal data assimilation algorithm based on the ensemble Kalman filter // Quarterly Journal of the Royal Meteorological Society. 2012. V. 138, P. 2079–2085.

[11] Klimova E.G. Application of ensemble Kalman filter in environment data assimilation // IOP Conference Series: Earth and Environmental Science. – Vol. 211 (??) 012049 . – doi:10.1088/1755-1315/211/1/012049.

[12] Lei, J., Bickel, P., Shyder, C. Comparison of ensemble Kalman filters under non-gaussianity //Monthly Weather Review. 2010. V. 138. P. 1293–1306.

[13] Nakamura G., Potthast R. Inverse Modeling. IOP Publishing. 2015. 509 pp. doi.org/10.1088/978-0-7503-1218-9.

Климова Екатерина Георгиевна — д.ф.-м.н., ст. науч.сотр. Института вычислительных технологий СО РАН;
e-mail: klimova@ict.nsc.ru
Дата поступления — 30 апреля 2019 г.

ИСПОЛЬЗОВАНИЕ ЧЕТЫРЕХМЕРНЫХ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ АВТОМАТИЗАЦИИ ПОСТРОЕНИЯ МОДЕЛЕЙ МЕСТНОСТИ

А. А. Колесников[1], П. М. Кикин[2]

[1] *Сибирский государственный университет геосистем и технологий, 630108, Новосибирск*
[2] *Санкт-Петербургский Политехнический Университет Петра Великого, 195251, Санкт-Петербург*

УДК 528.854
DOI: 10.24411/9999-016A-2019-10039

Описаны способы, особенности и проблемы использования многомерных сверточных нейронных сетей для семантической сегментации облаков точек, полученных как по результатам съемки в помещениях, так и на открытой местности. Приводятся предлагаемые классификации типов нейронных сетей для трехмерной семантической сегментации и моделей, распределенных по принципу ориентации на обработку определенного типа пространственных данных. Сформирован сводный список эталонных тестов для оценки результатов семантической сегментации облаков точек съемки помещений, городской застройки, отдельных строений, частей механизмов и бытовых инструментов.

Ключевые слова: облако точек, семантическая сегментация, сверточная нейронная сеть.

Введение

На сегодняшний день появляется все больше детальной трехмерной информации, поставщиками которой являются беспилотные летательные аппараты и лазерные сканеры. Но эти данные представляют собой неструктурированные статические картины, даже если речь идет о видеопотоке, который разделяется на отдельные изображения. Тем не менее сочетание трехмерной и семантической информации являются фундаментальными для обработки пространственных данных всех типов, использования робототехнических комплексов и автономных транспортных средств. Сегодня, после обработки и загрузки облака точек в какую либо ГИС и CAD программу, разделение массива точек на отдельные группы по их отнесению к реальному физическому объекту (далее семантическая сегментация) как правило выполняется вручную, либо автоматизации подлежит выделение только строго определенных классов объектов (как правило объектов растительности). Чтобы назначить правильную метку для каждой точки в трехмерной сцене, алгоритмы семантической классификации должны понимать геометрию сцены.

Поскольку существующий уровень развития глубоких нейронных сетей позволяет с высокой точностью (зачастую превышающей аналогичные показатели среднестатистического человека) производить семантическую сегментацию фотографий и спутниковых снимков, то логичным было использовать данную технологию и для обработки облаков точек. Поэтому в данной работе приведены существующие способы семантической сегментации облаков точек преимущественно средствами нейронных сетей, рассмотрены перспективные подходы и архитектуры сетей, а также тестовые наборы данных, представляющие собой облака точек ориентированные на различные задачи. Размещение лазерных сканеров на беспилотных летательных аппаратах и автомобилях позволяет получать динамические облака точек, что еще более усложняет задачу сегментации и приводит к разработке методики использования четырехмерных нейронных сетей. Далее под четырехмерными (4D) нейронными сетями подразумеваются сети ориентированные на обработку данных, где каждый объект описывается четырьмя или более параметрами, например, облака точек с указанием цвета, видео с данными камеры глубины, множественные съемки лазерным сканером одного и того же объекта с разных точек и т.д.

ISBN 978-5-901548-42-4

1 Многомерные нейронные сети для семантической сегментации облаков точек

Существует много технических проблем при использовании пространственно-временных данных для задач сегментации высокого уровня. Во-первых, трехмерные данные требуют гетерогенного представления и обработки или затрудняют интеграцию в более крупные системы (как правило геоинформационные). Во-вторых, производительность трехмерных сверточных нейронных сетей хуже или на одном уровне с двумерными сверточными нейронными сетями. В-третьих, существует ограниченное количество библиотек с открытым исходным кодом для быстрой обработки крупномасштабных 3D-данных. Пространственно-временное восприятие (4D) принципиально требует обработки в качестве основы трехмерных сцен, поскольку срез 4D информации в единый момент времени представляет собой трехмерное сканирование.

Поэтому в первую очередь рассмотрим трехмерную семантическую сегментацию облаков точек, преимущественно с использованием нейронных сетей. В настоящее время существуют различные работы описывающие принципы трехмерной сегментации: плотная трехмерная свертка [1], варианты pointnet [2, 3], непрерывные свертки [4, 5], свертки поверхностей [6, 7], и свертка octree [8]. По результатам анализа авторами работы все существующие решения были разделены на три класса: трехмерные сверточные нейронные сети, нейронные сети не использующие свертки и традиционные алгоритмические методы машинного обучения.

Трехмерные сверточные нейронные сети в свою очередь разделяются на три категории, зависящие от способа представления исходных данных. Первая категория трехмерных сверточных нейронных сетей использует прямоугольную сетку и плотное представление [1, 9], где пустое пространство представляется либо как 0, либо как функция расстояния со знаком. Это простое представление интуитивно понятно и поддерживается всеми основными свободно распространяемыми библиотеками работы с нейронными сетями. Однако, поскольку большая часть пространства в 3D-сканировании пуста, что приводит к высокому потреблению памяти и медленным вычислениям. Чтобы решить эту проблему OctNet [8] предложили использовать структуру Octree для представления трехмерного пространства и свертки на нем. Вторая категория - это разреженные трехмерные сверточные нейронные сети [10, 11]. Для больших размеров используются два метода квантования: прямоугольная сетка и пермутоэдральная решетка. Разработчики Splat-Net использовали пермутоэдральную решетку и прямоугольную сетку для построения трехмерной сверточной нейронной сети для 3D-классификации и семантической сегментации [10, 12]. Третью категорию представляют 3D-псевдонепрерывные сверточно-нейронные сети [4, 5]. В отличие от предыдущих вариантов, они определяют свертки, используя непрерывные ядра в непрерывном пространстве. Однако поиск соседей в непрерывном пространстве вычислительно затратен, поскольку требует поиска по дереву KD, а не хеш-таблицы, и подвержен неравномерному распределению облаков точек.

Нейронные сети не использующие свертки. В последнее время наблюдается огромный рост нейронных сетей без 3D сверток для 3D сегментации. Поскольку 3D-сканы состоят из тонких наблюдаемых поверхностей, было предложено использовать 2D-свертки на поверхности для семантической сегментации [6, 7]. Другое направление - это методы на основе PointNet [2, 3]. Варианты PointNet используют набор входных координат в качестве объектов для многослойного персептрона. Однако этот подход обрабатывает ограниченное количество точек, и, таким образом, скользящее окно для обрезания секции из входных данных использовалось для больших пространств, делая размер рецептивного поля довольно ограниченным. Разработчики PointCNN пытались разрешить такие недостатки с помощью повторяющейся сети поверх множества точечных сетей, а также предложили вариант непрерывной трехмерной свертки для нижних уровней PointNet и получили значительное повышение производительности [5].

Традиционные алгоритмические методы машинного обучения представлены методиками использования метаточечных графов и деревьев решений. Разработчики Superpoint Graphs [13] предлагают представлять большие облака точек в виде набора взаимосвязанных простых форм, объединенных метаточками, которые аналогичны суперпиксельным методам при сегментации фотографий. Эта структура представлена в виде аннотированного ориентированного графа, названного метаточечным графом. Его узлы представляют собой примитивы, а дуги описывают их семантические связи со смежными дугами. Представление в виде метаточечного графа имеет несколько преимуществ: вместо классификации отдельных точек или вокселей рассматриваются целые части объекта как единое целое, которые легче идентифицировать, становится возможно подробно описать отношения между соседними объектами, что имеет решающее значение для контекстной классификация (автомобили, как правило, находятся над дорогами, потолки окружены стенами и т. д.), размер графа определяется числом простых структур в облаке, а не общим количеством точек, которое обычно на несколько порядков больше. Другим методом, показывающим хорошие результаты яв-

ляется ансамблирование с использованием вероятностей класса алгоритмов деревьев решений и марковских случайных полей. В данном решении деревья решений являются классификатором, а для для обеспечения пространственной когерентности в классификации между точками используются марковские случайные поля [14].

Четырехмерная семантическая сегментация облаков точек во многом опирается на описанные методы и алгоритмы трехмерной сегментации. В одном из первых алгоритмов предложена модель динамического деформируемого "воздушного шара" для анализа 4D кардиограмм [15]. Позже также были использованы многомерные марковские случайные поля для решения задачи сегментации кардиограмм [16]. В качестве основы многомерных нейронных сетей часто используется разреженный тензор. В традиционной речи, тексте или данных изображения элементы извлекаются плотно. Таким образом, наиболее распространенными представлениями этих данных являются векторы, матрицы и тензоры. Однако для трехмерного сканирования или даже для многомерных пространств такие плотные представления неэффективны из-за разреженности. Вместо этого можно сохранить только заполненную часть пространства как его координаты и связанные с ними особенности. Это представление является N-мерным расширением разреженной матрицы, поэтому оно известно как разреженный тензор. Данное представление данных имеет ряд преимуществ: во-первых, он допускает однородное представление данных в традиционных библиотеках нейросетей, поскольку большинство из них поддерживают разреженные тензоры, во-вторых, разреженная свертка очень похожа на стандартную свертку, которая доказала свою эффективность как в 2D-сегментации, так и в 3D-реконструкции и семантической сегментации , в-третьих, разреженная свертка эффективна и обладает высокой скоростью [9,17,18]. Она вычисляет выходные данные только для предварительно определенных координат и сохраняет их в компактном разреженном тензоре. Это экономит как память, так и вычисления, особенно для трехмерного сканирования или многомерных данных, где большая часть пространства пуста. Как раз разряженный тензор используется в одной из последних архитектур, названной Spatio-Temporal CNN авторы которой объединили 3D-UNet с 1D-AutoEncoder для временных данных и применили модель для автоматического кодирования изображений МРТ головного мозга [19]. Один из самых комплексных вариантов предлагается авторами Minkowski Engine, где рассматриваются многомерные сверточные нейронные сети для четырехмерных пространственно-временных данных, трехмерного видео (видеоряд, дополненный данными с камеры глубины) и семимерного пространства дополненного данными о времени и цветности. По сравнению с другими подходами, которые комбинируют временные данные с рекуррентной нейронной сетью или поверхностной моделью (CRF), предложенная архитектура используют однородное представление и свертки последовательно во всех сетях. Вместо использования RNN используется свертка для временной оси, поскольку доказано, что она более эффективна при моделировании последовательности. Для этого используется технология разреженных тензоров и предложенная обобщенная (генерализованная) разреженная свертка, которая охватывает все дискретные свертки [11]. Для реализации обобщенной (генерализованной) разреженной свертки используется Open Source библиотеку автоматического дифференцирования для разреженных тензоров, которая реализует ряд функций для многомерных сверточных нейронных сетей. Чтобы преодолеть трудности в многомерном четырехмерном пространстве, предлагается использовать гибридное ядро, частный случай обобщенной разреженной свертки и трехсторонне-стационарное условное случайное поле, которое обеспечивает пространственно-временную согласованность в семимерном пространственно-временно-цветностном поле. Экспериментально обосновано, что сверточные нейронные сети с только обобщенными разреженными свертками могут превзойти 2D или 2D-3D-гибридные методы с большим отрывом. Также на 3D-видео 4D пространственно-временные сверточные нейронные сети устойчивы к шуму, превосходят трехмерные сверточные нейронные сети и работают быстрее, чем трехмерные аналоги в некоторых случаях. Кроме того, обобщенные разреженные сверточные сети 4D более устойчивы к шуму, а в некоторых случаях более эффективны, чем трехмерные аналоги [20]. Однако даже при эффективном представлении простое масштабирование трехмерной свертки в многомерные пространства приводит к значительным затратам вычислительных ресурсов и потреблению памяти из-за "проклятия размерности". Для 2D-преобразования с размером ядра 5 требуется $5^2 = 25$ весов, который экспоненциально увеличивается до $5^3 = 125$ в трехмерном кубе и 625 в 4D тессеракте . Это экспоненциальное увеличение, однако, не обязательно приводит к повышению производительности и значительно замедляет работу сети.

Наконец, предсказания многомерных пространственно-временных обобщенных разреженных сверточных сетей не обязательно являются последовательными во всем пространстве и времени. Для обеспечения согласованности предлагаются условные случайные поля большой размерности, определенные в 7-мерном трехстороннем пространстве (пространство-время-цвет) со стационарной функцией согласованности по парам. Для решения этой проблемы используются преобразования условного случайного поля в дифференцируе-

мые рекуррентные слои, которые могут быть реализованы в виде обобщенного разреженного сверточные сети 7D, и непрерывно обучаемых 4D и 7D сетей [20].

Кроме описанной классификации по используемым методам авторами данной работы также предлагается классификация конкретных программных решений, распределенных по ориентированию на обработку определенного типа данных:

- облака точек внутренних помещений (indoor): PointNet, SnapNet, Multiscale SphericalNeighborhoods, PointWeb, SEGCloud;
- облака точек открытой местности (outdoor): SqueezeSeg, LidarPerception, Multiscale SphericalNeighborhoods, Superpoint Graphs, 3D-FCNN-TI, SPLATNet;
- данных полученные с МРТ и электронных микроскопов: 3DUnetCNN, DenseVoxNet, LEKTRONN2;
- определение составных частей механизмов, транспортных средств, промышленных объектов и т.п.: PointNet, PointNet++, OctNet, Graph-CNN, SGPN, KPConv, PointSIFT, SPLATNet.

2 Эталонные тесты семантической сегментации

Для оценки новых методов и алгоритмов сегментации существует ряд наборов тестовых данных (бенчмарков) для трехмерной и четырехмерных сегментации и классификации. При этом это могут быть как наборы специализированных данных, например, данных содержащих облака точек внутренних помещений, открытых пространств, результатов МРТ [1, 21–24].

Далее приводится сводный список наиболее популярных оценочных наборов данных, сгруппированных по тематике облаков точек:

- помещения: ScanNet, 2D-3D-S, S3DIS;
- открытые пространства: Paris-Lille-3D, Semantic3D, Synthia, Varcity;
- данные для беспилотных транспортных средств: KITTY Vision;
- составные части механизмов, транспортных средств и бытовых предметов: ModelNet10, ModelNet40, ShapeNet;
- данные MRT: BraTS.

Заключение

Авторами работы в качестве апробации была выполена семантическая сегментация нескольких небольших участков (от 2-х до 5-и тысяч квадартных метров) малоэтажной застройки. В результате были сформлированы следующие выводы:

- модели нейронных сетей, размещенные в свободном доступе хорошо работают на определенных тестах, но на реальных облаках точек показывают точность ниже на 15-35 процентов (модели и архитектуры, ориентированные на сегментацию определенного типа, например, KITTY Vision показали точность ниже на 5-10 процентов);
- формат входных данных различается в зависимости от архитектуры сети и большое количество времени уходит на обработку исходных данных;
- в текущем варианте интеграция рассмотренных моделей в геоинформационные системы позволит значительно увеличить скорость обработки трехмерных данных, но сильно затруднена сложностями при сборке определенного набора программных библиотек.

Список литературы

[1] Dai A., Chang A.X., Savva M., Halber M.,Funkhouser T., and Niener M.. Scannet: Richly-annotated 3d reconstructions of indoor scenes. InProc. Computer Vision and Pattern Recognition (CVPR), IEEE, 2017

[2] Qi C.R., Su H., Mo K., and Guibas L.J.. Pointnet: Deep learning on point sets for 3d classifi-cation and segmentation.arXiv preprint arXiv:1612.00593,2016

[3] Qi C.R., Yi L., Su H., and Guibas L.J.. Pointnet++: Deep hierarchical feature learning onpoint sets in a metric space. In Advances in Neural Information Processing Systems, 2017

[4] Hermosilla P., Ritschel T., Vazquez P.P., Vinacua A., and Ropinski T. Monte carlo convolution for learning on non-uniformly sampled point clouds.ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2018), 2018

[5] Li Y., Bu R., Sun M., Chen B. Pointcnn. arXiv preprint arXiv:1801.07791, 2018

[6] Pan H., Liu S., Liu Y., Tong X.. Convolutional neural networks on 3d surfaces using parallel frames.arXiv preprint arXiv:1808.04952, 2018

[7] Tatarchenko M., Park J., Koltun V., Zhou Q.. Tangent convolutions for dense predictionin 3D.CVPR, 2018

[8] Riegler G., Ulusoy A.O., Geiger A.. Oct-net: Learning deep 3d representations at high resolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017

[9] Tchapmi L.P., Choy C.B., Armeni I., Gwak J.Y., Savarese S. Segcloud: Semantic segmentationof 3d point clouds.International Conference on 3D Vision(3DV), 2017

[10] Su H., Jampani V., Sun D., Maji S., Kalogerakis V., Yang M., Kautz J. Splat-net: Sparse lattice networks for point cloud processing.arXiv preprint arXiv:1802.08275, 2018

[11] Graham B. Sparse 3d convolutional neural networks.British Machine Vision Conference, 2015

[12] Adams A., Baek J., Davis M.A.Fast high-dimensional filtering using the permutohedral lattice. InComputer Graphics Forum, volume 29, Wiley Online Library, 2010, pp. 753–762

[13] Landrieu L., Simonovsky M.. Large-scale pointcloud semantic segmentation with superpoint graphs.arXivpreprint arXiv:1711.09869, 2017

[14] Gall Y.L., Thomas H., Goulette F., Deschaud J., Marcotegui B. Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods. 2018 International Conference on 3D Vision (3DV), Sep 2018, Verone, Italy. 10.1109/3DV.2018.00052

[15] McInerney T., Terzopoulos D. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4d image analysis.Computerized Medical Imaging and Graphics, 19(1), 1995, pp.69–83

[16] Lorenzo-Valds M., Sanchez-Ortiz G.I., Gelkington A., Mohiaddin R.H., Rueckert D.. Segmentation of 4D cardiac mr images using a probabilistic atlas and the em algorithm.Medical Image Analysis, 8(3), 2004, pp.255–265

[17] Choy C.B., Xu D., Gwak J., Chen K., Savarese S. 3D-r2n2: A unified approach forsingle and multi-view 3D object reconstruction. InProceed-ings of the European Conference on Computer Vision (ECCV),2016

[18] Zeng A.,Song S., Niener M., Fisher M.,Xiao J., Funkhouser T. 3dmatch: Learning the matching of local 3D geometry in range scans. InCVPR, 2017.

[19] Zhao Y., Li X., Zhang W., Zhao S., Makkie M., Zhang M., Li Q., Liu T. Modeling 4d fMRI data via spatio-temporal convolutional neural networks(st-cnn).arXiv preprint arXiv:1805.12564, 2018

[20] Choy C., Gwak J., Savarese S. 4D Spatio-Temporal ConvNet: Minkowski Convolutional Neural Network, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019

[21] Armeni I., Sener O., Zamir A.R., Jiang H., Brilakis I., Fischer M., Savarese S.. 3d semantic parsing of large-scale indoor spaces. InProceedings of theIEEE International Conference on Computer Vision and Pattern Recognition, 2016

[22] Riemenschneider H., Bdis-Szomor A., Weissenberg J., Gool L. Learning where to classify inmulti-view semantic segmentation. InEuropean Conference on Computer Vision. Springer, 2014

[23] Ros G., Sellart L., Materzynska J., Vazquez D., Lopez A.M. The synthia dataset: Alarge collection of synthetic images for semantic segmentation of urban scenes. InThe IEEE Conference on ComputerVision and Pattern Recognition (CVPR), 2016

[24] Menze B.H., Jakab A., Bauer S., Kalpathy-Cramer J., Farahani K., Kirby J.,Burren Y., Porz N., Slotboom J., Wiestet R., "The multimodal brain tumorimage segmentation benchmark (BRATS),"IEEE Transactions on Medical Imag-ing, vol. 34, no. 10, 2015, pp. 1993–2024

Колесников Алексей Александрович — к.т.н., доцент кафедры картографии и геоинформатики Сибирского государственного университета геосистем и технологий;
e-mail: alexeykw@yandex.ru;
Кикин Павел Михайлович — к.т.н., инженер научно-образовательного центра "Газпромнефть-Политех"Санкт-Петербургского Политехнического Университета Петра Великого;
e-mail: it-technologies@yandex.ru.
Дата поступления — 31 мая 2019 г.

MICOSI: ОТКРЫТЫЙ ПАКЕТ ДЛЯ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ МИТОЗА В СТАДИЯХ ПРОМЕТАФАЗЫ, МЕТАФАЗЫ И АНАФАЗЫ

М. А. Кривов[1], П. С. Иванов[1]

[1] *Московский государственный университет им. М.В. Ломоносова, 119991, Москва*

УДК 519.688
DOI: 10.24411/9999-016A-2019-10040

В работе рассмотрена концепция разрабатываемого авторами открытого пакета для численного моделирования трёх последовательных фаз митоза. Он основан на трёхмерной мета-модели клетки, которая эволюционно расширяется за счёт встраивания других моделей микроуровня, в том числе конкурирующих. Описаны ключевые компоненты пакета, такие как расчётный модуль, визуализатор и библиотеки для интеграции с языком программирования C#. Проиллюстрирована схема работа с пакетом на примере оценки числа меротельных зацеплений микротрубочек в метафазе.

Ключевые слова: компьютерное моделирование, митоз, деление клетки, прометафаза, метафаза, анафаза.

Введение

Актуальной проблемой биофизики является изучение процессов, протекающих при митотическом делении клетки. Сегодня, благодаря различным методам синхронизации делящихся культур, имеются достаточно точные описания каждой фазы митоза. С другой стороны, до сих пор отсутствует детальное понимание механизмов, позволяющих клетке при всем многообразии протекающих в ней процессов в результате митоза перейти в строго ожидаемое состояние и разделиться на две новых.

К настоящему времени предложено множество математических моделей [1, 2], описывающих отдельные стадии митоза на уровне как всей клетки, так и её частей. Получаемые результаты численного моделирования в большинстве случаев верифицируются путём сопоставления с данными экспериментов, однако это не всегда позволяет исключить конкурирующие интерпретации одних и тех же явлений. Среди причин, приводящих к появлению альтернативных моделей, можно отметить ограниченность временного масштаба, а также упрощение взаимосвязей между частями клетки, которые заменяются искусственными условиями или просто игнорируются.

В статье представлены промежуточные результаты разработки расширяемой мета-модели, описывающей происходящие в клетке изменения в рамках трёх последовательных фаз митоза и реализованной в виде программного комплекса с открытым исходным кодом [3]. В отличие от схожих работ, подобная концепция «модели-конструктора» позволяет оценить долгосрочный эффект от выбора конкретных моделей и параметров микроуровня, например, проверить, как влияет рост тубулиновых микротрубочек на процесс центрирования хромосом перед их расхождением.

1 Исследуемый процесс

Основной задачей митоза является разделение генетической информации, содержащейся в клеточном ядре, на две строго идентичные копии. Рассматриваемое явление сопровождается различными биохимическими реакциями и сложными белок-белковыми взаимодействиями, поэтому часто его рассматривают как процесс,

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта №19-07-01164а

ISBN 978-5-901548-42-4

протекающий одновременно в нескольких пространственно-временных масштабах. При этом пространственный масштаб должен охватывать как микрометры (процессы на уровне всей клетки), так и нанометры (белок-белковые взаимодействия). Характерные же времена отдельных митотических процессов находятся в диапазоне от нескольких секунд (события типа сцепления белков) до десятков минут и даже часов (механическое центрирование хромосом).

В протекании митоза на уровне клетки выделяют пять основных этапов (Рис. 1):

1. Профаза (Рис. 1a). В клеточном ядре происходит конденсация хромосом, в результате чего в каждой из них проявляются две связанные сестринские хроматиды (т. е. будущие хромосомы, «склеенные» в пары), в сумме содержащие два идентичных набора ДНК. Независимо от этого в пространстве клетки происходит формирование веретена деления — полимерной структуры, которая в дальнейшем отвечает за разделение хромосом и их расхождение к полюсам веретена деления.

2. Прометафаза (Рис. 1b). Ядерная мембрана разрушается, и пары хромосом начинают дрейфовать в пространстве клетки. В то же время в веретене деления выделяются два полюса, которые расходятся в диаметрально противоположные концы клетки. От каждого полюса в случайных направлениях начинают расти тонкие тубулиновые микротрубочки, которые при соударении с хромосомами либо прикрепляются к ним и тянут их к соответствующему полюсу, либо отталкивают их. Конкретный сценарий зависит от места соударения: центральная часть хромосом, именуемая кинетохором, обеспечивает зацепление, а белки хромокинезина, содержащиеся в боковых частях (так называемые руки хромосомы), приводят к появлению отталкивающих сил.

3. Метафаза (Рис. 1c). За счёт достижения баланса притягивающих и отталкивающих сил все пары хромосом располагаются в экваториальной плоскости. Для перехода клетки в данное устойчивое состояние требуется продолжительное время, которое может достигать нескольких часов и на протяжении которого наблюдаются осцилляции. Другой особенностью данного этапа является достижение хромосомами нужной ориентации, в соответствии с которой каждая из хроматид «направлена» на один из полюсов.

4. Анафаза (Рис. 1d). Одновременно во всех парах хромосом происходит разрыв центромеры — эластичной структуры, связывающей хромосомы в районе кинетохора. Как следствие, ставшие независимыми хромосомы под действием сил со стороны микротрубочек притягиваются к одному из полюсов, что обеспечивает требуемое распределение генетической информации по двум будущим дочерним клеткам.

5. Телофаза (Рис. 1e). На финальной фазе митоза происходит разборка веретена деления и формирование ядерных мембран вокруг каждой группы хромосом. Таким образом, в клетке образуются два новых ядра, которые потом на стадии цитокинеза (этап, следующий за митозом) станут ядрами двух дочерних клеток.

Если не учитывать белок-белковые взаимодействия, то изменения в ходе отдельных этапов выглядят случайными и хаотичными, тем не менее в результате пяти этапов митоза клетка практически всегда приходит к строго ожидаемому состоянию.

В настоящей работе рассматриваются три ключевые фазы митоза — прометафаза, метафаза и анафаза. Они представляют наибольший научный интерес, так как из-за возникающих на этих этапах ошибок клетка может стать анеуплоидной [4], что в ряде случаев приводит к развитию злокачественных опухолей [5]. Стоит отметить, что механизмы подавления подобных ошибок могут быть различны и в целом недостаточно изучены [6]. В этой связи понимание фундаментальных механизмов митоза может стать толчком к разработке новых классов лекарственных препаратов, замедляющих неконтроллируемое деление клеток-мишеней, что и определяет востребованность подобных исследований.

2 Идея проекта

Большинство предложенных математических моделей описывают отдельные аспекты митоза, например, рост и сокращение микротрубочек [7], их взаимодействие с хромосомами [8], осцилляция хромосом в пространстве клетки [9], их распределение по новым ядрам [10]. При всём разнообразии предлагаемых идей, можно выделить два ключевых подхода к такому моделированию:

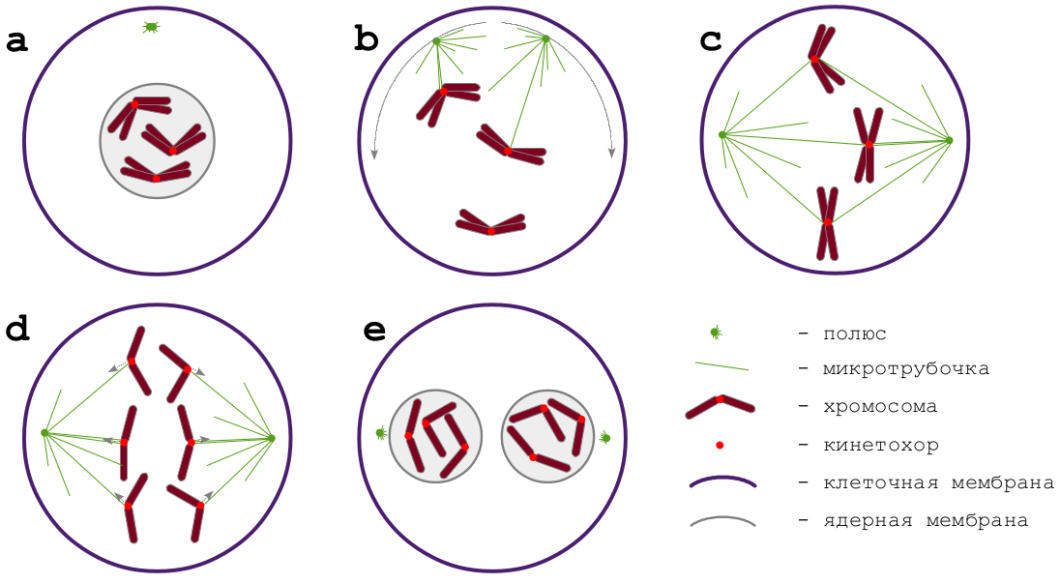


Рис. 1: Стадии митоза. (a) – профаза, (b) – прометафаза, (c) – метафаза, (d) – анафаза, (e) - телофаза

1. «Поиск-и-захват» [11]. Как отмечалось выше, центрирование и разделение хромосом осуществляется благодаря зацеплению микротрубочек за кинетохор. При этом остается открытым вопрос, каким образом микротрубочки находят столь небольшие области в пространстве клетки. Например, первые зацепления могут наблюдаться уже на десятой минуте, а к концу метафазы к каждому кинетохору прикрепляются десятки микротрубочек. Согласно принципу «поиск-и-захват», микротрубочки растут в случайных направлениях, постоянно переходя из состояния полимеризации (роста) в состояние деполимеризации (сокращения), и обратно. Именно благодаря подобной нестабильности они «сканируют» всё пространство клетки и обнаруживают хромосомы.

2. «Баланс сил» [12]. Расположение хромосом в экваториальной плоскости, в соответствии с данным принципом, обеспечивается исключительно за счёт достижения баланса действующих на них сил, в то время как сами хромосомы свободно перемещаются в вязкой цитоплазме клетки. Набор учитываемых сил зависит от модели, но обычно включает в себя притягивающие и отталкивающие силы со стороны микротрубочек, эффекты броуновского движения и вязкости цитоплазмы, а также силу растяжения центромеры.

Идея выполняемого авторами проекта заключается в разработке на базе подобных принципов самодостаточной каркасной мета-модели, которая непрерывно описывает динамику клетки в стадиях прометафазы, метафазы и анафазы. Подобная мета-модель допускает расширение путем встраивания сторонних моделей микроуровня, например, задающих скорости роста и сокращения микротрубочек или определяющих конкретный список действующих сил. Предложенная мета-модель реализуется в виде открытого пакета MiCoSi и дополнительно снабжается набором базовых виртуальных экспериментов с заранее известными ожидаемыми результатами.

Таким образом, построенная мета-модель позволяет исследователю сконструировать виртуальную клетку, задав биофизические параметры и описав взаимодействия её отдельных частей. После этого у него появляется возможность запустить стандартные вычислительные эксперименты и оценить, насколько в них воспроизводятся факты о митозе, установленные опытным путём. Независимо от этого, он может составить свою программу для выгрузки требуемой информации с целью проверки целевых гипотез. По мнению авторов, подобный инструмент будет полезен для изучения конкретных культур клеток и определения их биологических особенностей, а также для более комплексного сравнения конкурирующих моделей митоза.

3 Мета-модель

Подробное описание используемой авторами математической модели содержится в работе [13]. Данная модель относится к классу геометрических, так как все белковые структуры представлены в виде наиболее близких к ним геометрических фигур. Например, клеточная мембрана описывается сферой, хромосома — группой из трёх полуцилиндров, а микротрубочки — отрезками заданной длины. Настройка численных опытов под конкретную культуру клеток осуществляется за счёт задания значений для 26 варьируемых параметров. Они определяют размеры самих белковых структур, вероятности для биохимических событий (зацепление микротрубочки за кинетохор, переход микротрубочки из состояния роста в состояние сокращения и др.), а также характер белок-белковых взаимодействий (силы, точки их применения, изменение формы).

В начальный момент времени, соответствующий прометафазе, полюса деления расходятся по искусственному закону в диаметрально противоположные стороны клетки. Рост микротрубочек реализован в аналогии с [14] и задаётся четырьмя параметрами — скоростями полимеризации и деполимеризации, а также вероятностями катастрофы и восстановления (т. е. событий перехода из состояния роста в состояние сокращения, и наоборот). При наложении микротрубочки на хромосому она может либо оборваться, либо перейти в состояние деполимеризации. Если плюс-конец микротрубочки накладывается на кинетохор, то с определённой вероятностью происходит зацепление. В свою очередь, также имеются вероятностные события для отцепления микротрубочек.

Метафаза наступает после окончания расхождения полюсов и продолжается до момента разрыва центромеры, представляемой стержнем или гуковской пружинкой. За это время хромосомы перемещаются в пространстве клетки под действием трёх сил: притяжения со стороны микротрубочек, тепловой диффузии и сжатия центромеры. В случае, если требуется обеспечить центрирование пар хромосом искусственным образом, то имеется возможность ограничить их движение только экваториальной плоскостью (что равносильно подавлению осцилляции). Разрыв центромер, соответствующий началу анафазы, происходит либо при достижении пороговой силы, либо при зацеплении требуемого числа микротрубочек за каждый кинетохор.

Как отмечалось ранее, предложенная модель выступает в роли каркаса, в который могут встраиваться более сложные модели микроуровня. Среди подобных механизмов, которые планируется внедрить в первую очередь, стоит отметить предположение о неявном влиянии концентрации свободных белков тубулина на направления роста микротрубочек [16], взаимодействие микротрубочек с белками хромокинезина [9], а также возможность зацепления микротрубочки за хромосому путём сцепления с другой небольшой микротрубочкой, порождённой самим кинетохором [8]. Стоит отметить, что хотя в текущей мета-модели подобные клеточные механизмы ещё не отражены, она уже способна воспроизвести основные фактов о митозе. Например, авторы повторили виртуальный опыт по оценке влияния геометрических параметров клетки на подавление меротельных (ошибочных) зацеплений [15], уточнив ряд сделанных ранее выводов [13].

4 Структура пакета

В соответствии с используемым численным алгоритмом [13], переход от времени t к $t + \Delta t$ разделён на два этапа. На первом из них осуществляется логика выполнения белок-белковых взаимодействий и вероятностных событий, в результате чего становятся известными значения и точки приложения действующих на хромосомы сил. Второй этап заключается в обновлении состояния клетки и самом переходе на следующий временной слой. Для этого достаточно решить систему линейных уравнений с матрицей 6x6, что позволяет по списку сил определить текущие значения линейных и угловых скоростей. Далее, интегрируя найденные скорости с помощью метода Эйлера, удаётся вычислить новые координаты белковых структур для времени $t + \Delta t$.

Так как в модель уже внесена искусственная погрешность, отвечающая броуновскому движению молекул, то первый порядок точности численного метода не является недостатком. Также стоит отметить, что основные операции, на которые уходит наибольшая часть времени, заключаются в проверке наложения геометрических объектов друг на друга, а не в решении СЛАУ или в интегрировании скоростей. На базе данного численного метода разработан открытый пакет MiCoSi [3], состоящий из пяти компонент:

1. Расчётный модуль (Рис. 2а). Консольная программа Mitosis.exe реализует численный алгоритм и позволяет по заданным файлам проекта провести серию виртуальных опытов. Результаты сохраняются в

виде бинарных файлов .cell, содержащих полную информацию о клетке. Предполагается, что работа с данным модулем осуществляется через утилиты-помощники, которые подготавливают файлы проекта и вызывают расчётный модуль с нужными аргументами командной строки.

2. .NET интерфейсы. Две библиотеки для .NET-совместимых языков программирования (в первую очередь, C#) предназначены для настройки виртуальных опытов и анализа их результатов. Первая из них, Mitosis.API, содержит классы для загрузки результатов моделирования из файлов .cell. Вторая, Mitosis.Launcher, позволяет описать требуемый опыт, запустить соответствующие расчёты и напрямую из C# отслеживать ход их выполнения.

3. Визуализатор (Рис. 2b). Полноценная программа для воспроизведения результатов моделирования, которая позволяет просматривать файлы .cell в режиме плеера, выделять объекты и замерять их параметры. Дополнительно к этому, имеется возможность выгрузки интересующих моментов в виде картинок или видео-фрагментов формата .wmv.

4. Функциональные тесты (Рис. 2c).Набор из 42 автоматических тестов построен на базе библиотеки Google Tests и разделён на три уровня. В первом из них работа с пакетом осуществляется в режиме «чёрного ящика», т. е. проверяется только сам факт успешного или аварийного завершения расчётов. Тесты второго уровня проверяют корректность моделирования на простейших сценариях, например, произошло ли зацепление микротрубочки за кинетохор, соответствует ли растяжение центромеры прикладываемым силам и т. д.. Наконец, третий уровень содержит более комплексные проверки, в частности, на нём осуществляется сбор статистической информации о клетке для дальнейшего сопоставления полученных данных с ожидаемыми.

5. Эксперименты. В настоящий момент в проекте реализованы два виртуальных эксперимента по оценке числа меротельных зацеплений и осцилляции хромосом. Они представлены в виде консольной программы, которая для заданной конфигурации клетки выполняет соответствующие расчёты, выдавая на выходе ответ о переходе ансамбля целевых клеток в ожидаемое состояние (в формате «да или нет»), а также .csv таблицы с более подробной информацией о числе меротельных зацеплений и амплитудах колебаний соответственно.

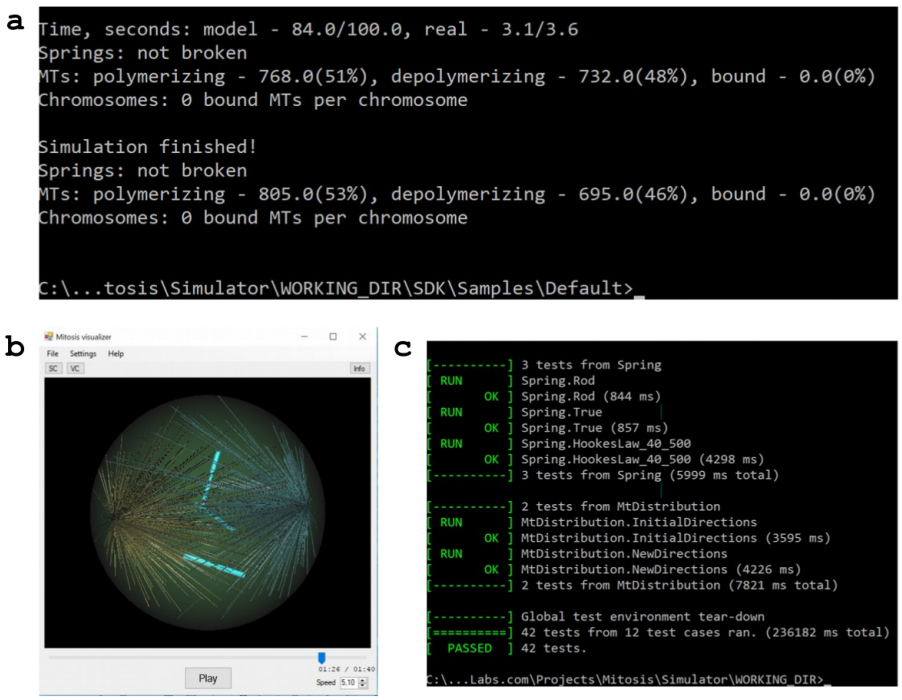


Рис. 2: Интерфейсы пакета MiCoSi. (a) – модуль для проведения расчётов, (b) — визуализатор результатов моделирования, (c) — набор встроенных автоматических тестов

5 Работа с пакетом

Схему использования пакета можно проиллюстрировать на примере работы виртуального эксперимента MerotelicKTs, в котором анализируется влияние геометрических особенностей клетки на подавление меротельных зацеплений в метафазе (версия от 16 апреля 2019 года). Для этого необходимо выполнить пять следующих шагов:

1. Подготовка. Требуется провести клонирование исходных кодов из git-репозитория, после чего осуществить автоматическую настройку с помощью утилиты CMake. В случае, если на системе установлена библиотека Google Tests, то также будут созданы правила сборки для функциональных тестов.

2. Компиляция. Подготовленный через утилиту CMake проект необходимо скомпилировать с помощью среды Visual Studio. В случае успеха все требуемые файлы будут помещены в директорию WORKING_DIR, непосредственно из которой можно осуществлять работу с пакетом. В частности, рекомендуется запустить функциональные тесты (приложение FunctionalTests), чтобы оценить работоспособность сборки.

3. Настройка опыта. К собственному C# проекту следует подключить библиотеки Mitosis.API и Mitosis.Launcher. С помощью содержащегося в них класса LaunchParameters можно задать значения для 26 параметров клетки (поле Config), описать начальное состояние клетки (поле InitialStates) и закон, по которому двигаются полюса (поле PoleCoords), а также переопределить аргументы командной строки решателя (например, указать зерно генератора случайных чисел). В рассматриваемом эксперименте этим операциям соответствуют строчки 48-108.

4. Анализ результатов. Чтобы запустить расчёты по ранее подготовленному сценарию, требуется задействовать класс Launcher (строчки 109-114). Далее, с помощью класса TimeStream необходимо открыть полученный файл формата .cell и обойти все временные слои, собирая нужную информацию о клетке (строчки 122-175). Для каждой белковой структуры имеется соответствующий ей класс (Pole, Chromosome, MT), через поля которого можно узнать значения её параметров. Непосредственный анализ результатов зависит от цели виртуального опыта — в рассматриваемом примере измеряется число мерoтельных зацеплений и угол поворота пары хромосом (строчки 183-331), которые в дальнейшем сохраняются в табличном формате .csv.

5. Визуализация. Крайне важным моментом является визуальная проверка корректности опыта, поэтому требуется дополнительно запустить визуализатор (приложение Visualizer) и открыть созданный файл .cell, проследив за динамикой клетки. Это позволит удостоверится в том, что настройки опыта верны, а также что в нём отсутствуют аномалии, которые могли быть упущены при экспорте данных в упрощённый табличный формат.

Заключение

Целью представленного проекта является создание настраиваемой математической модели делящейся клетки, в которую авторами поэтапно встраиваются сторонние модели микроуровня, в подавляющем большинстве случаев представленные только в виде статей. Ключевой особенностью является реализация данной мета-модели в виде пакета с открытым исходным кодом, что должно позволить исследователям проводить более качественный анали экспериментально наблюдаемых клеточных механизмов.

На текущем этапе развития предложенная модель уже достаточна для моделирования трёх последовательных фаз митоза — прометафазы, метафазы и анафазы. Так как в пакете предусмотрена возможность гибкой настройки виртуальных опытов за счёт интеграции с языком программирования C#, то это делает возможным его использование при изучении различных явлений, например, таких, как подавление мерoтельных зацеплений микротрубочек за кинетохор в рамках метафазы.

Список литературы

[1] McIntosh R. et. al., Quarterly Reviews of Biophysics. 2012. Vol. 45(2). P. 147–207.

[2] Civelekoglu-Scholey G., Cimini D., Modelling chromosome dynamics in mitosis: a historical perspective on models of metaphase and anaphase in eukaryotic cells // Interface Focus. 2014. Vol. 4(3). P. 1–9.

[3] Портал проекта MiCoSi: https://github.com/m-krivov/MiCoSi (дата обращения: 01.06.2019).

[4] Cimini D., Merotelic kinetochore orientation, aneuploidy, and cancer // Biochimica et Biophysica Acta (BBA) – Reviews on Cancer. 2008. Col. 1786(1). P. 32–40.

[5] Heim S., Mitelman F., Cancer Cytogenetics: Chromosomal and Molecular Genetic Aberrations of Tumor Cells, Fourth Edition. Wiley. 2015. P. 632.

[6] Gregan J., Merotelic kinetochore attachment: causes and effects // Trends in Cell Biology. 2011. Vol. 21(6). P. 374–381.

[7] Gregoretti I., et al., Insights into cytoskeletal behavior from computational modeling of dynamic microtubules in a cell-like environment // Journal of Cell Science. 2006. Vol. 119. P. 4781–4788.

[8] Vasileva V. et al., Molecular mechanisms facilitating the initial kinetochore encounter with spindle microtubules // Journal of Cell Biology. 2017. Vol. 216(6). P. 1609–1622.

[9] Campas O., Sens P., Chromosome Oscillations in Mitosis // Physical Review Letters. 2006. Vol. 97(12). P. 128102-1–128102-4.

[10] Gerlich D. et al., Global Chromosome Positions Are Transmitted through Mitosis in Mammalian Cells // Cell. 2003. Vol. 112. P. 751–764.

[11] Kirschner M., Mitchison T., Beyond self-assembly: From microtubules to morphogenesis // Cell. 1986. Vol. 45. P. 329–342.

[12] Joglekar A., Hunt A., A Simple, Mechanistic Model for Directional Instability during Mitotic Chromosome Movements // Biophysical Journal. 2002. Vol. 83(1). P. 42–58.

[13] Кривов М.А. и др., Моделирование деления биологических клеток в стадии метафазы на суперкомпьютере "Ломоносов-2"// Вычислительные методы и программирование: Новые вычислительные технологии. 2018. Том 19, С. 327–339.

[14] Gliksman N., Skibbens R., et al., How the Transition Frequencies of Microtubule Dynamic Instability (Nucleation, Catastrophe, and Rescue) Regulate Microtubule Dynamics in Interphase and Mitosis: Analysis Using a Monte Carlo Computer Simulation // Molecular Biology of the Cell. 1993. Vol. 4. P. 1035–1050.

[15] Zaytsev A., Grishchuk E., Basic mechanism for biorientation of mitotic chromosomes is provided by the kinetochore geometry and indiscriminate turnover of kinetochore microtubules // Molecular Biology of the Cell. 2015. Vol. 26. P. 3985–3998.

[16] Wollman R., et al., Efficient Chromosome Capture Requires a Bias in the 'Search-and-Capture' Process during Mitotic-Spindle Assembly // Current Biology. 2005. Vol. 15. P. 828–832.

Кривов Максим Андреевич — математик факультета ВМК МГУ им. М. В. Ломоносова
e-mail: m_krivov@cs.msu.su;
Иванов Павел Сергеевич — к.ф.-м.н., с.н.с. физического факультета МГУ им. М. В. Ломоносова
e-mail: psmart@rambler.ru;
Дата поступления — 01 июня 2019 г.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТУРБУЛЕНТНЫХ ТЕЧЕНИЙ В УСТОЙЧИВО СТРАТИФИЦИРОВАННОЙ СРЕДЕ: ЦИРКУЛЯЦИЯ НАД ГОРОДСКИМ ОСТРОВОМ ТЕПЛА

Л. И. Курбацкая

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск

УДК 532.517.4, 551.511.32
DOI: 10.24411/9999-016A-2019-10041

Аннотация Представлены результаты численного моделирования турбулентной структуры проникающей конвекции над городским островом тепла малого относительного удлинения в устойчиво стратифицированной покоящейся среде. Использованы представления градиентной диффузии для турбулентных потоков импульса и тепла, зависящие от трех параметров — кинетической энергии турбулентности, скорости её спектрального расходования и дисперсии температурных флуктуаций. Эти параметры находятся из замкнутых дифференциальных уравнений баланса в RANS приближении описания турбулентности. Распределения осредненных полей скорости и температуры, турбулентных характеристик хорошо согласуются с данными измерений в контролируемом лабораторном эксперименте. Показано, что термическая стратификация вызывает различия в коэффициентах турбулентного обмена между активным скаляром теплом и пассивным скаляром массой. Различие означает, что предположение об идентичности коэффициентов турбулентной диффузии тепла и массы, используемое в обычных моделях турбулентности, дает существенные ошибки в оценивании тепло и массопереноса в термически устойчиво стратифицированной среде.

Ключевые слова: турбулентность, планетарный пограничный слой, городской остров тепла, крупномасштабная циркуляция, численное моделирование.

Введение

Циркуляция воздуха над городским островом тепла генерируется за счет энергии антропогенных источников в пределах городской черты. Наиболее интенсивной такая циркуляция наблюдается в ночное время при ясном небе и слабом окружающем ветре. Фундаментальными характеристиками, отражающими структуру ночного городского острова тепла, являются распределения полей скорости и температуры, а также интенсивностей турбулентности. Вертикальный турбулентный тепловой факел от локализованного источника (острова тепла) и, связанная с ним циркуляция, развиваются вследствие различия температуры между островом тепла и его окружением, которое имеет более низкую температуру. Интерес также представляет исследование механизма тепло и массопереноса в течениях окружающей среды. Течения в атмосферном пограничном слое часто стратифицированы по плотности (температуре), и плавучесть оказывает сильное воздействие на перенос скаляра (тепло, масса). Для описания проникающей конвекции, индуцируемой тепловым источником (островом тепла) и воспроизведения полей скорости, температуры и структуры турбулентности используется RANS приближение. Основные уравнения термогидродинамики берутся в гидростатическом приближении [1] в безразмерном виде для осредненных значений скорости и температуры в приближении Буссинеска. Для численного решения основных уравнений необходимо моделировать турбулентные потоки импульса и тепла. Для этих целей используется трехпараметрическая $E-\varepsilon-<\theta^2>$ модель

Работа была выполнена в рамках госзадания Института вычислительной математики и математической геофизики СО РАН № 0315-2019-0004 и частичной поддержки Российского фонда фундаментальных исследований (код проекта № 17-01-00137) и Российского фонда фундаментальных исследований и Правительства Новосибирской области (код проекта 18-48-540005).

ISBN 978-5-901548-42-4

турбулентности. В этой модели кинетическая энергия турбулентности (КЭТ) $E =< u_i u_i >$, скорость ее диссипации ε и дисперсия флуктуаций температуры θ^2 находятся из дифференциальных уравнений переноса, а вектор турбулентного потока импульса и тепла определяются с помощью явной алгебраической модели, полученной из дифференциальных уравнений переноса в приближении локально равновесной турбулентности. Проникающая турбулентная конвекция, индуцируемая постоянным потоком тепла H_0 с поверхности пластины диаметра D моделирует в лабораторном эксперименте прототип городского острова тепла малого относительного удлинения $(z_i/D << 1)$ [4]. Движение среды рассматривается как обладающее осевой симметрией. Подробности численного моделирования проникающей турбулентной конвекции и связанной с ней циркуляцией могут быть найдены в [2] и [3], и здесь они не приводятся. Ниже будут представлены только результаты численного моделирования, описывающие основные характеристики проникающей турбулентной конвекции и связанной с ней циркуляцией. Основное внимание в статье было уделено описанию турбулентной диффузии пассивного скаляра (примеси).

1 Модель турбулентной диффузии примеси над нагретым источником

Для описания рассеяния пассивной примеси базовая мезомасштабная модель турбулентного атмосферного пограничного слоя должна быть дополнена уравнениями для осредненной концентрации C, вектора турбулентного потока примеси $< u_i c >$ и корреляции между флуктуациями концентрации и температуры $< c\theta >$. Поскольку результаты моделирования переноса активной примеси (тепла) над нагретым источником (в данном случае диском) [2], [3] удовлетворительно согласуются с данными измерений [4], для турбулентного потока пассивного скаляра (массы вещества) также может быть сформулирована полностью явная анизотропная алгебраическая модель при упрощении дифференциального уравнения переноса для вектора турбулентного потока скаляра $< u_i c >$ и корреляции $< c\theta >$ в приближении локально равновесной турбулентности. Таким образом, определяющие уравнения для поля концентрации пассивного скаляра в цилиндрической системе координат имеют вид [5]:

$$\frac{\partial C}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r} r C U_r + \frac{\partial}{\partial z} C U_z = -\frac{1}{r}\frac{\partial}{\partial r} r < u_r c > - \frac{\partial}{\partial z} < u_z c >, \tag{1}$$

$$\begin{aligned}\frac{\partial < u_r c >}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r} r < u_r c > U_r + \frac{\partial}{\partial z} < u_r c > U_z = - < u_r^2 > \frac{\partial C}{\partial r} - < u_r u_z > \frac{\partial C}{\partial z} - < u_r c > \frac{\partial U_r}{\partial r} - \\ - < u_z c > \frac{\partial U_z}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(r\alpha_{1s}\frac{E}{\varepsilon} < u_r^2 > \frac{\partial < u_r c >}{\partial r}) + \frac{1}{r}\frac{\partial}{\partial r}(r\alpha_{1s}\frac{E}{\varepsilon} < u_r u_z > \frac{\partial < u_r c >}{\partial z}) + \\ + \frac{\partial}{\partial z}(r\alpha_{1s}\frac{E}{\varepsilon} < u_r u_z > \frac{\partial < u_r c >}{\partial r}) + \frac{\partial}{\partial z}(r\alpha_{1s}\frac{E}{\varepsilon} < u_z^2 > \frac{\partial < u_r c >}{\partial z} - \alpha_{1c}\frac{\varepsilon}{E} < u_r c >,\end{aligned} \tag{2}$$

$$\begin{aligned}\frac{\partial < u_z c >}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r} r < u_z c > U_z + \frac{\partial}{\partial z} < u_z c > U_z = - < u_z^2 > \frac{\partial C}{\partial r} - < u_r u_z > \frac{\partial C}{\partial r} - < u_r c > \frac{\partial U_z}{\partial r} - \\ - \alpha_{1c}\frac{\varepsilon}{E} < u_z c > - < u_z c > \frac{\partial U_z}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(r\alpha_{1s}\frac{E}{\varepsilon} < u_r^2 > \frac{\partial < u_z c >}{\partial r}) + \frac{1}{r}\frac{\partial}{\partial r}(r\alpha_{1s}\frac{E}{\varepsilon} < u_r u_z > \frac{\partial < u_z c >}{\partial z}) + \\ + \frac{\partial}{\partial z}(r\alpha_{1s}\frac{E}{\varepsilon} < u_r u_z > \frac{\partial < u_r c >}{\partial r}) + \frac{\partial}{\partial z}(r\alpha_{1s}\frac{E}{\varepsilon} < u_z^2 > \frac{\partial < u_z c >}{\partial z} - (1-\alpha_{2c}) < c\theta > \frac{g}{\Theta_0},\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{\partial < c\theta >}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}\left[r < c\theta > U_r\right] + \frac{\partial}{\partial z}\left[< c\theta > U_z\right] == -\alpha_{3c}\frac{\varepsilon}{E} < c\theta > + \frac{1}{r}\frac{\partial}{\partial r}\left[r\alpha_{2s}\frac{\nu_t}{Sc_t}\right] \cdot \frac{\partial < c\theta >}{\partial r} \\ + \frac{\partial}{\partial z}\left[\alpha_{2s}\frac{\nu_t}{Sc_t}\right] \cdot \frac{\partial < c\theta >}{\partial z} - \left\{< u_r\theta > \frac{\partial C}{\partial r} + < u_z\theta > \frac{\partial C}{\partial z} + < u_r c > \frac{\partial \Theta}{\partial r} + < u_z c > \frac{\partial \Theta}{\partial z}\right\},\end{aligned} \tag{4}$$

$$\begin{aligned}- < u_j c >= C_D \frac{E^2}{\varepsilon}\sqrt{2R}\frac{\partial C}{\partial x_i} - \frac{\sqrt{2R}}{C_{1\theta}}\frac{E}{\varepsilon} \cdot \left\{[2\nu_t + (1-\alpha_{2C})D_t]\frac{1}{2}(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}) \times \right. \\ \left. \times (1-C_{2\theta})D_t\frac{1}{2}(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i})\right\}\frac{\partial C}{\partial x_j} + \frac{1-\alpha_{2c}}{C_{1c}}\frac{E}{\varepsilon}\frac{g}{\Theta_0} < c\theta > .\end{aligned} \tag{5}$$

В уравнениях (1)–(5): U_r — средняя горизонтальная скорость, U_z — средняя вертикальная скорость (вид уравнений для осредненных скоростей и температуры приведен в [5]), $< u_r c >$, $< u_z c >$ — вертикальный и горизонтальный потоки массы вещества, u_r, u_z — горизонтальная и вертикальная турбулентная флуктуация скорости, $< u_r^2 >$, $< u_z^2 >$ — горизонтальная и вертикальная составляющие КЭТ, $< u_r\theta >$, $< u_z\theta >$ — вертикальный и горизонтальный турбулентные потоки тепла, Θ_0 — средняя температура, $D_t = C_D\sqrt{2R}E^2/\varepsilon$ — коэффициент турбулентной диффузии скаляра, $\nu_t = C_\mu E^2/\varepsilon$ — турбулентная вязкость, $C_D = C_\mu/Sc_t$, $R = \tau_c/\tau$ — параметр отношения характерных масштабов времени скалярного и динамического турбулентных полей. Численные значения констант модели $C_\mu = 0.095$, $Sc_t = 0.9$, $R = 0.6$, $C_{1\theta} = 3.28$, $C_{2\theta} = 0.5$, $C_{1c} = 0.4$, $C_{3\theta} = 0.4$, $\alpha_{1c} = 4.0$, $\alpha_{2c} = \alpha_{3c} = 0.4$, $\alpha_{1s} = \alpha_{2s} = 0.22$. Сформулированная диффузионная модель (1)–(5) используется для решения задачи о распространении пассивного скаляра (массы вещества) от непрерывного источника заданной производительностью Q, расположенного на поверхности. На источнике задается постоянный вертикальный поток массы вещества

$$-D_m(\partial C/\partial z) = H_c \tag{6}$$

где $D_m = \nu/Sc$ — коэффициент молекулярной диффузии, ν — коэффициент кинематической вязкости, Sc — молекулярное число Шмидта.

2 Коэффициенты турбулентного обмена активного и пассивного скаляра

В обычных градиентно-диффузионных моделях (типа Буссинеска) вертикальные потоки тепла и массы в термически стратифицированном течении имеют вид

$$< w\theta >= -K_h\frac{\partial T}{\partial z}, \tag{7}$$

$$< wc >= -K_c\frac{\partial C}{\partial z}, \tag{8}$$

где w — турбулентная флуктуация скорости в вертикальном направлении, K_h и K_c — коэффициенты турбулентного обмена активного и пассивного скаляра.

Главное допущение в таких моделях состоит в том, что турбулентные потоки и массы пропорциональны локальным градиентам средней температуры и концентрации. При использовании таких моделей обычно полагается, что коэффициенты турбулентного обмена теплом K_h и массой K_c равны друг другу. Однако в [7] на основе модели турбулентности второго уровня замыкания для транспортируемых скаляров тепла и влажности показано, что эти коэффициенты не равны друг другу. Покажем различие в коэффициентах K_h и K_c. Для этого запишем уравнение для турбулентного потока тепла

$$\begin{aligned} - < u_j\theta >= C_T\frac{E^2}{\varepsilon}\sqrt{2R}\frac{\partial T}{\partial z} - \frac{\sqrt{2R}}{C_{1\theta}}\frac{E}{\varepsilon}\cdot\Big\{[2\nu_t + (1-C_{2\theta})\alpha_t]\frac{1}{2}\Big(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\Big)\times \\ \times(1-C_{2\theta})\alpha_t\frac{1}{2}\Big(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\Big)\Big\}\frac{\partial T}{\partial x_j} + \frac{1-C_{2\theta}}{\alpha_{1\theta}}\frac{E}{\varepsilon}\frac{g}{\Theta_0} < \theta^2 > . \end{aligned} \tag{9}$$

Для оценки коэффициентов K_h и K_c сохраним лишь ведущие члены с вертикальными градиентами температуры и концентрации, а также члены плавучести в соответствии с (5) и (9).

$$- < u_j\theta >= C_T\frac{E^2}{\varepsilon}\sqrt{2R}\left[(1+C_{3\theta})\frac{< w^2 >}{2/3E} - C_{3\theta}\right]\frac{\partial T}{\partial z} - \frac{1-C_{2\theta}}{C_{1\theta}}\frac{g}{\Theta_0}\sqrt{2R}\frac{E}{\varepsilon} < \theta^2 >, \tag{10}$$

$$- < u_j c >= C_D\frac{E^2}{\varepsilon}\sqrt{2R}\left[(1+\alpha_{3c})\frac{< w^2 >}{2/3E} - \alpha_{3c}\right]\frac{\partial C}{\partial z} - \frac{1-\alpha_{2c}}{\alpha_{1c}}\frac{g}{\Theta_0}\sqrt{2R}\frac{E}{\varepsilon} < c\theta > . \tag{11}$$

Деление (11) на (10) дает отношение коэффициентов

$$\frac{K_c}{K_h} = \frac{1-(1-\alpha_{2c})/(\alpha_{1c}C_D)\cdot\Delta_1^{-1}\frac{g}{\Theta_0} < c\theta > /E(\partial C/\partial z)^{-1}}{1-(1-C_{2\theta})/(C_{1\theta}C_T)\cdot\Delta_2^{-1}\frac{g}{\Theta_0} < \theta^2 > /E(\partial T/\partial z)^{-1}}, \tag{12}$$

где $K_c = \frac{-<u_j c>}{\partial C/\partial z}$, $K_h = \frac{-<u_j \theta>}{\partial T/\partial z}$ — коэффициенты турбулентного обмена для массы и тепла, через Δ_1 и Δ_2 обозначены выражения в квадратных скобках (11) и(10) соответственно. Из (12) следует, что если оба скаляра (тепло и масса) являются пассивными примесями (члены плавучести пренебрежимо малы), то $K_h = K_c$. Очевидно, что различие между коэффициентами K_h и K_c может возникать, когда один из скаляров (тепло) действует как активная примесь. Величина отклонения зависит от соотношения корреляций $< \theta^2 >$ и $< c\theta >$. Конечно, о различии в коэффициентах K_h и K_c можно судить по результатам численного моделирования, которое было проведено с использованием полных (неупрощенных) выражений для турбулентных потоков активного скаляра (9) и пассивного скаляра (5). Результаты обсуждаются ниже.

3 Результаты численного моделирования

Представленные результаты, описывающие развитие теплового факела над источником тепла, его турбулентную структуру, соответствуют квазиустановившемуся состоянию термической циркуляции от поступающего тепла от нагревателя, при котором интенсивность острова тепла $\Delta T_m = T_m - T_0$ и поверхностный поток тепла оставались неизменными с течением времени (T_m и T_0 — температура поверхности в центре острова и вне его, соответственно). На рис. 1*а* показана теневая фотография теплового факела, имеющего верхнюю границу в форме шляпы, что обусловлено воздействием устойчивой стратификации окружающей среды, которая создает запирающую инверсию в верхней части перемешанного слоя при $z/z_i \cong 1$ (z_i — высота перемешанного слоя, z — текущая высота). В лабораторном эксперименте [6] получены картины визуализаций движения воздуха над островом тепла с нагревателем в виде проволоки круглого сечения, а не круглого диска, как в экспериментах [4]. Движение воздуха было, приближенно, плоскосимметричным. (Реальную ситуацию городского острова тепла можно рассматривать, по-видимому, как комбинацию плоско-и осесимметричного движений). Общим для обоих экспериментов является подавление высоты плума устойчивой стратификацией среды, увеличение бокового движения и турбулентности плума. Движение среды над "несущей" вихревой парой (рис. 1*б*) фиксируется картинами визуализаций Noto [6] и может быть связано с локальными циркуляциями над реальным островом тепла. Тепло при этом не транспортируется от урбанизированной поверхности выше в атмосферу, а аккумулируется в локальных циркуляциях.

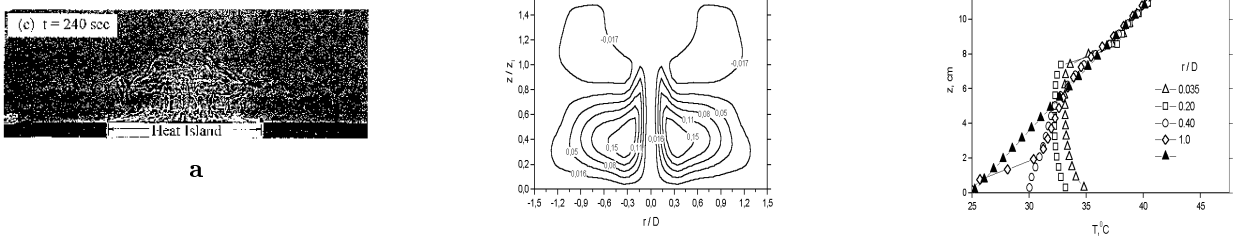
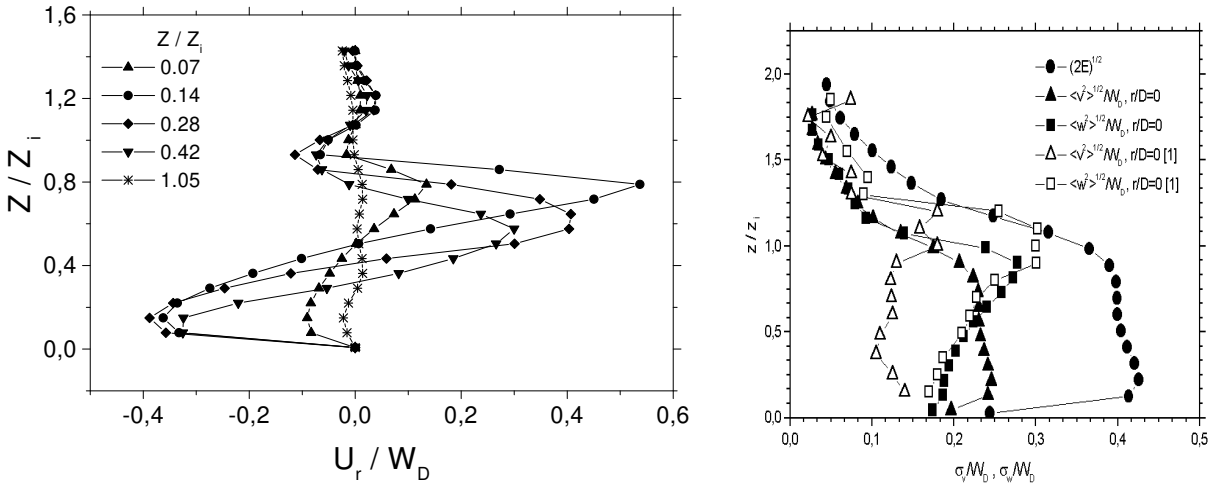


Рис. 1: Циркуляция над нагреваемым источником (теневая фотография) (*а*), картина линий тока циркуляционного течения (*б*), профили средней температуры в различных сечениях (*в*).

На рисунке 1*б* видны два крупномасштабных вихря с вращением левого вихря против часовой стрелки, правого — по часовой стрелке, которые формируют основное восходящее движение в центре, простирающееся до слоя вовлечения (z/z_i) $\cong 1$) и нисходящее движение на периферии. При этом высота теплового факела подавляется устойчивой стратификацией, а боковое движение и турбулентность факела увеличиваются. На рис. 1*в* хорошо прослеживается область сильной и положительной плавучести в нижней части теплового факела и область слабой отрицательной плавучести вблизи слоя вовлечения ($z/z_i \cong 1$). Вычисленные вертикальные распределения температуры в центре острова указывают на хорошее перемешивание в нижней и центральной части факела. Такой характер поведения распределения температуры с высотой относится к тем реальным ночным пограничным слоям, в которых преобладают неустойчивые (конвективные) условия вследствие восходящего потока тепла от урбанизированной поверхности при слабом ветре. На рис. 1*в* также

видно, что профили температуры внутри факела имеют характерное "вздутие": температура внутри факела оказывается ниже температуры вне его на той же высоте, фиксируя тем самым область отрицательной плавучести, вследствие возвышения факела в центре, связанное с эффектом перекрещивания. Поведение радиальной скорости по вертикали показывает возрастание скорости натекающего потока в нижней части теплового факела, который затухает в зоне вовлечения, скорость оттекающего потока наверху также возрастает от центра острова тепла к периферии (рис. 2а).

Структура турбулентности теплового факела представлена на рис. 2*б* в виде распределения среднеквадратичных флуктуаций горизонтальной и вертикальной турбулентных скоростей в зависимости от безразмерной высоты z/z_i в центре факела. Как данные лабораторных измерений, так и результаты вычислений показывают, что большие значения σ_u/W_D и σ_w/W_D внутри перемешанного слоя быстро убывают с высотой выше слоя вовлечения. Область вычисленного максимального значения σ_u/W_D простирается по горизонтали дальше, чем значение σ_w/W_D вследствие горизонтального дивергентного течения, вызываемого устойчивой стратификацией верхней части факела (W_D — конвективный масштаб скорости). Следует отметить, что простая градиентная модель не только правильно описывает характерные особенности распределения σ_u/W_D и σ_w/W_D, но и удовлетворительно отражает их анизотропный характер.

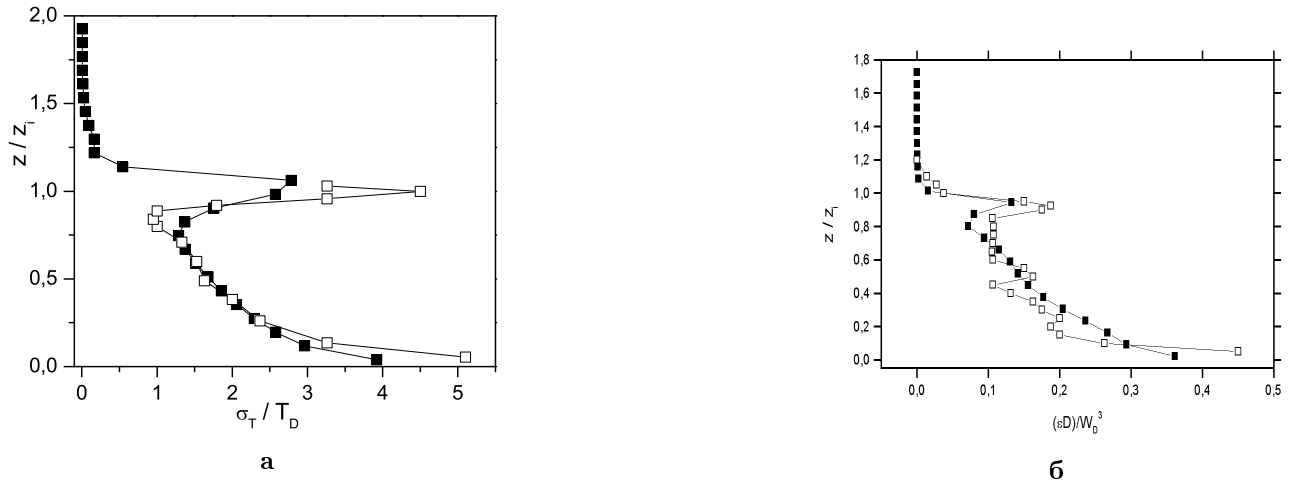


а б

Рис. 2: Профили горизонтальной средней скорости в различных сечениях по радиальной координате (*а*). Распределение горизонтальной σ_v/W_D и вертикальной σ_w/W_D дисперсий турбулентной скорости на оси факела над островом тепла (*б*). Светлые треугольники и квадраты — эксперимент [4], зачерненные треугольники и квадраты — вычисление, зачерненные окружности — вычисленный профиль $\sqrt{< u_i u_j >}$.

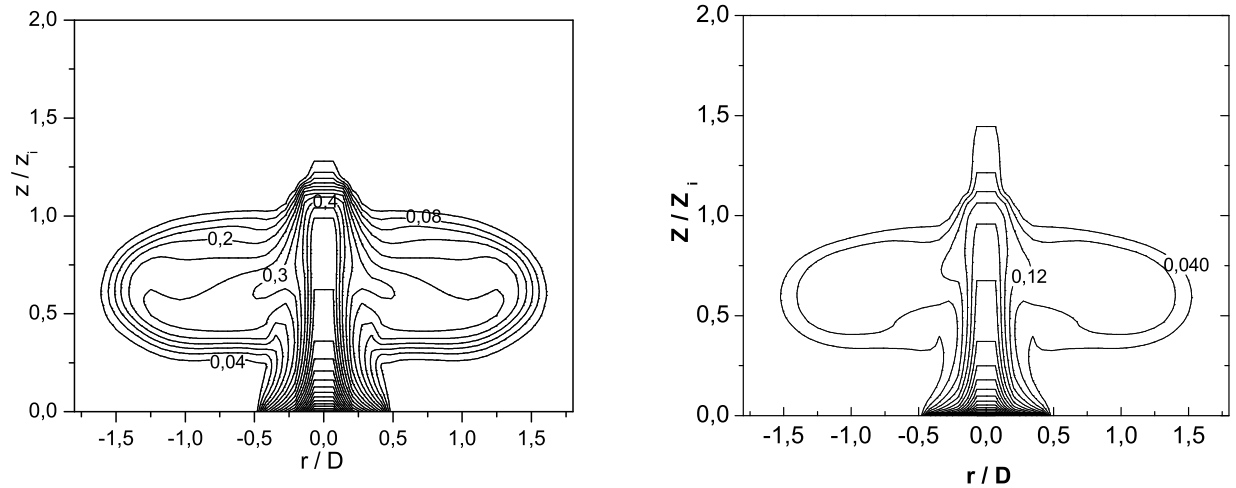
Турбулентная структура температурного поля представлена на рис. 3а распределением среднеквадратичных флуктуаций σ_T/T_D, где T_D — конвективный масштаб температурного поля. (Выражения для W_D и T_D см. в [3], [4]). Профиль имеет характерный вид с убыванием от максимального значения вблизи поверхности до минимального значения при $z/z_i \approx 0.85$, как в эксперименте, так и в RANS приближении с вычисляемыми из уравнений баланса параметрами E, ε, $< \theta^2 >$. Второй максимум вблизи слоя вовлечения обусловлен большим градиентом скорости генерации $< \theta^2 >$, подтверждается вертикальным распределением температуры на рис. 1*в*. Различие в профилях дисперсии в верхней части перемешанного слоя связано с известным дефектом использования упрощений градиентного типа при аппроксимации процессов турбулентной диффузии (третьих моментов термогидродинамических полей) в RANS приближении моделирования турбулентности с замыканием на уровне вторых моментов. На рис. 3*б* приведен вертикальный профиль диссипации кинетической энергии турбулентности, вычисленный по трехпараметрической модели, в сравнении с LES данными Moeng и Wyngaard [9] (взятыми из работы Canuto et.el. [8]). Максимум диссипации в области слоя вовлечения связан, по видимому, с высоким уровнем турбулентности вдоль "крышки" термического плума [4], [6], который в свою очередь может быть обусловлен дроблением вихрей, достигающих слоя вовлечения, и генерированием случайных сдвиговых слоев.

На рис. 4а приведено распределение концентрации, полученное в результате численного решения систе-



а б

Рис. 3: Вертикальный профиль дисперсии турбулентных флуктуаций температуры над островом тепла на оси факела: светлые квадраты — эксперимент [4], зачерненные квадраты — вычисления (*а*). Вертикальное распределение диссипации энергии турбулентности: светлые квадраты — эксперимент [9], зачерненные квадраты — вычисления (*б*).



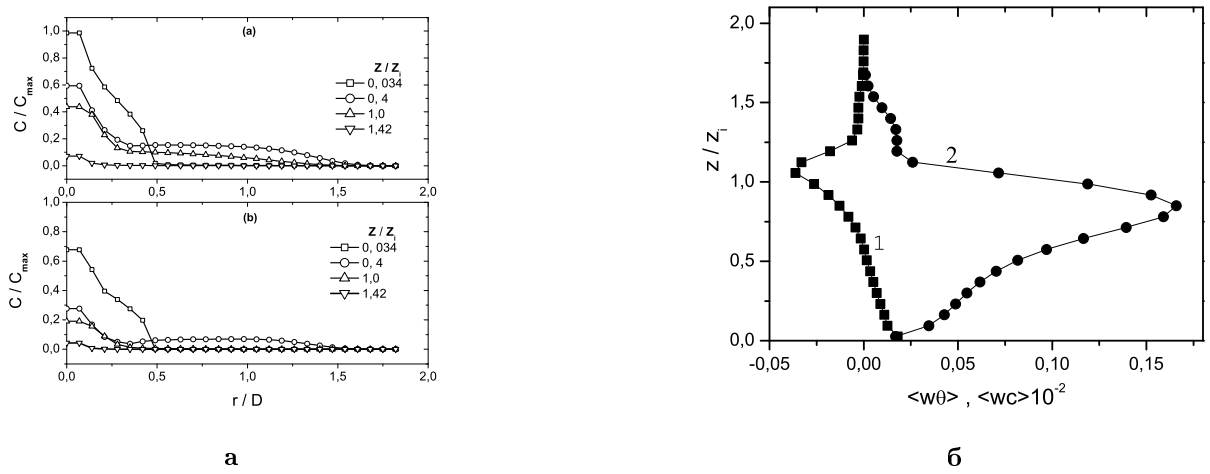
а б

Рис. 4: Изолинии концентрации пассивного скаляра (массы) в диффузионном факеле над поверхностным источником: (а) — анизотропная диффузионная модель, (*б*) — модель Буссинеска. Профили средней концентрации пассивного скаляра в различных сечениях по вертикальной координате над непрерывным поверхностным источником (*б*).

мы уравнений (1)–(5), т.е. эйлеровой диффузионной модели, а на рис. 4*б* показаны изолинии концентрации, вычисленные по модели вихревой диффузии Буссинеска: $- < wc >= \nu_t / Sc_t \cdot \partial C / \partial z$), в которой плавучесть не принимается во внимание. Рис. 4*а* и 4*б* показывают, что под влиянием проникающей турбулентной конвекции, порождаемой двумя вихрями, примесь переносится от поверхностного источника в перемешанный слой, распространяется далее в слой инверсии, претерпевает гравитационное рассеяние в горизонтальном направлении и проникает за пределы инверсионного слоя в область значений $z/z_i > 1$. Проникновение примеси за пределы инверсионного слоя получено в лабораторных экспериментах по распространению плавучих факелов пассивной примеси от непрерывного и мгновенного источников в конвективном пограничном слое [10].

Однако имеются различия в распределении средней концентрации. Модель Буссинеска без учета эффектов плавучести дает более медленную диффузию примеси в вертикальном (а, следовательно и горизонтальном) направлении по сравнению с моделью (5), учитывающей эффекты плавучести. При использовании модели (5) уровень концентрации оказывается выше как вблизи поверхности, так и по всей высоте перемешанного слоя. В области слоя инверсии (при $z/z_i \approx 1$) значение концентрации в центре острова тепла, вычисленное с помощью модели (5), примерно в два раза выше, чем рассчитанные по модели Буссинеска. На верхнем рис. 5*а* приведены результаты вычислений с использованием модели (1)–(5), на нижнем рис. 5*а* — модель Буссинеска.

На рис. 5*б* приведено поведение коэффициентов турбулентного обмена активного и пассивного скаляра. Можно видеть, что поток тепла меняет знак по высоте и достигает отрицательной максимальной величины в инверсионном слое (при $z/z_i = 1$). Поток массы по вертикали остается положительным.



а **б**

Рис. 5: Профили средней концентрации пассивного скаляра (массы вещества) в различных сечениях по вертикальной координате над непрерывным поверхностным источником (*а*). Вертикальные профили турбулентных потоков тепла $< w\theta >$ (кривая 1) и массы $< wc >$ (кривая 2) в центре острова: анизотропная диффузионная модель (*б*).

Заключение

Представленные результаты численного моделирования структуры циркуляции над островом тепла позволяют сделать следующие выводы. Описание турбулентности теплового факела в RANS приближении с включением в качестве искомых моментов гидротермодинамических полей второго порядка позволяет воспроизвести структурные особенности проникающей конвекции над источником тепла, включая такие тонкие эффекты, как перекрещивание вертикальных профилей температуры теплового факела с образованием области отрицательной плавучести, свидетельствующей о развитии куполообразной формы у верхней части факела. Вертикальная дисперсия турбулентной скорости и дисперсия турбулентных флуктуаций температуры удовлетворительно согласуется с опытными данными.

Результаты численного моделирования распространения пассивной примеси от непрерывного поверхностного источника показывают, что явные анизотропные выражения для турбулентных потоков тепла и массы с точным учетом эффектов плавучести дают более эффективную диффузию примеси от поверхностного источника в перемешанный слой и выше. Вычисленное отношение коэффициентов турбулентного обмена для пассивной примеси и активной примеси (тепло) в термически устойчивом стратифицированном течении над поверхностными источниками тепла и массы свидетельствуют об их различии.

Список литературы

[1] R. A. Pielke Mesoscale meteorological modeling // Academic Press, New York, Boston 2002. p. 676.

[2] Курбацкий А.Ф., Курбацкая Л.И. Трехпараметрическая модель турбулентности для атмосферного пограничного слоя над урбанизированной поверхностью // Изв. РАН. Физика атмосферы и океана. 2006. Т. 42, № 4, С. 476–494.

[3] Курбацкий А.Ф., Курбацкая Л.И. Турбулентная циркуляция над поверхностным источником тепла в устойчиво стратифицированной окружающей среде // Теплофизика и аэромеханика. 2016. Т.23, № 5, С. 703–719.

[4] Lu J., Arya P. et al. A Laboratory Study of the Urban Heat Island in a Calm and Stably Stratified Environment. Part I and II // J. Appl. Meteor. 1997. Vol.36, No.10, 1377–1402.

[5] A. F. Kurbatskii and L. I. Kurbatskaya Urban aeromechanics: Turbulent circulation and contaminant dispersion above city in stably stratified environment AIP Conference Proceedings 2027, 030024 2018. https://doi.org/10.1063/1.5065118

[6] Noto K. Dependence of heat island phenomena on stablestratification and heat quantity in a calm environment //Atmospheric Environment.1996.V.30.№3.P.475–485

[7] Warhaft Z. Heat and Moisture Flux in hte Stratified Boundary Layer // Q. J. Royal Meteor. Soc. 1976. V. 102. P. 703.

[8] Canuto, V.M., Minotti, F., Ronchi, C., Ypma, R.M.,1994: Second-Order Closure PBL Model with New Third-Order Moments: Comparison with LES Data. J.Atmos.Sci. 1605–1618.

[9] Moeng C.-H., Wyngaard J.C. Spectral analysis of large-eddy simulations of the convective boundary layer // J. Atmos. Sci. 1988. V.45. P.3574–3587.

[10] Snyder W.H., Lawson R.E.Je., Shipman M.S., Lu J. Fluid Modeling of Atmospheric in Convective Boundary Layer // Bound. Layer. Meteor. 2002. V. 102. 3. P.155.

Курбацкая Людмила Ивановна — ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: L.Kurbatskaya@ommgp.sscc.ru
Дата поступления — 27 мая 2019 г.

ВЕСОВЫЕ ОЦЕНКИ МЕТОДА МОНТЕ-КАРЛО ДЛЯ НЕИНВАЗИВНОГО ИЗМЕРЕНИЯ ОПТИЧЕСКИХ ПАРАМЕТРОВ БИОЛОГИЧЕСКИХ ТКАНЕЙ

А. В. Лаппа[1], А. Е. Анчугова[1], Д. Ю. Шакаева[1]

[1] *Челябинский государственный университет, 454001, Челябинск*

УДК 519.245
DOI: 10.24411/9999-016A-2019-10042

Предлагается новый метод для определения оптических параметров биологических тканей (коэффициента поглощения и транспортного коэффициента рассеяния), исходя из неинвазивных измерений диффузного отражения света. В методе используется кинетическая модель распространения света в веществе и новые весовые алгоритмы метода Монте-Карло. Алгоритмы, в частности, позволяют с высокой точностью, одновременно для большого набора сред с различными оптическими параметрами рассчитывать радиационные характеристики и их производные по этим параметрам. Этот результат получен благодаря использованию строгой теории метода Монте-Карло для интегральных уравнений. Метод реализован в виде программно-аппаратного комплекса и опробован на реальных биологических тканях. Проведенное сопоставление непосредственно измеренных полей излучения с рассчитанными по измеренным оптическим параметрам показало хорошее согласие в диапазоне значений параметров, характерных для кровенаполненных тканей.

Ключевые слова: метод Монте-Карло, весовые оценки, перенос излучения, кинетическая модель.

Введение

Настоящая работа посвящена использованию новых математических методов для решения центральной проблемы биомедицинской оптики — неинвазивного определения оптических параметров биологических тканей, исходя из измерений характеристик поля излучения. Здесь мы рассматриваем схему с непрерывным облучением и регистрацией диффузно отраженного излучения посредством многоволоконного оптического зонда. По одному из волокон зонда (излучающему) свет подводится от монохромного источника к исследуемому образцу, по остальным (считывающим) — свет отводится от образца к детекторам мощности. Это типичная обратная задача, решение которой включает разработку или выбор модели распространения излучения в веществе, множественное решение прямой задачи для заданного набора оптических параметров, проведение измерений и поиск оптических коэффициентов. В настоящей работе приводятся весовые оценки метода Монте-Карло, применимые как для данной конкретной задачи, так и для общего случая произвольных источников и сред.

1 Методы

1.1 Общие оценки в общей модели

Решение прямой задачи заключается в предварительной оценке показаний нескольких детекторов излучения для множества наборов оптических параметров, покрывающего с достаточно мелким шагом область возможных значений искомого набора. В работах [1], [4] предложено 2 семейства монте-карловских весовых оценок, позволяющих одновременно для большого набора сред с различными оптическими параметрами рассчитывать показания детекторов и даже их производные по этим параметрам. Оценки получены в общей кинетической модели и применимы для произвольных гетерогенных сред и источников. Они имеют высокую дифференциальную точность (необходимую в обратных задачах), т.е. малую погрешность расчета, как самих характеристик, так и их разностей при близких значениях параметров.

ISBN 978-5-901548-42-4

1.1.1 Уравнение и функционал

В кинетической модели световое поле в мутной среде от стационарных монохромных источников в пренебрежении неупругими взаимодействиями и поляризацией описываются скалярной лучевой интенсивностью $I = I(x)$ с фазовыми координатами $x = (\mathbf{r}, \mathbf{\Omega})$, описывающими положение точки в пространстве R^3, направление лучей на единичной сфере $O \subset R^3$. Параметрами модели являются показатель преломления n, коэффициенты поглощения μ_a и рассеяния μ_s, и индикатриса (фазовая функция) рассеяния $w(\mathbf{\Omega} \to \mathbf{\Omega}')$ (зависит от угла рассеяния $\arccos(\mathbf{\Omega}\mathbf{\Omega}')$).

Объект, вообще говоря, гетерогенный, т.е. занимаемая им область $\bar{G} \subset R^3$ состоит из одной или нескольких попарно непересекающихся открытых зон $G_1, G_2, ..., G_m$ и кусочно-гладкой границы Γ, состоящей из внутренней границы Γ_{in} (между зонами) и внешней границы Γ_{out}. Не ограничивая общности, будем полагать область $\bar{G}$ ограниченной, односвязной и обобщенно выпуклой. Последнее означает, что $\bar{G}$ невогнута и что любая прямая, проходящая через точку $\mathbf{r} \in G \equiv G_1 + ... + G_m$, имеет конечное число общих точек (пересечений и касаний) с границей Γ. Далее, без оговорок, термин граница и символы Γ, Γ_{in}, Γ_{out} применяются только в отношении гладких (открытых) частей соответствующих границ. Таким образом, ребра исключаются, и в любой точке границы $\mathbf{r} \in \Gamma$ определена нормаль $\mathbf{n_r}$; условимся считать, что на внутренней границе она направлена в сторону зоны с меньшим номером, а на внешней — из среды.

Предполагается, что детекторы линейны: их показания выражаются в виде линейных функционалов от интенсивности в облучаемом объекте. Следующий функционал является достаточно общим для задач биомедицинской оптики:

$$Q = (I, \mathrm{v}) + [I, b]_> \equiv \int_X I(x)\mathrm{v}(x)dx + \int_\Gamma d\Gamma \int_{O_{\mathbf{r}>}} d\mathbf{\Omega}\, |\mathbf{\Omega}\mathbf{n_r}|\, I(x)b(x) \tag{1}$$

Здесь и далее: $X \equiv \{G \times O\}$, $dx \equiv d\mathbf{r}d\mathbf{\Omega}$, $O_{\mathbf{r}<}(O_{\mathbf{r}>})$ — множество направлений $\mathbf{\Omega}$ таких, что $\mathbf{\Omega}\mathbf{n_r} \neq 0$ при $\mathbf{r} \in \Gamma_{in}$ и $\mathbf{\Omega}\mathbf{n_r} > 0\,(\mathbf{\Omega}\mathbf{n_r} < 0)$ при $\mathbf{r} \in \Gamma_{out}$. Принято, что интенсивность на границе есть интенсивность падающего на нее излучения. Функционал (1) задает показания детекторов двух типов: объемного и граничного. Показание первого формируются светом внутри среды, второго — светом, падающим на границу и покидающим ее; v, b — соответствующие функции отклика. Предполагается, что граничные детекторы, покидающего границу излучения, сосредоточены в фазовой области, свободной от поверхностных источников (детектирующая граница не излучает).

В сопряженной форме этот функционал имеет вид:

$$Q = (s, J) + [\sigma, J]_< \equiv \int_X s(x)J(x)dx + \int_\Gamma d\Gamma \int_{O_{\mathbf{r}<}} d\mathbf{\Omega}\, |\mathbf{\Omega}\mathbf{n_r}|\, \sigma(x)J(x), \tag{2}$$

где J — функция ценности, имеющая смысл показания детектора, создаваемого лучом света единичной мощности, испускаемым из фазовой точки x, s — функция (объемных) источников внутри объекта (фазовая плотность мощности), σ — интенсивность покидающего границу света, создаваемого поверхностными источниками (поверхностная плотность источников есть $\sigma(\mathbf{r}, \mathbf{\Omega})\,|\mathbf{\Omega}\mathbf{n_r}|$).

Прямое представление (1) и сопряженное представление (2) определяют одну величину, но условия их применимости различны. Для применимости прямого требуется регулярность источников, сопряженного — детекторов. Первое условие исключает, в частности, точечные, мононаправленные, и другие сингулярные источники, являющиеся естественными, широко используемыми моделями для множества реальных задач. Второе условие — гораздо менее ограничительно, оно отражает реальность и является естественным для алгоритмов метода Монте-Карло. Поэтому, далее мы развиваем сопряженный подход, допуская сингулярность источников, т.е. полагая функции σ и s в представлении (2), вообще говоря, обобщенными. Функция ценности и интегро-дифференциальное сопряженное уравнение, при этом, остаются классическими.

Интенсивность же при сингулярных источниках является обобщенной функцией и уравнение переноса необходимо рассматривать в обобщенном смысле. Разрывы коэффициентов приводят к затруднениям такого рассмотрения, для преодоления которых нужно обобщать обобщенный подход в духе теории Коломбо [6] и других, что громоздко и не сделано для уравнения переноса. Более адекватный подход требует введения мер для описания поля излучения вместо интенсивности. Не все физики готовы к такому переходу, в частности из-за потери уравнения переноса в интегро-дифференциальной форме. Желание избежать всех этих усложнений, явилось важной причиной выбора сопряженного подхода.

В работе [1] мы получили интегральное уравнение для функции ценности, корректно учитывающие внутренние и внешние граничные условия:

$$J = (\mathbf{LS} + \mathbf{\Theta P}_\chi)J + (\mathbf{L}\mathrm{v} + \mathbf{\Theta}b), \tag{3}$$

где сопряженные операторы пробега, рассеяния, достижения границы и отражения-преломления определяются, соответственно, выражениями:

$$\mathbf{L}f(x)=\int_0^{l_x} e^{-\tau(x,l)}f(\mathbf{r}+\mathbf{\Omega}l,\mathbf{\Omega})dl, \mathbf{S}f(x)=\mu_s(\mathbf{r})\int_O f(\mathbf{r},\mathbf{\Omega}')w(\mathbf{r},\mathbf{\Omega},\to\mathbf{\Omega}')d\mathbf{\Omega}', \mathbf{\Theta}f(x)=e^{-\tau(x,l_x)}f(\mathbf{r}_\mathbf{\Omega},\mathbf{\Omega}),$$

$$\mathbf{P}_\chi=\chi_{in}(\mathbf{P}_<+\mathbf{P}_>)+\chi_{out}\mathbf{P}_<,\ \mathbf{P}_{</>}g(x)=\int_{(\mathbf{n_r\Omega}')(\mathbf{n_r\Omega})^{</>}0}|\mathbf{n_r\Omega}'|\,g(\mathbf{r},\mathbf{\Omega}')p_b(\mathbf{r},\mathbf{\Omega}\to\mathbf{\Omega}')d\mathbf{\Omega}'.$$

Здесь $\tau(x,l)=\int_0^l\mu(\mathbf{r}+\mathbf{\Omega}l')dl'$ — оптическое расстояние между точками $\mathbf{r}$ и $\mathbf{r}+\mathbf{\Omega}l$; $\mu=\mu_a+\mu_s$ — коэффициент взаимодействия, l_x — расстояние вдоль луча, выходящего из точки $\mathbf{r}$ в направлении $\mathbf{\Omega}$ до первого достижения границы, $\mathbf{r}_\mathbf{\Omega}=\mathbf{r}+\mathbf{\Omega}l_x$ — точка этого достижения, $p_b(\mathbf{r},\mathbf{\Omega}\to\mathbf{\Omega}')$ — дифференциальный коэффициент отражения-преломления (при отрицательных $(\mathbf{n_r\Omega})(\mathbf{n_r\Omega}')$ — отражение, при положительных — преломление), $\chi_{in}(r),\chi_{out}(\mathbf{r})$ — индикаторы внутренней и внешней границ.

Это уравнение определено и справедливо внутри объекта, т.е. в X, и на границах для «пересекающих» направлений: входящих на внешней границе и всех на внутренних, т.е. на $X_<=\{x:\mathbf{r}\in\Gamma,\Omega\in O_{\mathbf{r}<}\}$. Объединение $\bar{X}=X\bigcup X_<$ совпадает с областью интегрирования функции ценности в (2), так что $\bar{X}$ является естественным фазовым пространством для уравнения (3) и задачи в целом. Естественным пространством для решений уравнения (3) является банахово пространство ограниченных (строго) в $\bar{X}$ функций. Уравнение имеет в нем единственное итерационное решение, если функции отклика v,b и введенные выше операторы ограничены, а среда всюду имеет ненулевое поглощение: $\inf\limits_{\mathbf{r}\in G}|\mu_a(\mathbf{r})|>0$. Для существования функционала (2) достаточно интегрируемости распределений источников s,σ (конечности соответствующих мер). Все эти требования естественны, и мы их далее будем считать выполненными без оговорок.

1.1.2 Простые оценки сигналов

Оператору уравнения (3) и заданным источникам отвечает аналоговая («физическая») марковская цепь в $\bar{X}$. Цепь эта однородная и обрывающаяся: $n<\infty$. Реализацию этой цепи $x_0,...,x_n$ естественно интерпретировать как последовательность фазовых координат событий на траектории частицы: старт, столкновения двух типов (рассеяние и поглощение) и достижение границы двух типов (с вылетом из среды и без такового). Последним, n-м событием реализации является событие, предшествующее поглощению или вылету. Дополним реализацию этими событиями, приписав им номер $n+1$ и направления до поглощения и после вылета.

Физическая цепь полностью задается геометрией объекта, функциями источников и оптическими параметрами, т.е. набором *исходных* характеристик: $s,\sigma,\mu_a,\mu_s,w,p_b$. Введем другую цепь, *модифицированную,* отвечающую той же геометрии объекта, но другим источникам и оптическим параметрам, т.е. определенную в том же фазовом пространстве $\bar{X}$, но обусловленную набором *модифицированных* характеристик: $\tilde{s},\tilde{\sigma},\tilde{\mu}_a,\tilde{\mu}_s,\tilde{w},\tilde{p}_b$. Потребуем, чтобы выполнялись условия *регулярности*: каждая модифицированная характеристика переходит в исходную посредством умножения на некоторую классическую функцию: $\mu_s(\mathbf{r})=\underset{\sim}{\mu}_s(\mathbf{r})\tilde{\mu}_s(\mathbf{r})$ и т.д. Эти функции будем обозначать символами характеристик, подчеркнутых волной.

Определим на реализациях модифицированной цепи 6 функционалов:

$$\phi_{\alpha,\beta}=\sum_{i=0}^{n}q_i(\tilde{\mathrm{v}}_i^\alpha+\tilde{b}_i^\beta),\ \alpha=e,t,c,\ \beta=e,a, \tag{4}$$

где вес $q_i=e^{\Delta\tau_i}\prod_{m=0}^i u_m$, $u_m=\chi_\Gamma(\mathbf{r})\underset{\sim}{p}(\mathbf{r}_m,\mathbf{\Omega}_{m-1}\to\mathbf{\Omega}_m)+\chi_G(\mathbf{r})\underset{\sim}{\mu}_s(\mathbf{r}_m)\underset{\sim}{w}(\mathbf{r}_m,\mathbf{\Omega}_{m-1}\to\mathbf{\Omega}_m),m>0;$, $\Delta\tau_i=\sum_{j=0}^{i-1}\int_0^{l_j}\Delta\mu(\mathbf{r}_j+\mathbf{\Omega}l)dl$ — разность оптических путей от старта до i-го события, вычисленных с коэффициентами $\tilde{\mu}$ и $\mu(\Delta\mu=\tilde{\mu}-\mu)$, $l_i=|\mathbf{r}_{i+1}-\mathbf{r}_i|$, χ_Γ и χ_G — индикаторы Γ и G, $\Delta\tau\,(x_i,l)=\int_0^l\Delta\mu(r_i+\Omega_i l')dl'$

$$\begin{gathered}\tilde{\mathrm{v}}_i^e=\mathbf{L}\mathrm{v}(x_i),\ \tilde{\mathrm{v}}_i^t=\int_0^{l_i}e^{\Delta\tau(x_i,l)}\mathrm{v}(\mathbf{r}_i+\mathbf{\Omega}_i l,\mathbf{\Omega}_i)dl,\ \tilde{\mathrm{v}}_i^c=\chi_G(\mathbf{r}_{i+1})e_i\mathrm{v}(\mathbf{r}_{i+1},\mathbf{\Omega}_i)/\tilde{\mu}(\mathbf{r}_{i+1}),\\ \tilde{b}_i^e=\mathbf{\Theta}b(x_i),\ \tilde{b}_i^a=\chi_\Gamma(\mathbf{r}_{i+1})e_ib(\mathbf{r}_{i+1},\mathbf{\Omega}_i),e_i\equiv e^{\Delta\tau(x_i,l_i)}.\end{gathered} \tag{5}$$

Каждый из этих функционалов является несмещенной оценкой функции ценности, объемной и граничной частей искомого функционала (2) : $J(x)=\tilde{\mathbf{M}}_x\phi_{\alpha,\beta}$, $(s,J)=\tilde{\mathbf{M}}^v\phi_{\alpha,\beta}$, $[\sigma,J]_<=\tilde{\mathbf{M}}^b\phi_{\alpha,\beta}$ на траекториях,

начинающейся, соответственно, из неслучайной точки x и из случайных точек в X и $X_<$, распределенных с модифицированными нормированными плотностями $\tilde{s}(x)$ и $\tilde{\sigma}(x)\,|\mathbf{\Omega n_r}|$ ($(\tilde{s},1) = [\tilde{\sigma},1]_< = 1$). Начальные веса u_0 оценок равны 1, $\underset{\sim}{s}(x_0)$, $\underset{\sim}{\sigma}(x_0)$, соответственно. Символы $\tilde{\mathbf{M}}$ с индексами обозначают математические ожидания на модифицированных траекториях с описанными стартами.

Оценки $\phi_{\alpha,\beta}$ отличаются степенью аналитического осреднения «физических» вкладов, она возрастает в порядке $\alpha = c,t,e;\ \beta = a,e$. При этом дисперсия оценок уменьшается, но увеличивается их сложность. Поэтому, вообще говоря, нельзя заранее выбрать лучшую из них. Но, в важном частном случае, когда фазовый объем детектора (области в $\bar{X}$, где отличаются от нуля функции отклика) мал, и вероятность попадания в него траекторией мала, следует ожидать увеличения эффективности с увеличением степени аналитического осреднения: $\phi_{e,e}$ тогда, вероятно, будет лучшей в семействе (4). При $\mathrm{v} = 0$ (имеются только граничные детекторы) семейство состоит из двух оценок: «по достижениям» границ ϕ_a ($\beta = a$) и «по событиям» ϕ_c ($\beta = c$). Количество вкладов в последней гораздо больше (из-за аналитического осреднения), что и обуславливает ее большую эффективность.

Еще больше можно увеличить эффективность оценок, проведя дополнительное аналитическое осреднение вкладов в духе локальных оценок [7, 8, 9].

1.1.3 “Локальные” оценки сигналов

Идея перехода к таким оценкам заключается в применении весовой оценки (4) к вычислению итерированной ценности $J_1 = (\mathbf{LS} + \mathbf{\Theta P}_\chi)J$, удовлетворяющей уравнению (3) с другим свободным членом: $(\mathbf{LS} + \mathbf{\Theta P}_\chi)(\mathbf{L}\mathrm{v}(x)+\mathbf{\Theta}b(x))$. Его носитель шире носителя исходного свободного члена: $\mathbf{L}\mathrm{v}(x)+\mathbf{\Theta}b(x)$, и потому оценка J_1 с вкладами $(\mathbf{LS}+\mathbf{\Theta P}_\chi)(\mathbf{L}\mathrm{v}+\mathbf{\Theta}b)$ будет иметь большую вероятность ненулевых реализаций, чем оценка J с вкладами $\mathbf{L}\mathrm{v}+\mathbf{\Theta}b$. В результате новый алгоритм может быть существенно эффективнее старого. Вместо (4) получаем семейство функционалов на траекториях модифицированной цепи:

$$\xi_{\alpha,\beta} = \sum_{i=0}^{n} q_i(\tilde{\mathrm{v}}_{1i}^{\alpha} + \tilde{b}_{1i}^{\beta}),\ \alpha = e,t,c,\ \beta = e,a, \tag{6}$$

являющихся несмещенными оценками итерированных функции ценности и искомых функционалов: $J_1(x) = \tilde{\mathbf{M}}_x\xi_{\alpha,\beta}$, $(s,J_1) = \tilde{\mathbf{M}}^{\mathrm{v}}\xi_{\alpha,\beta}$, $[\sigma,J_1]_< = \tilde{\mathbf{M}}^{b}\xi_{\alpha,\beta}$ при тех же стартах и весах, что и в (4). Вклады $\tilde{\mathrm{v}}_{1i}^{\alpha}$, $\tilde{b}_{1i}^{\beta}$ рассчитываются по формулам (5) с заменой v, b на $\mathrm{v}_1 \equiv \mathbf{SL}\mathrm{v} + \mathbf{S\Theta}b, b_1 \equiv \mathbf{P}_\chi\mathbf{L}\mathrm{v}+\mathbf{P}_\chi\mathbf{\Theta}b$. Переход к искомым (не итерированным) ценностям и функционалам осуществляется с помощью связи: $J = J_1 + \mathbf{L}\mathrm{v}+\mathbf{\Theta}b$. Заметим, что при граничных детекторах ($\mathrm{v} = 0$) даже простейшая оценка из (6) $\xi_{c,a}$ имеет ненулевые вклады от событий, причем, гораздо чаще, чем любая из (4).

1.1.4 Оценки производных

Пусть η — некоторый параметр и существуют и ограничены производные по нему от всех заданных функций, определяющих систему: источники s,σ, среду μ,μ_s,w,p_b и детекторы v,b. Тогда, при выполнении принятых ранее допущений, обеспечивающих существование ценности и искомой характеристики, также существуют и их производные по этому параметру: J', Q', причем операцию дифференцирования можно ввести под интегралы в (2).

Выберем неаналоговую модифицированную цепь так, чтобы ее характеристики $\tilde{s},\tilde{\sigma},\tilde{\mu},\tilde{\mu}_s,\tilde{w},\tilde{p}_b$, не зависели от η. Тогда операции дифференцирования и осреднения перестановочны и возможно почленное дифференцирование оценок. Получается, что функционалы от модифицированной цепи

$$\phi'_{\alpha,\beta} = \sum_{i=0}^{n}[q_i(\tilde{\mathrm{v}}_i^{\alpha} + \tilde{b}_i^{\beta})' + q_i'(\tilde{\mathrm{v}}_i^{\alpha} + \tilde{b}_i^{\beta})],\quad \xi'_{\alpha,\beta} = \sum_{i=0}^{n}[q_i(\tilde{\mathrm{v}}_{1i}^{\alpha} + \tilde{b}_{1i}^{\beta})' + q_i'(\tilde{\mathrm{v}}_{1i}^{\alpha} + \tilde{b}_{1i}^{\beta})], \tag{7}$$

являются несмещенными оценками производных от функций ценности J и J_1, соответственно: $J'(x) = \tilde{\mathbf{M}}_x\phi'_{\alpha,\beta}$,$J_1'(x) = \tilde{\mathbf{M}}_x\xi'_{\alpha,\beta}$.

Оценки для Q' получаются дифференцированием выражения (2) и применением для J' оценок $\phi'_{\alpha,\beta}$ или $\xi'_{\alpha,\beta} + (\mathbf{L}\mathrm{v}(x_0)+\mathbf{\Theta}b(x_0))'$. Если функции источников s,σ не зависят от параметра η, то получаются оценки вида (4), (6), в которых вместо величин ϕ,ξ,v,b стоят их производные$\phi',\xi',\mathrm{v}',b'$.

1.2 Оптические параметры

1.2.1 Модель и оптические параметры

Оптические параметры биологической ткани — это параметры модели, принятой для описания распространения света в ткани. В кинетической модели это μ_a, μ_s, w, n. Ясно, что объем информации, содержащейся в измеряемых характеристиках, недостаточен для определения всех этих параметров, и потому необходимо перейти к редуцированной модели с меньшим количеством параметров. Редуцированная кинетическая модель (сформулированная в [3], фактически используемая и в других работах) получается из общей простым фиксированием индикатрисы, т.е. априорным назначением ее (индикатрисы) формы и параметров. Тогда в модели остаются 3 варьируемых кинетических параметра: n, μ_a, μ_s, однозначно связанных с известным 3-х параметрическим набором, используемым для описания полей излучения: n, μ_a, μ_s', где $\mu_s' = \mu_s(1-g)$ — редуцированный (транспортный) коэффициент рассеяния, g — средний косинус угла рассеяния (фактор анизотропии). Имея в виду биомедицинские применения, мы выбрали в качестве индикатрисы функцию Хейни-Гринштейна [5] с априорно назначаемым параметром анизотропии вблизи реальных его значений. Показатель преломления может быть определен из независимых экспериментов и полагается известным. Определение двух остальных и является задачей метода.

1.2.2 Схема

Вместо пары μ_a, μ_s' удобнее определять пару $\mu_s, p \equiv \mu_a/\mu$, однозначно связанную с искомой: $\mu_a = \mu_s p/(1-p)$, $\mu_s' = \mu_s(1-g)$. Они определяются в диффузно-отражательной схеме (рис. 1): Облучение и регистрация от-

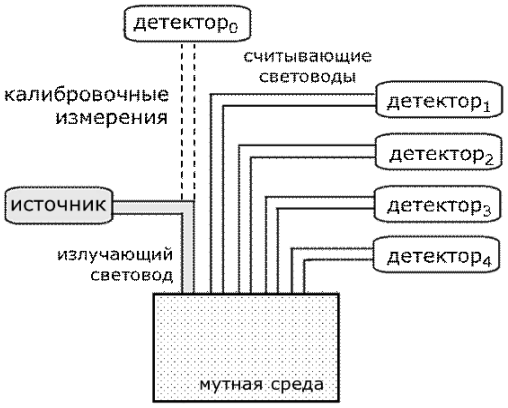


Рис. 1: Схема измерений оптических параметров

раженного излучения осуществляется с помощью специального многоволоконного зонда. Он состоит из 1 излучающего и n (не менее двух) считывающих световодов, расположенных на разных расстояниях от излучающего. По излучающему световоду подводится свет от монохромного источника к исследуемому образцу среды, по считывающим — свет отводится от образца к детекторам мощности, в качестве которых мы используем фотодиоды. Сначала измеряется сигнал d_0 от излучающего световода, затем зонд прикладывается к образцу, и измеряются сигналы $d_1, ..., d_n$ от считывающих световодов. Оптические параметры μ_s, p извлекаются из n не зависящих от мощности источника отношений $Q_i = d_i/d_0,\ i = 1, ..., n$, с использованием таблиц теоретических оценок этих относительных сигналов и их производных по этим параметрам для заданного достаточно широкого набора пар μ_s, p в описанной выше модели.

1.2.3 Оценки сигналов и производных

Оценка сигналов производилась в следующих, достаточно реалистических допущениях:
Теоретическая оценка относительных сигналов производилась в следующих, достаточно реалистических допущениях:

A1. Среда однородна и полубесконечна.

A2. Отражающие свойства зонда постоянны по всему торцу. Он достаточно велик в поперечнике, чтобы можно было считать его бесконечным.

A3. Излучающий световод испускает свет равномерно по торцу его сердцевины и азимутальному углу. Угловое распределение мощности $S_a(\vartheta)$, измеренное в воздухе (проинтегрированное по торцу), нормированное на 1, определяет источник для оценки Q_i.

A4. Мощность света, выходящего из детектирующего световода, не зависит от азимутального угла и точки падения луча на торец его сердцевины. Зависимость мощности $B_{ai}(\vartheta)$от полярного угла падения луча единичной мощности, измеренная в воздухе, определяет соответствующий детектор.

A5. Чувствительность всех детекторов одинакова и ее угловой зависимостью можно пренебречь в пределах угловой апертуры световодов.

A6. Все световоды цилиндричны с одинаковым радиусом сердцевины r_0.

В этих предположениях относительный сигнал Q_i для заданной пары μ_s, p может быть представлен в виде второго (граничного) члена функционала (1), $[I, b_i]_>$, в котором граница $\Gamma = \Gamma_{out}$ — плоская поверхность раздела среды и зонда, $I(\mathbf{r}, \mathbf{\Omega}) = I(\rho, z, \mathbf{\Omega})$ — интенсивность, рассчитанная с единичным коэффициентом рассеяния и заданным p от точечного источника, сосредоточенного в начале координат (на оси пучка на границе), имеющего угловое распределение, отвечающее «воздушному» распределению $S_a(\vartheta)$. Зависимость от μ_s и радиуса физического источника r_0 перенесена на функцию отклика:

$$b_i(\rho, \mathbf{\Omega}) = B_i(\theta)\varphi_{l_i/r_0}(\rho/(\mu_s r_0)), \quad \rho = |\rho|\,, \quad \cos\theta = \mathbf{n_r}\mathbf{\Omega}, \quad B_i(\theta) = 2\pi^{-2}[T_{mf}(\theta)/T_{af}(\theta^{ma})]B_{ai}(\theta^{ma}),$$

$$\varphi_\lambda(t) \equiv \left\{ \int_{\max(t-1,\lambda-1)}^{\min(t+1,\lambda+1)} \arccos\left(\frac{t^2-1+t'^2}{2t't}\right) \arccos\left(\frac{\lambda^2-1+t'^2}{2t'\lambda}\right) t'dt',\ |t-\lambda| \le 2;\ 0,\ |t-\lambda| > 2 \right\},$$

l_i — межцентровое расстояние между излучающим и i-м считывающим световодами, a, m, f — символы сред: воздуха, зондируемой среды и сердцевины волокна (кварца), соответственно; $\theta^{\alpha\beta}$ — угол преломления при угле падения θ в переходе из среды α в среду β; $T_{\alpha\beta}(\theta)$ — мощность преломленного луча в таком переходе при единичной мощности падающего луча (дифференциальный коэффициент преломления).

Оценку такого функционала будем строить на модифицированной марковской цепи, отвечающей среде с прежней (физической) индикатрисой рассеяния, единичным коэффициентом рассеяния и назначенной вероятностью поглощения $\tilde{p}$. Частица стартует из начала координат, ее направление моделируется в 2 этапа. Сначала выбирается полярный угол ϑ_* из «воздушного» распределения S_a, затем он пересчитывается как $\vartheta_0 = \vartheta_*^{am}$, т. е. выполняется преломление из воздуха в среду. Азимутальный угол старта полагается нулевым.

Веса в оценках (4), (6) при такой модификации достаточно просты:
$q_k = (T_{fm}(\vartheta_0^{mf})/T_{fa}(\vartheta_0^{mf}))\exp\{[\tilde{\nu}-\nu]L_k\}$, где $\nu = (1-p)^{-1}$, $\tilde{\nu} = (1-\tilde{p})^{-1}$, L_k — путь, пройденный частицей до k-го события. Сами оценки (4), (6) в простейших вариантах принимают вид:

$$\phi_a = \sum_a q_a b_i(\rho_a, \theta_a), \quad \xi_{c,a} = \sum_k q_k \tilde{\nu}^{-1} \int_{\mathbf{\Omega}\mathbf{n_r}>0} d\mathbf{\Omega} w(\mathbf{\Omega}'_k \to \mathbf{\Omega}) \exp\{-l_k\nu\} B_i(\theta)\varphi_{l_i/r_0}(|\mathbf{r}_k + \mathbf{\Omega}_k l_k|/(\mu_s r_0)) \qquad (8)$$

где $l_k \equiv z_k/\mathbf{n_r}\mathbf{\Omega}$. Поскольку в рассматриваемых условиях полная и итерированная функции ценности стартующих частиц совпадают, $J_1 = J$, обе оценки являются несмещенными для Q_i.

Первая оценка строится по достижениям границы с углом падения внутри апертуры детектора $A \equiv \max\{\theta : B_{ai}(\theta) > 0\}$. Вторая оценка формируется из гораздо более частых событий — столкновений, дающих больше ненулевых вкладов, так как для их расчёта не требуется попадание в угловую апертуру. Дисперсия этой оценки намного ниже, но ее сложность и время вычисления больше. Можно упростить оценку, рандомизируя интеграл, например, вычисляя его методом Монте-Карло по одному случайному направлению внутри апертуры $(0, A) \times (0, 2\pi)$.

Дифференцируя почленно оценки по параметрам, получаем оценки для производных $Q'_{i\mu_s}$ и Q'_{ip}. Из $\xi_{c,a}$получаются оценки производных по столкновениям:

$$\xi'_{\mu_s} = \sum_k q_k(\tilde{\nu}\mu_s)^{-1} \int_{\mathbf{\Omega}\mathbf{n_r}>0} d\mathbf{\Omega} w(\mathbf{\Omega}'_k \to \mathbf{\Omega}) \exp\{-l_k\nu\} B_i(\theta)\psi_{l_i/r_0}(|\mathbf{r_k} + \mathbf{\Omega}_k l_k|/(\mu_s r_0)) \qquad (9)$$

$$\xi'_p = -\sum_k q_k \nu^2 \tilde{\nu}^{-1} \int_{\mathbf{\Omega n_r}>0} d\mathbf{\Omega} w(\mathbf{\Omega}'_k \to \mathbf{\Omega})(L_k + l_k) \exp\{-l_k \nu\} B_i(\theta) \varphi_{l_i/r_0}(|\mathbf{r_k} + \mathbf{\Omega}_k l_k| / (\mu_s r_0), \theta) \qquad (10)$$

$$\psi_\lambda(t) = \left\{ \int_{\max(t-1,\lambda-1)}^{\min(t+1,\lambda+1)} \frac{t'(t^2+1-t'^2)}{((2tt')^2 - (t^2-1+t'^2)^2)^{1/2}} \arccos\left(\frac{t'^2+\lambda^2-1}{2t'\lambda}\right) dt',\ |t-\lambda| \le 2;\ 0,\ |t-\lambda| > 2 \right\}.$$

Здесь также для упрощения можно применить рандомизацию интегралов.

Поскольку модифицированная цепь не связана с параметрами среды и геометрией зонда, оценку всех сигналов Q_i и их производных для множества заданных наборов параметров можно выполнять в одном расчете на одном ансамбле траекторий. Функции $\varphi_\lambda(t)$, $\psi_\lambda(t)$ имеет смысл предварительно затабулировать.

Для определения оптических параметров сигналы детекторов вычисляются по формулам (8) одновременно для большого набора значений заданных оптических параметров. Полученные оценки сравниваются с измеренными значениями сигналов с помощью взвешенного метода наименьших квадратов, где веса компенсируют отличия порядков функций

2 Результаты

На основе «локальных» оценок по столкновениям (8) был разработан метод получения оптических параметров биологических тканей. Определение пары происходит в два этапа: предварительного и рабочего. Первый этап включает

1. измерение угловых распределений интенсивности источника $S_a(\vartheta)$ и чувствительности детектора в воздухе $B_{ai}(\theta^{ma})$, характеристик зонда a, l;

2. расчет и табулирование показаний детекторов Q_i ($i = \overline{1,4}$ — номер детектора) и их производных $Q_i^{\mu_s}, Q_i^p$ для сред с различными заданными оптическими параметрами;

На рабочем этапе выполняется:

1. освещение исследуемой среды низкоинтенсивным монохроматическим излучением и измерение сигналов d_0, d_i;

2. решение обратной задачи по поиску оптических параметров исследуемой среды μ_s, p и пересчет на искомые μ_a, μ'_s.

Эта схема была реализована в программно-аппаратном комплексе. Его программная часть служит для управления лабораторным оборудованием, ввода исходных данных, проведения всех необходимых расчетов. На рис. 2 представлены результаты расчётов сигналов детекторов Q_i ($i = 1, \ldots, 4$ — номер детектора) и их производных по оптическим параметрам $Q_i^{\mu_s}, Q_i^p$ для 17×21 различных пар параметров μ_s,p, заданных в диапазоне $[1, 500] \times [0.001, 0.1]$. Все 4284 расчетов были проведены на одном ансамбле из 107 траекторий модифицированной марковской цепи менее чем за несколько минут, выбрана оценка по столкновениям. Аппаратная часть комплекса включает в себя используемые для проведения эксперимента приборы и инструменты, где основные из них это оптоволоконный зонд (для измерения диффузно отраженного света), источник монохроматического света, детектор излучения, моторизированная поворотная платформа (для снятия угловых распределений).

Оптические параметры не являются строгими константами среды (зависят от принятой модели), и проверка метода состоит в оценке точности прогнозирования полей излучения по определенным параметрам. Мы провели такие испытания в эксперименте по прохождению широкого пучка света.

Излучение с помощью световода направляли на смесь чернил, интралипида и воды и регистрировали с помощью оптического зонда, погруженного в смесь и расположенного на заданном расстоянии от торца световода. Ось пучка совпадала с осью излучающего световода зонда. Мы подготовили две смеси с разной толщиной слоя жидкости (3.85 и 4 мм) и измерили их параметры методом диффузного отражения с использованием 2, 3 и 4 детекторов. На рис. 3 измеренные относительные сигналы в этих смесях сравниваются с расчетными. Расчеты проводились с использованием той же программы Монте-Карло, которая применялась для подготовительных расчетов. Мы оцениваем результаты тестирования как хорошие. Ошибка предсказания полей внутри ткани на глубинах до 4 транспортных длин не превосходит 10%, а на глубинах до 2 транспортных длин — не превосходит 5%.

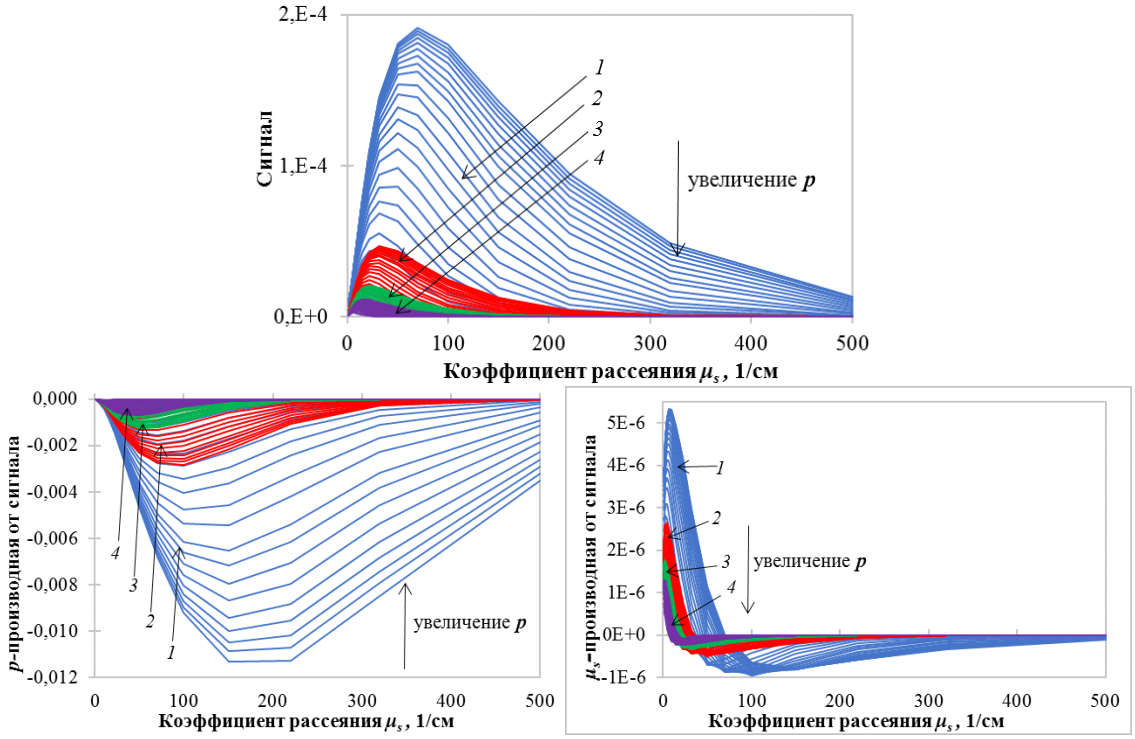


Рис. 2: Сигналы четырех детекторов и их производных по оптическим параметрам полученные одновременно в одном расчете. Разные цвета линий соответствуют разным расстояниям между центрами излучающего и считывающих волокон. Значения расстояний указаны у кривых в мм.

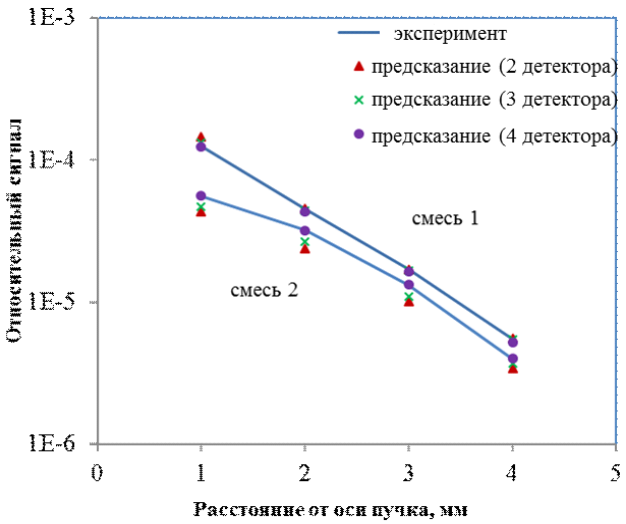


Рис. 3: Измеренные сигналы от четырех детекторов, расположенных на разных расстояниях от оси пучка, в эксперименте на прохождение и их предсказания, рассчитанные из измерений по диффузному отражению (по данным 2, 3 и 4 детекторов) для двух смесей

В качестве примера реального применения метода мы измерили оптические параметры цельной неразбавленной венозной крови человека. Полученные результаты согласуются с данными других авторов (Таблица 1). В ней использованы следующие обозначения: Htc — гематокрит — отношение объема эритроцитов к объему жидкой части крови, SatO2 — сатурация — уровень насыщения кислородом гемоглобина крови, выраженный в процентах, γ — скорость кровотока.

Таблица 1: Оптические параметры человеческой крови

Статья	**Образец**	λ, **нм**	g	μ_s', **см**$^{-1}$	μ_a, **см**$^{-1}$
Наша работа	Цельная кровь (Htc=0.52), SatO2 = 98%, без течения	657	0.98	11.3	12.7
Yaroslavsky и др. [9]	Цельная кровь (Htc=0.45-0.46), разбавленная до Htc=0.001 с помощью фосфатного буферного раствора, SatO2 > 98%, без течения	633	0.982	11.6	15.5
Meinke и др. [10] (данные из графиков)	4.8×106 RBCs/mL в плазме (Htc=0.4), SatO2 > 98%, γ=600 c-1	650	0.988	12	1
Roggan и др. [11] (данные из графиков)	Эритроцитарный концентрат, разбавленный до Htc=0.5 с помощью фосфатного буферного раствора, SatO2 > 98%, γ=500 c-1	633	0.978	20	8

Заключение

Представленный метод позволяет проводить неинвазивное определение коэффициента поглощения и транспортного (редуцированного) коэффициента рассеяния в мутных средах. Измеренные коэффициенты обеспечивают высокую точность прогнозирования полей отраженного, проникшего и прошедшего излучения для полубесконечной однородной среды. Метод подтвердил свою эффективность в экспериментах по определению оптических параметров некоторых биологических тканях. Высокие результаты были достигнуты благодаря использованию нового неаналогового алгоритма метода Монте-Карло весового типа. Одной из главных его особенностей является возможность расчета производных сигналов детекторов. Именно она сыграла ключевую роль в высокой точности измерений.

Список литературы

[1] Lappa A.V. and Anchugova A.E., Monte Carlo weighted algorithms for calculation of radiation characteristics and their derivatives in the biomedical optics problems// Proc. SPIE 10876, 1087626, 2019.

[2] Lappa A.V., Anchugova A.E. and Shakaeva D.Yu. Extraction of tissue optical parameters from diffuse reflectance measurements with a new able to count derivatives inverse Monte Carlo method// Proc. SPIE 10876, 108768, 2019.

[3] Lappa A.V., Kulikovskiy Ar.N., Kulikovskiy An.N., Makarova T.A. A non-destructive diffuse reflectance calibration-free method for determine optical parameters of biological tissues," Proc. SPIE 8579, 857912, 2013.

[4] Лаппа А.В., Анчугова А.Е. Весовые оценки метода Монте-Карло для решения прямых и обратных задач оптики мутных сред// Труды Международной конференции "Марчуковские научные чтения-2017 с. 521–527.

[5] Henyey L. G., and Greenstein J. L. Diffuse radiation in the galaxy// Astrophys. J.„V. 93, p. 70–83, 1941.

[6] Colombeau J. F. New Generalized Functions and Multiplication of the Distributions// North Holland, Amsterdam, 1984.

[7] Михайлов Г. А., Войтишек А. В. Численное статистическое моделирование. Методы Монте-Карло// Изд.: Академия, 2006.

[8] Kalos M.H. On the estimation of the flux at a point by Monte Carlo// IEEE Transactions on Nuclear Science,V. 16, iss. 3, 1963.

[9] Yaroslavsky A.N., Yaroslavsky I.V., Goldbach T. and Schwarzmaier, H.J. The optical properties of blood in the near infrared spectral range// Proc SPIE 2678, 314–324, 1996.

[10] Meinke M. C., Miller G. J., Helfmann J. and Friebel M. Optical properties of platelets and blood plasma and their influence on the optical behavior of whole blood in the visible to near infrared wavelength range// J. Biomed. Opt., V. 12, iss. 1, 014024, 2007.

[11] Roggan A., Friebel M., Doerschel K., Hahn A. and Mueller G. J. Optical properties of circulating human blood in the wavelength range 400-2500 nm// J. Biomed. Opt., V. 4, iss. 1, 1999.

Лаппа Александр Владимирович — д.ф.-м.н., проф. кафедры теоретической физики Челябинского государственного университета;
e-mail: lappa@csu.ru;
Анчугова Анастасия Евгеньевна — аспирант кафедры теоретической физики Челябинского государственного университета;
e-mail: anchugova.ae@gmail.com;
Шакаева Дарина Юсуповна — аспирант кафедры теоретической физики Челябинского государственного университета;
e-mail: sh.darina@mail.ru.

Дата поступления — 30 июня 2019 г.

ЭКСПЕРИМЕНТАЛЬНЫЕ И ЧИСЛЕННЫЕ ИССЛЕДОВАНИЯ ПРОЦЕССОВ РАСПРОСТРАНЕНИЯ ПРИМЕСЕЙ ОТ ВЫСОТНЫХ ИСТОЧНИКОВ

А. А. Леженин[1], В. Ф. Рапута[1,2], Р. А. Амикишиева[3]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[2] *Новосибирский НИИ гигиены Роспотребнадзора, 630108, Новосибирск*

[3] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 551.511.42.001.572
DOI: 10.24411/9999-016A-2019-10043

В работе изложен способ вычисления вертикальных профилей ветра и коэффициента турбулентного обмена в пограничном слое атмосферы. Метод основан на использовании спутниковых снимков дымовых факелов от разновысотных труб ТЭЦ, данных метеорологических и аэрологических наблюдений. Модель оценивания базируется на решениях уравнений экмановского пограничного слоя. Для г. Барнаула приведен анализ распространения дымового шлейфа ТЭЦ-3 и представлены расчеты компонентов скорости ветра в нижней атмосфере. Показана возможность применения предложенного подхода для исследования загрязнения атмосферы города.

Ключевые слова: спутниковые наблюдения, дымовой факел, метеорологические параметры, уравнения Экмана, обратная задача

Введение

Спутниковые снимки позволяют прослеживать траектории дымовых шлейфов от высотных труб ТЭЦ и промышленных предприятий [1–2]. При определенных метеорологических условиях на снимках отчётливо видна активная фаза подъема дымового факела, а его распространение происходит по направлениям ветра на высотах источников. Отметим, что с помощью наземных наблюдений невозможно оперативно получить полномасштабную картину распространения дымовых шлейфов. Использование спутниковой информации позволяет решить эту проблему. На снимках из космоса при определенных метеорологических условиях дымовые факела от труб крупных ТЭЦ могут прослеживаться на несколько десятков километров. Поперечные размеры шлейфов примеси отражают интенсивность диффузионных процессов в атмосфере [3–5]. Численное моделирование создает большие возможности для описания процессов распространения атмосферных примесей [6-10]. В рамках постановок обратных задач использование данных измерений позволяет проводить оценивание параметров атмосферных процессов и характеристик источников примесей [1, 11–13]. В частности, представляет интерес задача определения эффективной высоты выбросов в условиях реальной атмосферы с учетом динамических и тепловых характеристик примеси [14, 15]. Целью исследования является оценивание составляющих скорости ветра по аэрологической и спутниковой информации.

1 Материалы и объекты исследования

Материалами для исследований послужили снимки из космоса г. Барнаула в зимний период времени. На них хорошо видны дымовые факелы от труб ТЭЦ-2 и ТЭЦ-3, имеющих высоту от 100 до 230 метров соответственно. На снимке с ИСЗ "Landsat-8"за 29 января 2019 года (рис. 1), полученном в Сибирском центре ФГБУ

Работа выполнена в рамках Госзадания № 0315-2019-0004, программы президиума РАН № 51 "Изменение климата: причины, риски, последствия, проблемы адаптации и регулирования" (№ 0315-2018-0016), финансовой поддержке РФФИ и Правительства Новосибирской области в рамках научного проекта № 19-47-540008

ISBN 978-5-901548-42-4

«Научно-исследовательский центр космической гидрометеорологии «ПЛАНЕТА» (http://www.rcpod.ru), представлены траектории шлейфов дыма от рассматриваемых ТЭЦ.



Рис. 1: Спутниковый снимок г. Барнаула за 29 января 2019 года на 12 часов местного времени с ИСЗ "Landsat-8". 1 — ТЭЦ-2, 2 — ТЭЦ-3

Данные метеорологической станции г. Барнаула (индекс ВМО 29838) за 28-30 января 2019 г. (рис. 2) показывают, что в течение 29 января в приземном слое атмосферы наблюдались штилевые условия и ветра до 2 м/с южного, юго-восточного направлений, суточный ход температуры воздуха был слабовыраженным.

В табл. 1 приведены распределения по высоте основных метеорологических величин в нижней атмосфере на рассматриваемый день: давление, температура, направление и скорость ветра в 7 часов местного времени, полученные на аэрологической станции. Видно, что в это время были слабые ветра и температурная стратификация от поверхности земли до 300 метров, близкая к нейтральной, а далее до 800 м наблюдался инверсионный ход температуры.

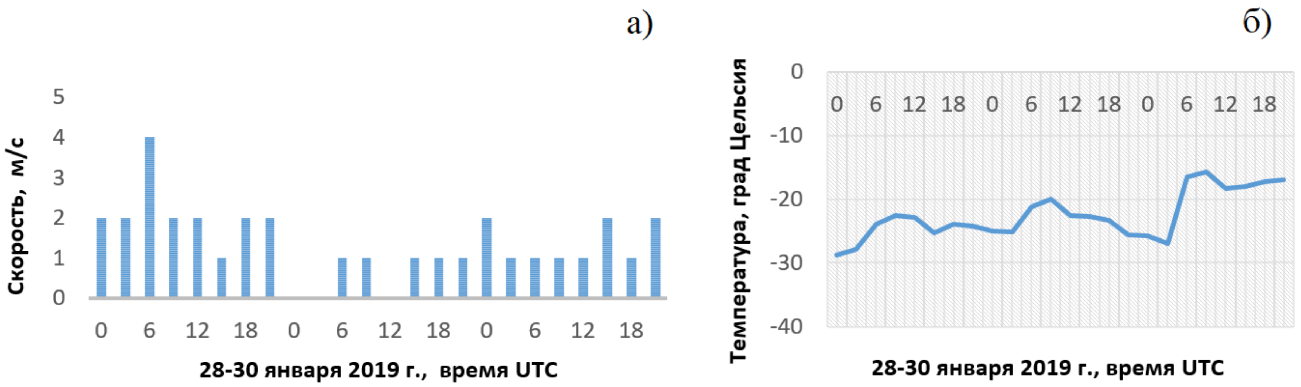


Рис. 2: Ход скорости ветра (а) и температуры воздуха (б) с 28 по 30 января 2019 г. на метеостанции г. Барнаула

Таблица 1: Метеорологические параметры в нижней атмосфере по данным аэрологической станции г. Барнаула на 07 часов местного времени (00 UTC) 29 января 2019 г.

Высота над уровнем моря, м	Давление hPa	Температура C	Направление ветра, град	Скорость ветра м/с
159	1031	-25.1	210	2
369	1000	-25.1	215	3
405	995	-25.1	219	4
472	986	-22.4	225	5
570	973	-18.5	260	6
916	929	-16.7	260	6
948	925	-16.9	260	6
1420	868	-21.5	264	8
1539	854	-21.2	265	8
1574	850	-21.1	265	9

2 Модель оценивания вертикального профиля ветра

Для определения скорости ветра воспользуемся уравнениями экмановского пограничного слоя [16,17]:

$$\frac{\partial}{\partial z}k(z)\frac{\partial u}{\partial z} + 2\omega_z(v - V_g) = 0, \quad \frac{\partial}{\partial z}k(z)\frac{\partial v}{\partial z} - 2\omega_z(u - U_g) = 0 \tag{1}$$

с граничными условиями
при $z = 0$

$$u = 0, \quad v = 0,$$

при $z = H$

$$u = U_g, \quad v = V_g. \tag{2}$$

Здесь $u(z,\vec{\theta}), v(z,\vec{\theta})$ — горизонтальные компоненты скорости ветра в направлении осей x, y соответственно, ось z направлена вертикально вверх, U_g, V_g — горизонтальные составляющие скорости геострофического ветра, ω_z — угловая скорость вращения Земли, k — коэффициент вертикального турбулентного обмена, H — высота атмосферного пограничного слоя, $\vec{\theta}$ — вектор неизвестных параметров. При выводе системы уравнений (1) предполагалось, что ось x совпадает с направлением приземного ветра.

Для задания коэффициента вертикального турбулентного обмена могут использоваться различные подходы в рамках моделей К-теорий [16]. Оценивание вектора параметров $\vec{\theta}$ проводится с использованием измерений скорости ветра на определенных высотах от поверхности земли путем минимизации следующего функционала:

$$J(\vec{\theta}) = \sum_{i=1}^{M} \{[u(z_i, \vec{\theta}) - u_i]^2 + [v(z_i, \vec{\theta}) - v_i]^2\}, \tag{3}$$

где u_i, v_i — измеренные компоненты скорости ветра на высотах z_i . Задача оценивания существенно упрощается, когда коэффициент турбулентного обмена принимается постоянным по высоте. В этом случае решение уравнений (1) имеет вид [16, 17]

$$u(z, \vec{\theta}) = U_g - e^{-az}[U_g \cos(az) + V_g \sin(az)],$$

$$\nu(z, \vec{\theta}) = V_g - e^{-az}[V_g \cos(az) - U_g \sin(az)], \tag{4}$$

где $\vec{\theta} = (U_g, V_g, a)$, $a = \sqrt{\frac{\omega_z}{k}}$.

Для определения вектора параметров $\vec{\theta}$ необходимо решить задачу минимизации функционала (3) с учетом (4). Поскольку компоненты скорости геострофического ветра входят линейно в соотношение (4), размерность задачи (3) может быть существенно снижена и сводится к задаче минимизации функции одного переменного, зависящего от параметра a . Соотношения (3), (4) позволяют получить оценку вектора параметров $\vec{\theta}$ по измерениям составляющих скорости ветра на высотах $z_(i), i = 1, \ldots, M$. Подставляя найденное значение $\vec{\theta}$ в соотношение (4), можно определить вертикальные профили составляющих скорости ветра в пограничном слое атмосферы.

3 Результаты и обсуждения

Анализ рис. 1 показывает, что вынос газоаэрозольной смеси от высотной трубы ТЭЦ-3 происходит в северо-восточном направлении. Траектория дымового шлейфа от трубы ТЭЦ-2 претерпевает более сложную эволюцию. На начальном участке траектории, в пределах приземного слоя атмосферы, шлейф направлен на восток-северо-восток. По мере его подъема, обусловленного динамическим импульсом и потоком плавучести, направление выноса смещается в северо-восточный сектор. Для ТЭЦ-3 в рассматриваемых метеоусловиях эффективная высота подъема примеси находится в верхней части пограничного слоя атмосферы. Как правило, на таких высотах фиксируется правый поворот ветра [16-18]. Анализ рис. 1 подтверждает наличие этого эффекта. Угол между траекториями дымовых шлейфов от труб ТЭЦ-2 и ТЭЦ-3 составляет около 20 градусов. С использованием модели (3)–(4), информации о величине угла между шлейфами и данных табл. 1 проведено оценивание компонентов скорости ветра.

На рис. 3 представлены результаты численного восстановления составляющих скорости ветра для г. Барнаула на 12 часов местного времени 29 января 2019 года.

Из анализа рис. 3 следует, что в нижней атмосфере при температурной стратификации, близкой к нейтральной, рассчитанный профиль ветра качественно согласуется с типичным распределением. При этом, скорость ветра совпадает с скоростью геострофического ветра на высотах 600–700 метров от поверхности земли.

Заключение

Предложена модель оценивания горизонтальных составляющих скорости ветра в нижней атмосфере. Метод базируется на использовании аэрологических данных и спутниковой информации о траекториях дымовых шлейфов от разновысоких труб ТЭЦ. В основу модели оценивания положены уравнения экмановского пограничного слоя. Проведен анализ распространения дымовых факелов от труб ТЭЦ-2 и ТЭЦ-3 г. Барнаула, также представлены результаты моделирования компонентов скорости ветра в пограничном слое атмосферы.

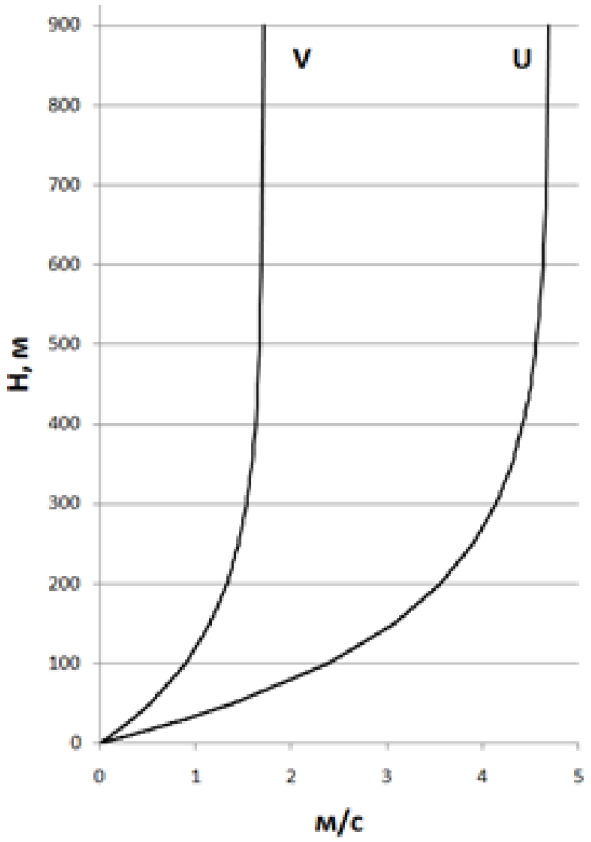


Рис. 3: Восстановленные по модели (3)–(4) составляющие скорости ветра u(z),v(z)

Список литературы

[1] Б. М. Балтер, Д. Б. Балтер, В. В. Егоров, М. В. Стальная //Использование данных ИСЗ Landsat для определения концентрации загрязнителей в шлейфах от продувки газовых скважин на основании модели источника // Исследование Земли из космоса. 2014. № 2. С. 55–66.

[2] К. Я. Кондратьев, А. А. Григорьев, О. М. Покровский, Е. В. Шалина. //Космическое дистанционное зондирование атмосферного аэрозоля. Л. :Гидрометеоиздат, 1983. 216 с.

[3] Obolkin V. A., Potemkin V. L., Makukhin V. L., Chipanina Y. V., Marinayte I. I. Low-level atmospheric jets as main mechanism of long-range transport of power plant plumes in the Lake Baikal Region // Int. J. Environ. Studies. 2014. Vol. 71 (3). P. 391-397.

[4] Balter B. M., Faminskaya M. V. Irregularly emitting air pollution sources: acute health risk assessment using aermod and the Monte Carlo approach to emission rate // Air Quality, Atmosphere and Health. 2017. Vol. 10 (4). P. 401-409.

[5] Solomos S., Amiridis V., Zanis P., Gerasopoulos E., Sofiou F. I., Herekakis T., Brioude J.,Stohl A., Kahn R. A., Kontoes C. // Smoke dispersion modeling over complex terrain using high resolution meteorological data and satellite observations — The FireHub platform // Atmos. Environ. 2015. Vol. 119. P. 348–361.

[6] Lezhenin A. A., Raputa V. F., Yaroslavtseva T. V. Numerical Analysis of Atmospheric Circulation and Pollution Transfer in the Environs of Norilsk Industrial Region // Atmospheric and Oceanic Optics. 2016. Vol. 29 (6). P. 565–569.

[7] Gao Z., Bresson R.,Qu Y., Milliez M., Munck C., Carissimo B.//High resolution unsteady RANS simulation of wind, thermal effects and pollution dispersion for studying urban renewal scenarios in a neighborhood of Toulouse// Urban Climate. 2018. Vol. 23. P. 114-130.

[8] Sokhi R., Baklanov A., Schluenzen H. (Eds.) Air Pollution and Meteorological Modelling for Atmospheric Research and Policy Applications, Anthem Press, 2017. 260 p.

[9] Li H., Cui G., Zhang Z. A New Scheme for the Simulation of Microscale Flow and Dispersion in Urban Areas by Coupling Large-Eddy Simulation with Mesoscale Models // Boundary-Layer Meteorol. 2018. Vol. 167. P. 145–170.

[10] Fallah-Shorshani M., Shekarrizfard M., Hatzopoulou M. Evaluation of regional and local atmospheric dispersion models for the analysis of traffic-related air pollution in urban areas // Atmos. Environ. 2017. Vol. 167. P. 270–282.

[11] Рапута В. Ф. , Шлычков В. А., Леженин А. А., Романов А. Н., Ярославцева Т. В.// Численный анализ данных аэрозольных выпадений примесей от высотного источника // Оптика атмосферы и океана. 2014. Т. 27. № 8. С. 713–718.

[12] Леженин А. А., Ярославцева Т. В., Рапута В. Ф. Использование спутниковой информации о траекториях дымовых факелов для расчета полей ветра // Интерэкспо ГЕО-Сибирь-2016. Т. 4. № 1. С. 63–67 .

[13] Islam M. A., Roy G. D. A mathematical model in locating an unknown emission source // Water, Air, and Solid Pollut. 2002. Vol. 136. № 1/4. P. 331–345.

[14] Бызова Н. Л., Гаргер Е. К., Иванов В. Н. Экспериментальные исследования атмосферной диффузии и расчеты рассеяния примеси. Л. : Гидрометеоиздат, 1991. 278 с.

[15] Рапута В. Ф., Леженин А. А, Ярославцева Т. В. Оценка параметров выбросов Новосибирских ТЭЦ с использованием спутниковой информации // Интерэкспо ГЕО-Сибирь. 2018. Т.1. № 4. С. 137–146.

[16] Браун Р. А. Аналитические методы моделирования планетарного пограничного слоя. Л : Гидрометеоиздат, 1978. 151 с.

[17] Динамическая метеорология. / Под ред. Д. Л. Лайхмана. Л : Гидрометеоиздат, 1976. 608 с.

[18] Зилитинкевич С. С. Атмосферная турбулентность и планетарные пограничные слои. М. : ФИЗМАТЛИТ, 2013. 252 с.

Леженин Анатолий Александрович – к.ф.-м.н., ст. науч. сотрудник Института вычислительной математики и математической геофизики СО РАН;
e-mail: lezhenin@ommfao.sscc.ru;
Рапута Владимир Федотович – д.ф.-м.н., главный научный сотрудник Института вычислительной математики и математической геофизики СО РАН;
Новосибирский НИИ гигиены Роспотребнадзора;
e-mail: raputa@sscc.ru;
Амикишиева Руслана Александровна – магистрант Новосибирского государственного университета;
e-mail: ruslana215w@mail.ru.
Дата поступления — 31 мая 2019 г.

КРИТЕРИИ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ ПРИ АНАЛИЗЕ БОЛЬШИХ ВЫБОРОК: ПРОБЛЕМЫ И ИХ РЕШЕНИЕ

Б. Ю. Лемешко, С. Б. Лемешко, И. В. Веретельникова, П. Ю. Блинов[1]

[1] *Новосибирский государственный технический университет, 630073, Новосибирск*

УДК 519.24
DOI: 10.24411/9999-016A-2019-10044

В работе рассмотрены методы построения оценок при анализе больших данных (Big Data). Демонстрируется влияние на результаты выводов по критерию χ^2 Пирсона выбора числа интервалов и способа группирования. Показывается, как влияет на распределения статистик непараметрических критериев согласия ограниченная точность представления данных в больших выборках. Даются рекомендации по применению критериев для анализа больших выборок. Показано, что на распределения статистик критериев однородности влияет неравноточность представления данных в сравниваемых выборках.

Ключевые слова: Big Data, оценивание параметров, проверка гипотез, критерии согласия, критерии однородности, статистическое моделирование.

Введение

В связи с увеличением объёмов информации в последние годы резко возрос интерес к вопросам применения статистических методов для анализа больших массивов данных (Big Data). При попытках применения для анализа больших выборок классического аппарата прикладной математической статистики, как правило, сталкиваются со специфическими проблемами, ограничивающими возможности корректного применения этого аппарата.

В настоящей работе мы будем касаться только методов и критериев, связанных с анализом одномерных случайных величин, реальные проблемы которых нам наиболее знакомы. Можно рассмотреть, по крайней мере, три ситуации, где рост размерности выборок вызывает проблемы в применении методов или критериев.

Во-первых, вследствие “проклятия размерности” хорошо зарекомендовавшие себя методы и алгоритмы становятся неэффективными. При оценивании параметров моделей с ростом размерности анализируемых выборок кардинально растут вычислительные затраты, ухудшается сходимость итерационных алгоритмов, используемых при нахождении оценок. Естественным способом разрешения данной ситуации видится применение методов, предусматривающих группирование данных. В таком случае возникают вопросы, как использование оценок по группированным данным отразится на свойствах критериев проверки гипотез, в которых будут использоваться эти оценки [1]?

Во-вторых, многие критерии проверки статистических гипотез не приспособлены даже для анализа выборок порядка тысячи наблюдений, так как информация о распределениях статистик этих критериев при справедливости проверяемой гипотезы представлена лишь краткими таблицами критических значений. Возможность применения таких критериев при “разумных” величинах объёмов выборок n легко разрешается статистическим моделированием распределений статистик в интерактивном режиме в ходе статистического анализа [2].

В-третьих, замечено, что применение, например, непараметрических критериев согласия, для которых известны предельные (асимптотические) распределения статистик, с ростом объёмов выборок всегда приводит к отклонению даже справедливой проверяемой гипотезы. В [3] показано, что корни этой проблемы

Работа выполнена при поддержке Министерства образования и науки РФ в рамках государственной работы «Обеспечение проведения научных исследований» (№ 1.4574.2017/6.7) и проектной части государственного задания (№ 1.1009.2017/4.6).

ISBN 978-5-901548-42-4

связаны с ограниченной точностью представления анализируемых данных. В данном случае будет показано, что причина некорректности применения к большим выборкам критериев проверки гипотез об однородности кроется в неравноточности измерений в анализируемых выборках.

1 Оценивание параметров законов распределения

Оценки параметров законов могут находиться различными методами. Наилучшими асимптотическим свойствами обладают оценки максимального правдоподобия (ОМП), вычисляемые в результате максимизации функции правдоподобия

$$\hat{\theta} = arg \max_{\theta} \prod_{j=1}^{n} f(x_j, \theta), \tag{1}$$

или её логарифма, где θ — неизвестный вектор параметров, $f(x, \theta)$ — функция плотности закона распределения, $x_1, x_2, \ldots, x_n$ — выборка, по которой вычисляется оценка $\hat{\theta}$. ОМП вектора параметров в большинстве случаев находятся в результате использования некоторого итерационного метода.

При вычислении MD-оценок (оценок минимального расстояния) по θ минимизируется некоторая мера близости (расстояние) $\rho(F(x, \theta), F_n(x))$ между теоретическим $F(x, \theta)$ и эмпирическим $F_n(x)$ распределениями. MD-оценки находятся в процессе решения задачи

$$\hat{\theta} = arg \min_{\theta} \rho(F(x, \theta), F_n(x)), \tag{2}$$

В качестве мер близости можно использовать, например, статистики непараметрических критериев согласия (Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Купера, Ватсона и других).

ОМП параметров законов распределения, как правило, не являются робастными. Наличие аномальных наблюдений или ошибочность предположения о виде закона приводят к построению моделей с функциями распределения, неприемлемо отклоняющимися от эмпирических распределений. MD-оценки обладают большей устойчивостью.

Очевидно, что при очень больших выборках вычисление оценок (1) и (2) связано с серьёзными вычислительными трудностями.

В случае группированной выборки имеющаяся в нашем распоряжении информация связана с множеством непересекающихся интервалов, которые делят область определения случайной величины на k непересекающихся интервалов граничными точками

$$x_{(0)} < x_{(1)} < \cdots < x_{(k-1)} < x_{(k)},$$

где $x_{(0)}$ — нижняя грань области определения случайной величины X; $x_{(k)}$ — верхняя грань области определения случайной величины X.

ОМП по группированной выборке вычисляется в результате максимизации функции правдоподобия

$$\hat{\theta} = arg \max_{\theta} \prod_{i=1}^{k} P_i^{n_i}(\theta), \tag{3}$$

где $P_i(\theta) = \int\limits_{x_{(i-1)}}^{x_{(i)}} f(x, \theta) dx$ — вероятность попадания наблюдения в i-й интервал значений, n_i — количество наблюдений, попавших в i-й интервал, $\sum\limits_{i=1}^{k} n_i = n$.

При наличии негруппированных данных к оценкам по группированным данным обращаются редко. Связано это с большей трудоёмкостью вычислительного процесса и необходимостью многократного использования численного интегрирования при вычислении $P_i(\theta)$ и требует соответствующей программной поддержки. В случае больших объёмов выборок ситуация меняется. При фиксированном числе интервалов группирования с ростом объёмов выборок вычислительные затраты не меняются, а возрастают только с увеличением количества интервалов k. Это свидетельствует о целесообразности в условиях Big Data использовать ОМП по группированным выборкам. Это асимптотически эффективные и, к тому же, робастные оценки. При малом k качество оценок можно улучшать, используя асимптотически оптимальное группирование (АОГ) [4], при котором минимизируются потери в информации Фишера, связанные с группированием.

2 Применение критерия χ^2 Пирсона к большим выборкам

Статистику критерия согласия χ^2 Пирсона вычисляют по формуле

$$X_n^2 = n\sum_{i=1}^{k}\frac{(n_i/n - P_i(\theta))^2}{P_i(\theta)} \tag{4}$$

В случае проверки простой гипотезы при $n \to \infty$ эта статистика подчиняется χ_r^2-распределению с $r = k-1$ степенями свободы, если верна нулевая гипотеза.

При проверке сложной гипотезы и оценивании по выборке m параметров закона статистика (4) в случае справедливости H_0 подчиняется χ_r^2-распределению с $r = k - m - 1$ степенями свободы, если оценки получаются минимизацией (4) этой статистики, или используются ОМП (3) (или другие асимптотически эффективные оценки по группированным данным).

При оценивании параметров по негруппированным данным распределение статистики (4) не подчиняется χ^2_{k-m-1}-распределению. При использовании ОМП по негруппированным данным рекомендуется применять критерий Никулина–Рао–Робсона [5,6]. Хотя, опираясь на статистическое моделирование и интерактивный режим [2] моделирования распределений статистики, не исключается использование критерия χ^2 Пирсона [7].

Принципиальные проблемы, препятствующие применению критерия χ^2 Пирсона для анализа Big Data, отсутствуют: возможны только вычислительные трудности.

Проиллюстрируем результаты применения критерия χ^2 Пирсона на примере достаточно большой выборки объёмом $n = 10^7$, смоделированной по стандартному нормальному закону $N(0,1)$.

В таблице 1 представлены результаты применения критерия при проверке простой гипотезы о принадлежности выборки закону $N(0,1)$ при различном числе интервалов в случае равночастотного группирования (РЧГ) и при $k = 15$ в случае (АОГ). При АОГ максимизируется мощность критерия χ^2 Пирсона относительно близких конкурирующих законов [8]. В таблице приведены значения X_n^{2*} статистики (4), вычисленные по выборке, и соответствующие значения достигнутого уровня значимости $p_{value} = P\{X_n^2 \geq X_n^{2*}|H_0\}$. Как можно видеть, результаты зависят как от способа разбиения, так и от числа интервалов. От этого же зависит и мощность критерия [9].

Таблица 1: Результаты проверки простой гипотезы

	АОГ	РЧГ						
	$k = 15$	$k = 15$	$k = 50$	$k = 75$	$k = 100$	$k = 500$	$k = 1000$	$k = 2000$
X_n^{2*}	7.75162	9.18380	56.8942	79.4904	96.5701	493.995	1044.57	2099.91
p_{value}	0.90186	0.81910	0.20475	0.31026	0.55038	0.55482	0.15403	0.05702

Таблица 2: Результаты проверки сложной гипотезы

	АОГ	РЧГ						
k	15	15	50	75	100	500	1000	2000
$\hat{\theta}_0$	0.000276	0.000301	0.0002440	0.000270	0.000268	0.000277	0.000273	0.000274
$\hat{\theta}_1$	1.007150	1.002629	1.001730	1.001338	1.001123	1.000399	1.000305	1.000236
X_n^{2*}	927.9202	99.99627	101.7669	104.5111	112.1514	493.7161	1043.471	2098.605
p_{value}	0.0	5.58e-16	6.50e-06	0.007396	0.139377	0.533166	0.149218	0.055723

В таблице 2 представлены результаты проверки сложных гипотез. Там приведены ОМП $\hat{\theta}_0$ параметра сдвига и $\hat{\theta}_1$ параметра масштаба нормального закона по группированным данным, полученные при соответствующем числе интервалов k, значения статистик X_n^{2*} и p_{value}.

ОМП параметров по полной негруппированной выборке $\hat{\theta}_0 = 0.000274$, $\hat{\theta}_1 = 1.000177$. В [7] построены модели распределений статистики (4) для случая проверки сложной гипотезы относительно нормального закона с использованием ОМП по негруппированным данным и применением АОГ. Вычисленное по выборке

значение статистики $X_n^{2*} = 6.600521$ при $k = 15$, а полученная в соответствии с приведенной в [7] моделью предельного распределения оценка $p_{value} = 0.886707$, что свидетельствует о хорошем согласии полной выборки с нормальным законом $N(0.000274, 1.000177)$.

В случае анализа больших выборок [3] критерий χ^2 Пирсона демонстрирует как свои положительные качества [8], так и недостатки: зависимость результатов анализа (и мощности критерия) как от способа разбиения, так и от числа интервалов [9].

3 Непараметрические критерии согласия при больших выборках

Как правило, в Big Data объёмы выборок практически неограничены, но сами данные представлены с ограниченной точностью (округлены с некоторым $\triangle$). По сути, "нарушается предположение" о том, что наблюдается непрерывная случайная величина. Именно этот факт, как подробно показано в [3], является основной причиной возможной некорректности выводов при анализе больших данных с использованием непараметрических критериев согласия.

В данном случае мы ограничимся демонстрацией зависимости от $\triangle$ распределения статистики критерия согласия Крамера–Мизеса–Смирнова. Статистика критерия имеет вид

$$S_\omega = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ F(x_i, \theta) - \frac{2i-1}{2n} \right\}^2, \tag{5}$$

и при проверке простой гипотезы в пределе подчиняется закону с функцией распределения $a1(s)$ [10].

Покажем зависимость распределения статистики от $\triangle$ для случая принадлежности выборок стандартному нормальному закону. При округлении с точностью до 1 в выборках, принадлежащих $N(0,1)$, может появляться 9 уникальных значений, при округлении с точностью до $\triangle = 0.1$ — порядка 86 уникальных значений, с точностью $\triangle = 0.01$ — порядка 956, с точностью до $\triangle = 0.001$ — порядка 9830.

На рис. 1 представлена зависимость распределений статистики (5) критерия Крамера–Мизеса–Смирнова от степени округления $\triangle$ при объёмах выборок $n = 1000$ для случая проверки простой гипотезы о принадлежности выборки стандартному нормальному закону. На рисунке приведено предельное распределение $a1(S)$, а также реальные распределения $G(S_{1000}|H_0)$ статистики при степени округления $\triangle$ =0.01, 0.05, 0.1, 0.2, 0.3. Как можно видеть, при $\triangle = 0.01$ распределение $G(S_{1000}|H_0)$ практически не отличается от $a1(S)$, но с ростом $\triangle$ отклонение $G(S_{1000}|H_0)$ от $a1(S)$ быстро увеличивается.

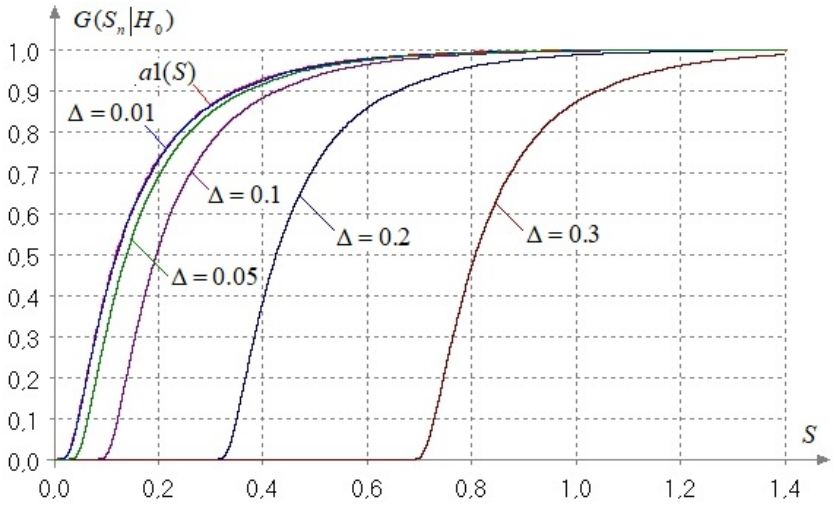


Рис. 1: Распределения статистики $G(S_n|H_0)$ критерия Крамера–Мизеса–Смирнова в зависимости от $\triangle$ при $n = 1000$

Таким же образом меняются распределения статистик непараметрических критериев согласия при проверке сложных гипотез вида H_0: $F(x) \in \left\{ F(x, \theta), \theta \in \Theta \right\}$, где Θ — область определения параметра θ. Следует напомнить, что при проверке сложных гипотез уже нельзя использовать классические результаты [11], так

как предельные распределения зависят от ряда факторов, связанных с оцениванием параметров [12]. И на всё это ещё накладывается влияние степени округления данных.

В работе [3] на основании результатов исследований сформулированы следующие рекомендации. Для того чтобы при анализе больших выборок с применением соответствующего непараметрического критерия согласия можно было использовать классические результаты, статистика должна вычисляться не по всему большому массиву, а по выборкам, извлекаемым по равномерному закону из той "генеральной совокупности", роль которой играет анализируемый большой массив данных. Объём извлекаемой выборки должен учитывать точность $\triangle$ и не превышать некоторой величины n_{max}, при которой (при данном $\triangle$) распределение статистики $G(S_{n_{max}}|H_0)$ критерия при справедливости H_0 ещё реально не отличается от предельного распределения $G(S|H_0)$ статистики этого критерия.

Вышесказанное в полной мере относится к применению к большим выборкам любых непараметрических критериев согласия.

Распределения статистик 3-х критериев согласия Жанга [13], представляющих собой развитие критериев Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга, зависят от объёмов выборок n. В этом случае не может идти речи об использовании предельных распределениях статистик. Но распределения статистик $G(S_n|H_0)$ этих критериев таким же образом зависят от от степени округления $\triangle$. Проблема может разрешаться статистическим моделированием (в том числе, в интерактивном режиме [2]) распределений статистик при заданных n и $\triangle$ при справедливости проверяемой гипотезы H_0. Именно так в подобной ситуации эта проблема разрешается в развиваемой программной системе ISW [17].

Отметим, что подобным же образом степень округления регистрируемых данных влияет на свойства множества других критериев, в частности, специальных критериев, ориентированных на проверку гипотезы о принадлежности выборок нормальному закону, о принадлежности выборок равномерному закону, показательному закону и др.

4 Критерии однородности при больших выборках

В критериях однородности, где сравнивается 2 и более выборок, на распределения статистик влияет неравноточность данных, представленных в выборках.

Двухвыборочный критерий однородности Лемана–Розенблатта предложен в работе [15] и исследован в [16]. Статистика, построенная по двум выборкам $x_{11}, x_{12}, \ldots, x_{1,n_1}$ и $x_{21}, x_{22}, \ldots, x_{2,n_2}$, используется в форме

$$S_{LR} = \frac{1}{n_1 n_2 (n_1 + n_2)} \Big[n_1 \sum_{i=1}^{n_1} (r_i - i)^2 + n_2 \sum_{j=1}^{n_2} (s_j - j)^2 \Big] - \frac{4 n_1 n_2 - 1}{6(n_1 + n_2)}, \tag{6}$$

где r_i — порядковый номер (ранг) x_{1i}; s_j — порядковый номер (ранг) x_{2j} в объединенном вариационном ряде. Предельным распределением статистики (6) при справедливости проверяемой гипотезы H_0 является то же самое распределение $a1(s)$, которое является предельным для статистики критерия согласия Крамера–Мизеса–Смирнова.

Рис. 2 демонстрирует зависимость распределения статистики $G(S_{LR}|H_0)$ критерия однородности Лемана–Розенблатта от степени округления $\triangle_2$ наблюдений во второй выборке при округлении в первой выборке $\triangle_1 = 0.01$ при объёмах выборок $n_i = 1000$. Уже при $\triangle_2 = 0.05$ отклонение $G(S_{LR}|H_0)$ от $a1(S)$ оказывается существенным. При фиксированном $\triangle_2$ с ростом объёмов выборок отклонение $G(S_{LR}|H_0)$ от $a1(S)$ быстро увеличивается. Отклонение увеличивается с ростом $\triangle_2$ и фиксированном объёме выборки. Распределения статистики $G(S_{LR}|H_0)$ критерия однородности Лемана–Розенблатта зависят также от разности $\triangle_1$ и $\triangle_2$.

Аналогичным образом от различия в точности регистрируемых данных в выборках зависят распределения статистик других критериев однородности, рассмотренных в [17].

Заключение

При построении вероятностных моделей по большим выборкам целесообразно использование методов оценивания параметров, предусматривающих группирование данных. В отличие от оценок по негруппированным данным такие оценки робастны, а вычислительные затраты не зависят от объёмов выборок.

Критерий χ^2 Пирсона при анализе больших выборок сохраняет как свои положительные качества, так и свойственные ему недостатки.

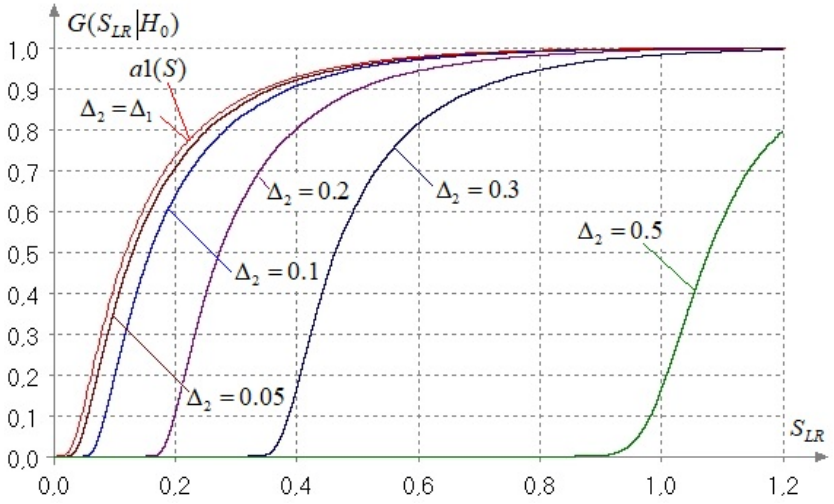


Рис. 2: Распределения статистики $G(S_{LR}|H_0)$ критерия однородности Лемана–Розенблатта при $n_i = 1000$ в зависимости от $\triangle_2$ при $\triangle_1 = 0.01$

Корректность применения непараметрических критериев согласия для анализа больших выборок можно обеспечить ограничением объёмов выборок, извлекаемых из больших совокупностей. Можно также использовать методы статистического моделирования для нахождения реальных распределений статистик (при соответствующих $\triangle$).

Возможную некорректность выводов при использовании критериев однородности, связанную с неравноточностью измерений в выборках, можно также устранить применением статистического моделирования для нахождения реальных распределений статистик.

Подобная стратегия действий при анализе больших выборок реализуется в программной системе ISW [14].

Список литературы

[1] Lemeshko B. Yu., Lemeshko S. B., Postovalov S. N. Statistic Distribution Models for Some Non-parametric Goodness-of-Fit Tests in Testing Composite Hypotheses // Communications in Statistics — Theory and Methods. 2010. Vol. 39, № 3. — P. 460-471.

[2] Lemeshko B. Yu., Lemeshko S. B., Rogozhnikov A. P. Interactive investigation of statistical regularities in testing composite hypotheses of goodness of fit // Statistical Models and Methods for Reliability and Survival Analysis : monograph. — Wiley-ISTE , 2013. — Chap. 5. — P. 61–76.

[3] Лемешко Б. Ю., Лемешко С. Б., Семёнова М. А. К вопросу статистического анализа больших данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. № 44. — С. 40-49. DOI: 10.17223/19988605/44/5

[4] Лемешко Б. Ю. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход /Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. Новосибирск : Изд-во НГТУ, 2011. — 888 с.

[5] Никулин М. С. О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. 1973. Т. XVIII. № 3. — С.75-676.

[6] Rao K. C., Robson D. S. A chi-squared statistic for goodness-of-fit tests within the exponential family // Commun. Statist. 1974. Vol. 3. — P.1139-1153.

[7] Лемешко Б. Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению. М.: ИНФРА-М, 2015. — 160 с. DOI: 10.12737/6086

[8] Денисов В. И., Лемешко Б. Ю. Оптимальное группирование при обработке экспериментальных данных // Измерительные информационные системы. Новосибирск, 1979. — С.5-14.

[9] Лемешко Б. Ю., Чимитова Е. В. О выборе числа интервалов в критериях согласия типа χ^2 // Заводская лаборатория. Диагностика материалов. 2003. Т. 69. № 1. С. 61-67.

[10] Большев Л. Н., Смирнов Н. В Таблицы математической статистики. — М. : Наука, 1983. — 416 с.

[11] Kac M., Kiefer J., Wolfowitz J. On tests of normality and other J. tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. Vol. 26. — P.189-211.

[12] Лемешко Б. Ю. Непараметрические критерии согласия: Руководство по применению. М.: ИНФРА-М, 2014. — 163 с. DOI: 10.12737/11873

[13] Zhang J. Powerful goodness-of-fit tests based on the likelihood ratio // Journal of the Royal Statistical Society: Series B. 2002. V.64. № 2. — P.281-294.

[14] ISW–Программная система статистического анализа одномерных наблюдений. https://ami.nstu.ru/ headrd/ISW.htm. (дата обр. 30.04.2019)

[15] Lehmann E. L. Consistency and unbiasedness of certain nonparametric tests // Ann. Math. Statist. — 1951. — Vol. 22, № 1. — P. 165–179.

[16] Rosenblatt M. Limit theorems associated with variants of the von Mises statistic // Ann. Math. Statist. — 1952. — Vol. 23. — P. 617–623.

[17] Лемешко Б. Ю. Критерии проверки гипотез об однородности. Руководство по применению. М.: ИНФРА-М, 2017. — 208 с. DOI: 10.12737/22368

Лемешко Борис Юрьевич — д.т.н., профессор, г.н.с. кафедры прикладной и теоретической информатики Новосибирского государственного технического университета, e-mail: Lemeshko@ami.nstu.ru;

Лемешко Станислав Борисович — с.н.с., к.т.н., кафедры прикладной и теоретической информатики Новосибирского государственного технического университета, e-mail: skyer@mail.ru;

Веретельникова Ирина Викторовна — м.н.с. кафедры прикладной и теоретической информатики Новосибирского государственного технического университета, e-mail: ira-veterok@mail.ru;

Блинов Павел Юрьевич — м.н.с. кафедры прикладной и теоретической информатики Новосибирского государственного технического университета, e-mail: Blindizer@yandex.ru

Дата поступления — 30 апреля 2019 г.

СРАВНЕНИЕ РАЗНОСТНЫХ СХЕМ РАЗЛИЧНЫХ ПОРЯДКОВ ТОЧНОСТИ НА АДАПТИВНЫХ СЕТКАХ ДЛЯ РАСЧЕТА ПОГРАНИЧНЫХ И ВНУТРЕННИХ СЛОЕВ

В. Д. Лисейкин[1,2], В. И. Паасонен[1,2]

[1] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.6
DOI: 10.24411/9999-016A-2019-10045

В работе на примере ОДУ второго порядка с малым параметром описывается комплексный подход для расчета задач с пограничными и внутренними слоями, основанный симбиозе разностных схем повышенного порядка точности и применении специальных адаптивных сеток, сгущающихся в зонах быстрого изменения решения. Алгоритм построения адаптивных сеток представляет собой обобщение методики, разработанной ранее для схемы с односторонними разностями, и опирается на априорные оценки старших производных решения. Разработанная технология протестирована на задачах с различными типами слоев с разными масштабами и расположением. В серии численных экспериментов проведено сравнение компактных схем второго и третьего порядка аппроксимации на неравномерных сетках и схемы первого порядка, которое подтвердило практическую применимость обобщенных отображений, генерирующих адаптивные сетки, и высокое качество расчетов с помощью компактных схем на этих сетках.

Ключевые слова: уравнение с малым параметром, погранслой, внутренний слой, компактная схема, схема повышенной точности, адаптивная сетка.

Введение

Рассматривается первая краевая задача для обыкновенного дифференциального уравнения

$$-\varepsilon u'' + a(x)u' + F(x,u) = 0, \quad x \in (0,1), \quad F'_u(u,x) \geq 0, \quad u(0) = U_0, \quad u(1) = U_1, \tag{1}$$

где $\varepsilon > 0$ малый параметр.

Для численного решения задачи (1), являющейся модельной для многих задач механики, связанных с исследованием ударных переходов, пограничных и внутренних слоев и других узких зон чрезвычайно быстрого изменения искомых функций, рассматриваются три различные схемы на трехточечном шаблоне, аппроксимирующие уравнение соответственно с первым, вторым и третьим порядком относительно локального значения шага неравномерной сетки $0 = x_0 < x_1 < ... < x_N = 1$, являющейся образом равномерной сетки при гладком отображении.

В зависимости от поведения функции $a(x)$, расположения ее нулей и значений ее производных реализуются различные по структуре, по числу и по размещению пограничные и внутренние слои [1], [2]. В этих работах по априорной информации о производной решения для схемы первого порядка точности с односторонними разностями были разработаны алгоритмы явного задания адаптивной сетки, квазиравномерной относительно модуля приращения решения на каждом шаге. Опираясь на свойство обратной монотонности, для схемы с учетом знака в работе [3] теоретически доказана сходимость, равномерная по ε.

Обобщение алгоритма задания сеток для случая высокоточных схем построено по аналогии на основе априорных оценок старших производных решения и конструирования для генерации сеток преобразований необходимого класса гладкости [4]. При этом поскольку схемы порядков аппроксимации выше первого не

Работа выполнена при частичной финансовой поддержке Российского научного фонда (код проекта 17-72-30006).

ISBN 978-5-901548-42-4

обладают свойством обратной монотонности, то провести теоретические оценки решения не представляется возможным. Однако является ли это свойство необходимым для равномерной сходимости или оно избыточно, но лишь облегчает процесс доказательства, неизвестно. Ответ на этот вопрос могло в таком случае дать численное исследование, которое и проведено в данной работе для компактных схем второго и третьего порядков точности.

1 Компактные разностные схемы

В качестве классической схемы для сравнения результатов возьмем схему

$$-\varepsilon\Lambda u + a\Delta_{\pm}u + F = 0. \tag{2}$$

первого порядка точности со стандартной аппроксимацией второй производной

$$\Lambda u = \frac{\Delta_+ u - \Delta_- u}{s/2}, \quad s = h_+ + h_-$$

и односторонней разделенной разностью $\Delta_{\pm}u$, выбираемой в зависимости от знака $a(x_i)$. Символами h_+ и h_- обозначены местные значения шагов сетки справа и слева от данного узла.

Во второй схеме первая производная аппроксимируется по трем узлам со вторым порядком, а главный член погрешности разностного аналога второй производной Λu компенсируется дополнительными слагаемыми так, чтобы в целом получилась схема второго порядка аппроксимации

$$-\varepsilon\Lambda u + a\Delta u + \frac{d}{3}\Omega u + (E + \frac{d}{3}\Delta)F = 0, \quad d = h_+ - h_-, \tag{3}$$

где

$$\Delta = \frac{h_-\Delta_+ + h_+\Delta_-}{s}, \quad \Omega = \frac{a_+\Delta_+ - a_-\Delta_-}{s/2}, \quad s = h_+ + h_-,$$

E — тождественный оператор, $a_{\pm}$ — местные значения функции $a(x)$ в полуцелых узлах справа и слева.

Вычислим главные члены разложения погрешности схемы (3) и понизим в них порядок производных с помощью продолженной системы уравнений, а затем аппроксимируем полученные слагаемые разностными выражениями, добавив их с противоположными знаками в схему для компенсации погрешности. В результате получим схему третьего порядка аппроксимации вида

$$(-\varepsilon + \frac{p}{6}\Delta a)\Lambda u + (a + \frac{p}{12}\Lambda a)\Delta u + (\frac{d}{3} - ar)\Omega u + \Sigma F = 0, \quad p = h_-h_+, \tag{4}$$

где

$$\Sigma = E + (\frac{d}{3} - ar)\Delta + \frac{p + d^2}{12}\Lambda, \quad r = \frac{3p + d^2}{36\varepsilon}.$$

Построенные схемы являются модельными по отношению к схемам для решения нелинейных двумерных задач [5], [6], описывающих стационарные течения вязкой несжимаемой жидкости.

Следует ожидать что в данном случае порядок аппроксимации схемы (4) будет выше на единицу, а именно четвертым, так как для квазиравномероной сетки (т. е. являющейся образом равномерной при гладком преобразовании $x(\xi, \varepsilon) \in C^2(0, 1)$) разность соседних шагов $d = h_+ - h_-$ имеет второй порядок малости. Для обеспечения равномерной по ε сходимости разностного решения к точному требуется равномерная по ε оценка производных решения относительно новой сеточной координаты ξ

$$\left|\frac{d^p u(x(\xi, \varepsilon), \varepsilon)}{d\xi^p}\right| \le M, \quad p \le n \tag{5}$$

с константой M, не зависящей от ε.

В тестовых задачах под «точным» решением на приводимых графиках подразумевается численное решение, полученное на самой подробной сетке в процессе ее детализации. Для оценки C-нормы ошибки на сгущающихся сетках при неизвестном точном решении использовалась разность приближенных решений на двух последовательных сетках:

$$\delta^h = \max_{i=0}^{N} |u_{2i}^{2N} - u_i^N|, \quad h = 1/N.$$

где u^N и u^{2N} — приближенные решения на сетках с N и $2N$ шагами. По оценкам ошибки вычислялась апостериорная оценка реально наблюдаемого порядка точности $p^h = \log_2(\delta^h/\delta^{h/2})$. Во всех таблицах оценки ошибки δ и порядка p снабжены нижними индексами со значениями 1, 2 и 3, соответствующими порядку точности схем, по которым получены результаты.

Важной характеристикой алгоритма явного задания узлов адаптивной сетки является равномерность модуля приращения решения на шаге, т. е. оценка $|u^N_{i+1} - u^N_i| \le m/N$, где m не зависит от ε и N. Практически во всех расчетах это свойство имело место. Нарушалось оно лишь при грубой сетке, не достаточно детальной для данного значения вязкости, и обычно проявлялось в осциллирующем характере приближенных решений.

2 Базовое преобразование, генерирующее адаптивную сетку

Для построения разностных сеток, сгущающихся в экспоненциальных слоях возле точки $x_0 = 0$, используем преобразование класса $C^l[0,\ 1]$, устраняющее в слое экспоненциальные особенности масштаба k до порядка n. Базовое преобразование, отображающее взаимно-однозначно отрезок $0 \le \xi \le 1$ на отрезок $0 \le x \le 1$, строится следующим образом. На части отрезка $0 \le \xi \le \xi_0$ определяется предложенная в [7] функция

$$x = \psi(\xi) = \varepsilon^k \left((1 - d\xi)^{-1/a} - 1 \right), \quad 0 \le \xi \le \xi_0, \tag{6}$$

где $d = (1 - \varepsilon^{k\beta})/\xi_0 \ge 1 + m_1 > 1$, $\beta = a/(1 + na)$, при этом a — произвольная положительная константа, отделенная от нуля ($a \ge m_2 > 0$), а k — масштаб слоя. Функция (6) устраняет экспоненциальную особенность в погранслое и гладко сопрягается с полиномом, определенным на остальной части отрезка $\xi_0 \le \xi \le 1$ так, чтобы производные до порядка l составной функции были непрерывны в точке склейки. Для построения этого полинома достаточно знать значения функции (6) и ее производных до порядка l в точке $\xi = \xi_0$, которые выражаются явно

$$\psi^{(i)}(\xi_0) = d^i \frac{1}{a} \Big(\frac{1}{a} + 1 \Big) \dots \Big(\frac{1}{a} + i - 1 \Big) \varepsilon^{ka(n-i)/(1+na)}, \quad i \ge 1. \tag{7}$$

После этого обе ветви склеенной гладкой функции множатся на постоянный коэффициент c_1, с тем чтобы растяжением (сжатием) обеспечить равенство единице значения склеенной функции на правом конце $\xi = 1$. В результате получается составное базовое преобразование

$$x_b(\xi, \varepsilon, a, k, l, n, \xi_0) = \begin{cases} c_1 \psi(\xi), & 0 \le \xi \le \xi_0, \\ c_1 \Big[\displaystyle\sum_{m=0}^{l} \frac{y^l}{l!} \psi^{(l)}(\xi_0) + + c_0 y^{l+1} \Big], & \xi_0 \le \xi \le 1, \end{cases} \tag{8}$$

где $y = \xi - \xi_0$, $l \le n$, $c_0 > 0$. Константа $c_0 > 0$ обеспечивает строгую монотонность составной функции, а $c_1 > 0$ выбирается из условия $x_b(1, \varepsilon, a, k, ...) = 1$.

Преобразование (8) учитывает оценки старших производных и является обобщением на случай $n > 2$ формулы из работы [7], ориентированной для схемы первого порядка точности. Число ξ_0 в (8) означает долю отрезка, относящуюся к ветви (6) составного преобразования отображаемую в погранслой. В случае равномерной сетки по ξ число ξ_0 означает долю числа шагов сетки, попадающих в погранслой. Числа n (порядок производной в главном члене погрешности аппроксимации схемы) и l (гладкость склейки) выбираются в зависимости от порядка аппроксимации схемы p (обычно $l = p$, $n = p + 1$).

Преобразование (8) является довольно универсальным для устранения экспоненциальных особенностей функции $u(x, \varepsilon)$ до порядка n, так как его константа $a \ge m > 0$ годится при любых значениях $a(0)$. Кроме того, преобразование (8) при некоторых ограничениях на константу a пригодно также и для устранения степенных сингулярностей.

3 Результаты численных исследований

Задача 1. Рассмотрим задачу с экспоненциальным слоем масштаба $k = 1$ в окрестности левой границы $x = 0$, который реализуется при всюду отрицательном $a(x)$, отделенном от нуля константой (см. [1] и [3]).

Таблица 1: Результаты численного решения задачи 1 по схемам $O(h^p)$ $(p = 1, 2, 3)$ при $\varepsilon = 0.001$

N	δ_1	p_1	δ_2	p_2	δ_3	p_3
80	2.40e-02	5.384	1.51e-02	6.053	1.62e-02	5.951
160	1.20e-02	0.991	3.44e-03	2.129	3.40e-03	2.250
320	6.10e-03	0.982	8.28e-04	2.056	1.18e-03	1.531
640	3.06e-03	0.998	2.05e-04	2.015	2.09e-04	2.489
1280	1.53e-03	0.997	5.11e-05	2.004	1.58e-05	3.728
2560	7.66e-04	0.999	1.28e-05	2.001	9.84e-07	4.005

Сетка определяется с помощью функции (8) при параметрах $k = 1$, $a = 2$, $\xi_0 = 1/2$. Краевая задача имеет вид

$$-\varepsilon u'' - (1 + x)u' + xu + \cos(4\pi x) = 0, \quad u(0) = 0, u(1) = 1.$$

В таблице 1 сравниваются расчеты по трем схемам на последовательности сеток при $\varepsilon = 1e-03$. В первой колонке приведено число шагов, затем пара колонок содержит апостериорные оценки C-нормы ошибки δ_1 и порядка точности p_1, вычисленные по результатам, полученным на двух последовательных сетках по схеме первого порядка аппроксимации. Следующие две пары колонок имеют тот же смысл, но в отношении схем второго и соответственно третьего порядков аппроксимации. На рисунке 1 приведены результаты расчетов

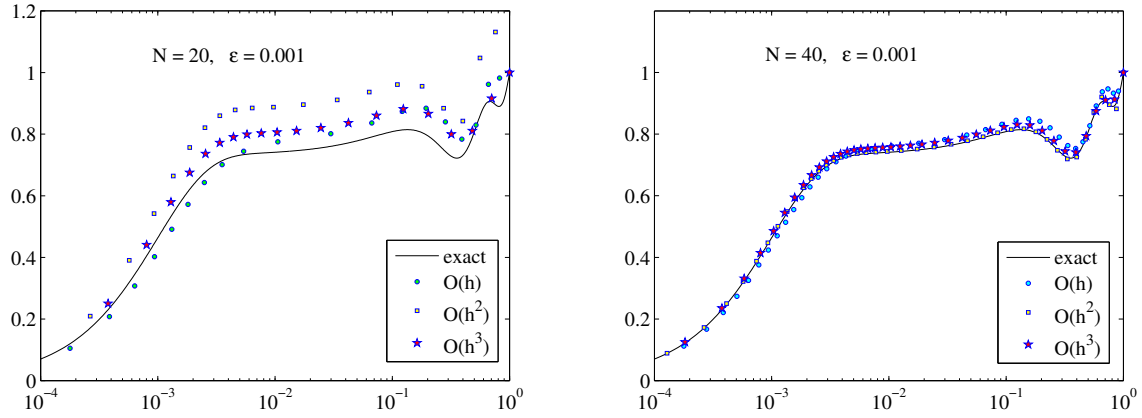


Рис. 1: Сравнение результатов расчетов при $N = 20$ (слева) и $N = 40$ (справа). Вязкость $\varepsilon = 0.001$.

задачи 1 на сетках с числом шагов $N = 20$ (слева) и $N = 40$ (справа) по схемам первого, второго и третьего порядков.

Как показывают результаты таблицы 1, на грубых сетках схема третьего порядка аппроксимации может давать не лучшую точность, но при детализации сетки расчеты по схеме третьего порядка становятся более точными. Это объясняется тем, что в главном члене погрешности порядок производной выше у схем более высокого порядка точности, а локальные значения производных высших порядков в погранслое могут быть существенно бóльшими, чем значения младших производных, и различие компенсируется только детализацией сетки за счет более высокого показателя степени в главном члене погрешности $O(h^p)$ схем повышенной точности. Именно это происходит на практике.

Заметим, что если погранслой расположен не на левой, а на правой границе (реализуется при отделенном от нуля положительном $a(x)$), то в качестве отображения, генерирующего сетку, можно взять функцию

$$y(\xi, \varepsilon, a, k, l, n, \xi_0) = 1 - x_b(1 - \xi, \varepsilon, a, k, l, n, \xi_0) \tag{9}$$

Задача 2. Рассмотрим случай, когда слои одинакового масштаба $k = 1$ расположены на концах отрезка $(0, 1)$. Такая конфигурация реализуется при $a(0) < 0$, $a(1) > 0$ и $a'(x_0) \geq 0$, если $a(x_0) = 0$. Например, это имеет место при возрастающей функции $a(x)$, меняющей знак внутри отрезка $(0, 1)$. Для расчета была взята следующая задача

$$-\varepsilon u'' + (x - \frac{1}{2})\, u' + u + \sin(2\pi x) \ = 0, \quad u(0) = \frac{1}{2}\,, \quad u(1) = 0.$$

Таблица 2: Результаты численного решения задачи 2 по схемам $O(h^p)$ $(p = 1, 2, 3)$ при $\varepsilon = 1e - 03$

N	δ_1	p_1	δ_2	p_2	δ_3	p_3
40	8.33e-02	2.815	9.39e-02	2.789	2.60e-01	1.013
80	6.08e-02	0.454	2.05e-02	2.192	3.75e-02	2.797
160	3.77e-02	0.687	5.04e-03	2.028	4.11e-03	3.188
320	2.15e-02	0.808	1.25e-03	2.010	2.63e-04	3.966
640	1.14e-02	0.915	3.12e-04	2.001	1.63e-05	4.011
1280	5.88e-03	0.960	7.81e-05	2.001	1.01e-06	4.004

Ввиду наличия в данной задаче двух пограничных слоев необходимо сгущать сетку на обоих концах отрезка.

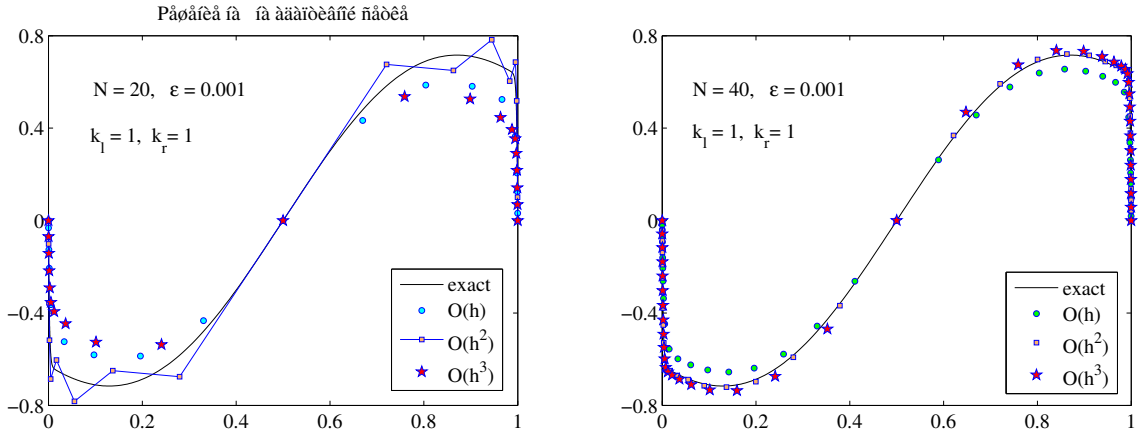


Рис. 2: Решение задачи 2 на сетке с $N = 20$ (слева) и $N = 40$ (справа) при $\varepsilon = 0.001$ разными схемами

Это можно осуществить как суперпозицией левостороннего и правостороннего преобразований (8), (9), так и отображением, склеенным из двух отображений, построенных на основе базовых. Последнее имеет вид

$$z(\xi) = \frac{1}{2}(1 + \bar{y}(\eta)), \quad \eta = 2\xi - 1, \quad 0 \le \xi \le 1,$$

где

$$\bar{y}(\eta) = \begin{cases} -y(-\eta, ...), & -1 \le \eta \le 0, \\ y(\eta, ...), & 0 \le \eta \le 1, \end{cases}$$

а y — отображение (9). В нашем примере параметры отображения $a = 2$, $\xi_0 = 0.5$. На рисунке 2 представлены результаты расчета задачи 2 по различным схемам при $\varepsilon = 0.001$. Результаты хорошо согласуются между собой, заметно некоторое размазывание решения монотонной схемой первого порядка точности.

Более подробное представление о результатах расчета задачи 2 на последовательности адаптивных сеток можно составить по приведенным в таблице 2 данным.

Задача 3. Для построения сеток для задач со степенными пограничными слоями можно использовать базовое преобразование (8), однако параметр a в (8) уже не будет произвольной константой, как это было в задаче с экспоненциальным погранслоем.

Пусть в нашем уравнении $F_u(x, u) \ge c > 0$, $a(0) = 0$, $a'(0) > 0$, $a(1) < 0$ и $a'(x_0) \ge 0$ для каждого x_0, для которого $a(x_0) = 0$, $0 < x_0 < 1$. В этом случае производные решения, согласно [1] и [3], оцениваются следующей формулой:

$$|u^{(p)}(x, \varepsilon)| \le M[1 + \varepsilon^{b/2}(\varepsilon^{1/2} + x)^{-b-p}], \quad p \le n, \quad 0 \le x \le 1, \tag{10}$$

где $0 < m \le b \le c/a'(0)$, т. е. решение имеет единственный степенной погранслой масштаба $1/2$ возле $x_0 = 0$. Поэтому для построения сетки, сгущающейся в слое, можно использовать преобразование $x_b(\xi, \varepsilon, a, 1/2, ...)$ из (8), в котором $0 < m \le a \le b/n^2$.

В качестве примера рассматривается следующая задача

$$-\varepsilon u'' - 8x(x-1/2)^3 u' + 16u - \exp(x) \ = 0, \quad u(0) = 0\ , \quad u(1) = 0.2\ .$$

Для этого примера $c = 16$, $a'(0) = 1$, поэтому можно положить

$$l = 2\ , \quad b = 16\ , \quad \xi_0 = 1/2\ , \quad a = 16/n^2\ , \quad \beta = 16/(n(16+n)).$$

Результаты расчета степенного слоя (задача 3) на сгущающихся равномерных и адаптивных сетках приведены в таблице 3.

Таблица 3: Результаты расчета степенного слоя масштаба 1/2 при $\varepsilon = 0.001$.

	равномерная сетка			адаптивная сетка		
N	δ_1	δ_2	δ_3	δ_1	δ_2	δ_3
20	9.87e-04	2.24e-03	1.40e-02	2.00e-03	7.07e-04	7.03e-03
40	2.87e-03	2.03e-03	8.06e-03	8.61e-04	3.62e-04	5.61e-03
80	3.09e-03	2.46e-03	2.15e-03	8.72e-04	3.33e-04	1.07e-03
160	1.88e-03	1.54e-03	2.87e-04	5.58e-04	7.83e-05	1.55e-04
320	6.00e-04	4.70e-04	1.99e-05	3.21e-04	2.30e-05	1.36e-05
640	1.97e-04	1.27e-04	1.29e-06	1.75e-04	6.53e-06	9.06e-07
1280	6.73e-05	3.24e-05	8.15e-08	9.13e-05	1.65e-06	7.02e-08

Заключение

Формат данной работы не позволяет привести более полно результаты многочисленных тестов, проведенных авторами, однако характер зависимостей ошибок от порядка точности схем во всех расчетах примерно такой же, как в приведенных тестах. Помимо рассмотренных здесь задач были были проведены расчеты экспоненциальных слоев с различными масштабами, внутренних слоев, логарифмических и смешанных [4]. Численные эксперименты свидетельствуют об эффективности симбиоза компактных схем и специальных адаптивных сеток, явно задаваемых на основе априорных оценок производных решения. В большинстве расчетов реальная точность тем выше, чем выше порядок точности схем, аномалии наблюдаются лишь при весьма грубых сетках. Это объясняется тем, что используемые оценки производных носят асимптотический характер при $h \to 0$, и рекомендации по распределению узлов сетки тем более точны, чем детальнее сетка. Кроме того, значения констант в оценках производных высших порядков неизвестны, в то время как они могут влиять на ошибку при крупном шаге сетки.

Расчеты различных типов слоев показывают, что хотя монотонная схема первого порядка работает исключительно надежно на любых сетках и при любых значениях ε, однако при недостаточной детализации сетки, когда аппроксимационная вязкость превышает физическую, результаты при фиксированной сетке перестают зависеть от уменьшающегося значения малого параметра. Это свойство вполне согласуется с опытом расчета течений вязкой несжимаемой жидкости с большими числами Рейнольдса — на фиксированной сетке схема первого порядка работает при сколь угодно малой вязкости, генерируя однако ошибочные результаты. В то же время схема второго порядка при превышении ограничения на разностное число Рейнольдса перестает работать ввиду переполнения. У схемы третьего порядка, обладающей некоторой аппроксимационной вязкостью, порог переполнения, в отличие от схемы первого порядка, имеется, но он выше, чем у схемы второго порядка.

Список литературы

[1] Лисейкин В. Д. Оценки производных решения дифференциальных уравнений с пограничными и внутренними слоями // Сибирский математический журнал. 1992. Т. 33, № 6, С. 106–117.

[2] Liseikin V. D. Grid Generation Methods// Springer, third ed., Berlin, 2017.

[3] Liseikin V. D. Layer Resolving Grids and Transformations for Singular Perturbation Problems// VSP, Utrecht, 2001.

[4] Лисейкин В. Д., Паасонен В. И. Компактные разностые схемы и адаптивные сетки для численного моделирования задач с пограничными и внутренними слоями // Сиб. Журн. Вычисл. математики. 2019. Т. 22, № 1, С. 41–56.

[5] Паасонен В. И. Схема третьего порядка аппроксимации на неравномерной сетке для уравнений Навье-Стокса// Вычислительные технологии. 2000. Т. 5, № 5, С. 78–85.

[6] Глуховский А. С., Паасонен В. И. Компактные разностные схемы для уравнений Навье-Стокса на неравномерных сетках // Марчуковские научные чтения 2017. Труды международной конференции «Вычислительная и прикладная математика 2017». Изд. ИВМиМГ. Новосибирск. 25 — 30 июня 2017. С. 211–217.

[7] Лисейкин В. Д. О численном решении уравнений со степенным пограничным слоем// Журн. Вычисл. матем. и матем. физики. 1986. Т. 26, № 12, С. 1813–1820.

[8] Лисейкин В. Д. О численном решении сингулярно возмущенных уравнений с точками поворота // Журн. Вычисл. матем. и матем. физики. 1984. Т 24, № 12, С. 1812–1818.

Лисейкин Владимир Дмитриевич —д.ф.-м.н., вед. науч.сотр. Института вычислительных технологий СО РАН;
e-mail: lis@ict.nsc.ru;
Паасонен Виктор Иванович — к.ф.-м.н., ст. науч.сотр. Института вычислительных технологий СО РАН;
e-mail: paas@ict.nsc.ru;
Дата поступления — 30 апреля 2019 г.

РЕШЕНИЕ МЕТОДОМ МОНТЕ–КАРЛО НЕЛИНЕЙНОГО УРАВНЕНИЯ ШРЕДИНГЕРА

В. Л. Лукинов

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск
Новосибирский государственный университет, 630090, Новосибирск
Сибирский государственный университет телекоммуникаций и информатики, 630102, Новосибирск

УДК 519.633
DOI: 10.24411/9999-016A-2019-10046

Проведено сравнение численных решений нелинейного уравнения Шредингера путем применения нелинейных аналогов прямого (FNFT) и обратного преобразования (BNFT) Фурье и прямой схемы Эйлера. При FNFT происходит разложение спектра поступающего сигнала путём решения уравнения Захарова–Шабата. Распространение нелинейной части спектра описывается известным уравнением. Для нахождения получаемого сигнала при BNFT применяются разработанные автором методы Монте–Карло для численного решения интегрального уравнения Гельфанда-Левитана-Марченко. Для оценки влияния аддитивного гауссовского белого шума использовалось численное решение на основе схемы Эйлера. В обоих численных схемах моделировался случайный гауссовский белый шум источника. Проведено моделирование потери мощности сигнала при прохождении случайной гетерогенной среды и изучено влияние шумов, возникающих при максимальных нагрузках оптического волокна. Проведено моделирование взаимодействия солитонов, распространяющихся навстречу друг другу с разных концов оптоволоконной среды. Для моделирования использовалось сочетание численной схемы на основе FNFT и BNFT до момента взаимодействия и неявной безусловно-устойчивой схемы Кранка-Николсона для решения дискретной дифференциальной задачи на равномерных сетках по пространству и времени.
Ключевые слова: статистическое моделирование, нелинейное уравнений Шредингера.

Введение

Данная работа посвящена изучению влияния случайных шумов при распространении и взаимодействии оптических сигналов в оптоволоконных каналах. На пропускную способность траффика нелинейных каналов оказывают влияние шумы с различного рода источниками: обратное рэлеевское рассеяние, спонтанное рамановское рассеяние и в усилителях вследствие спонтанной эмиссии фотонов. Оптический шум совместно с дисперсией и нелинейностью это три ключевых физических эффекта, влияющих на распространение оптических сигналов в оптоволоконных каналах, которые точно описываются нелинейным уравнением Шредингера (NLSE), учитывающим непрерывное взаимодействие между дисперсией и нелинейностью [1]. Известно, что NLSE (без возмущения) принадлежит классу интегрируемых нелинейных систем [2]. В частности, это обосновывает применение нелинейных аналогов преобразования Фурье (NFT) для перехода от пространственно-временных к спектрально-временным переменным. В настоящее время можно выделить два подхода: нелинейное частотное разделение каналов (NFDM) и нелинейный обратный синтез (NIS), использующие для передачи сигналов дискретную (солитонную) и непрерывные части нелинейного спектра [3, 4].

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-01-00698 А) и Института вычислительной математики и математической геофизики СО РАН (гос. задание 0315-2016-0002).

ISBN 978-5-901548-42-4

1 Основные нелинейные преобразования в оптической связи

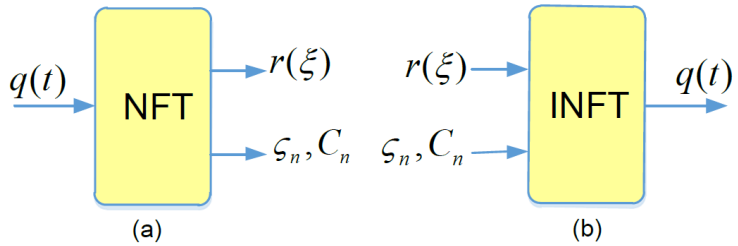


Рис. 1: Схема прямого NFT (a) и обратного INFT (b) для входного потенциала $q(t)$, стремящегося к 0 при $t \to \pm\infty$

В передатчике передаваемая последовательность двоичных данных сначала кодируется цифровым способом в комплексную форму волны $s(t)$, используя произвольный формат модуляции и метод кодирования. Линейный спектр Фурье этой кодированной комплексной формы сигнала определяется следующим образом

$$S(\omega) = \int_{-\infty}^{\infty} s(t) \exp(-j\omega t) dt. \tag{1}$$

После кодирования линейный спектр Фурье кодированной комплексной формы волны $S(\omega)$ отображается на непрерывную часть нелинейного спектра комплексного сигнала $q(t)$ с использованием обратного нелинейного преобразования Фурье BNFT (то-есть путем решения уравнения GLME). Обозначим через $r(\xi)$ непрерывную часть нелинейной спектра $q(t)$. Тогда операция отображения в блоке BNFT может быть выражена как:

$$r(\xi)|_{\xi=-\omega/2} = -S(\omega)$$

Комплексный сигнал, $q(t)$, затем подается в IQ модулятор для прямого преобразования в оптическую среду и передается в оптоволокно.

BNFT блок отображает закодированную информацию на непрерывную часть нелинейного спектра сложного сигнала. Комплексный передаваемый сигнал (выход BNFT блока) в этом случае не содержит солитонов, иначе говоря, нелинейный спектр не содержит дискретной части. На приеме реальные и мнимые части передаваемого сигнала обнаруживаются с помощью когерентного приемника. Нелинейный спектр принятого сигнала получается прямым нелинейным преобразованием Фурье NFT (т. е. путем решения ZSP).

При распространении по каналу оптического волокна без потерь (например, с распределенным Рамановским усилением), эволюция $r(\xi)$ тривиальна [1]:

$$r(L, \xi) = r(\xi) \cdot e^{-2j\xi^2 L},$$

где L — расстояние передачи.

Восстановление линейного спектра Фурье исходного кодированного комплексного сигнала после декодера возможно путем применения одноступенчатого линейного дисперсионного удаления:

$$\bar{S}(\omega) = -r(L, \xi) \cdot e^{2j\xi^2 L}\Big|_{\xi=-\omega/2}$$

После комплексный сигнал $s(t)$ можно восстановить используя операцию IFFT, наконец, сигнал можно подавать в стандартный декодер для обнаружения данных.

2 Нелинейное уравнение Шредингера (NLS)

Распространение комплексной медленно изменяющейся огибающей оптического поля $q(z, t)$ вдоль оптоволокна без потерь можно описать нелинейным уравнением Шредингера

$$iq_z - \frac{\beta_2}{2} q_{tt} + \gamma q|q|^2 = 0 \tag{2}$$

где z — расстояние на которое распространяется сигнал, t — время распространения сигнала. Для определенности будем рассматривать NLSE фокусирующего типа, при котором постоянный коэффициент хроматической дисперсии $\beta_2 < 0$ в уравнении (2). Дисперсионные члены более высокого порядка здесь не рассматриваются. Мгновенный коэффициент нелинейности Керра γ равен

$$\gamma = n_2\omega_0/cA_{em}, \tag{3}$$

где n_2 — показатель преломления, A_{em} — эффективная область моды, c — скорость света в вакууме, ω_0 — угловая несущая частота. Здесь рассматривается случай, когда постоянный коэффициент хроматической дисперсии $\beta_2 < 0$ в уравнении (2), то-есть так называемого фокусирующиюго типа NLSE. Дисперсионные члены более высокого порядка здесь не рассматриваются. Мгновенный коэффициент нелинейности Керра γ равен

$$\gamma = n_2\omega_0/cA_{em}, \tag{4}$$

где n_2 — показатель преломления, A_{em} — эффективная область моды, c — вакуумная скорость света и ω_0 — угловая несущая частота.

В дальнейшем NLSE будет рассматриваться в нормализованном виде:

$$iq_z - q_{tt} + q|q|^2 = 0, \tag{5}$$

который может быть получен с помощью следующей замены:

$$\frac{t}{T_s} \longrightarrow t, \qquad \frac{z}{Z_s} \longrightarrow z, \qquad \sqrt{\gamma Z_s} \longrightarrow q \tag{6}$$

где T_s — свободный параметр (например, характерный временной масштаб входного сигнала) и связанная пространственная шкала равна $Z_s = T_s^2\beta_2$. Заметим, что все величины q, t и z в нормированном уравнении теперь безразмерны.

2.1 Прямое нелинейное преобразование Фурье

Подобно прямому FT, целью FNFT является разложение сигнала в IST спектральные данные. Это преобразование производится путем решения задачи Захарова–Шабата (ZSP) [2]. Последнее соответствует задаче рассеяния для неэрмитовой (в случае аномальной дисперсии) системы уравнений Дирака для двух вспомогательных функций $v_1(t)$, $v_2(t)$ с профилем входного импульса NLSE $q(0,t) \equiv q(t)$, являющимся эффективным потенциалом в уравнениях:

$$\frac{dv_1}{dt} = q(t)v_2 - i\zeta v_1, \quad \frac{dv_2}{dt} = -\bar{q}(t)v_1 + i\zeta v_2. \tag{7}$$

Здесь ζ является (как правило, комплексным) собственным значением, $\zeta = \xi + i\eta$, $\bar{q}(t)$ является комплексным сопряжением потенциала $q(t)$, который считается распадающимся при $t \longrightarrow \pm\infty$ (при выполнении определенных условий, наложенных на скорость затухания, см. [2]).

Для определения непрерывной части нелинейного спектра (вещественных $\zeta = \xi$), фиксируются два линейно-независимых решения уравнения Йоста (7):

$$\Phi(t,\xi) = [\phi_1, \phi_2]^T, \quad \widetilde{\Phi}(t,\xi) = [\bar{\phi}_2, -\bar{\phi}_1]^T, \tag{8}$$

с “начальным” состоянием слева:

$$\Phi(t,\xi)|_{t\to-\infty} = [e^{-i\xi t}, 0]^T \tag{9}$$

Таким же образом, справа фиксируются два других решения Йоста

$$\Psi(t,\xi) = [\psi_1, \psi_2]^T, \quad \widetilde{\Psi}(t,\xi) = [\bar{\psi}_2, -\bar{\psi}_1]^T, \tag{10}$$

$$\Psi(t,\xi)|_{t\to+\infty} = [0, e^{i\xi t}]^T \tag{11}$$

Эти два набора решений линейно зависимы и могут быть выражены через коэффициенты рассеяния Йоста $a(\xi)$ и $b(\xi)$:

$$\begin{aligned} \Phi(t,\xi) &= a(\xi)\widetilde{\Psi}(t,\xi) + b(\xi)\Psi(t,\xi) \\ \widetilde{\Phi}(t,\xi) &= -\bar{a}(\xi)\Psi(t,\xi) + \bar{b}(\xi)\widetilde{\Psi}(t,\xi) \end{aligned} \tag{12}$$

Коэффициент отражения, производящий непрерывную часть нелинейного спектра, определяется как:

$$r(\xi) = \bar{b}(\xi)/a(\xi) \tag{13}$$

Солитоны соответствуют комплексным собственным значениям ζ_n, где $a(\zeta_n) = 0$. Прямое NFT отображает начальное поле $q(0,t)$ на множество данных рассеяния $\sum = [(r(\xi),\ \xi$ — вещественное); $(\zeta_n,\ \gamma_n = (b(\zeta_n)a'_\zeta(\zeta_n))^{-1})]$, где индекс n пробегает все дискретные собственные значения ZSP (дискретная недисперсионная часть нелинейного спектра). В этом случае комплекснозначная функция $r(\xi)$ вещественного аргумента ξ (нелинейный спектр) аналогична обычному спектру Фурье. Поэтому в рамках метода NIS эта непрерывная часть нелинейного спектра, $r(\xi)$ используется для кодирования и передачи информации, а ξ играет роль частоты. Нелинейная спектральная (NS) функция, определяемая как:

$$N(\omega) = -r(\xi)|_{\xi=-\omega/2} \tag{14}$$

служит прямым нелинейным аналогом спектра Фурье, стремясь к обычному преобразованию Фурье $q(0,t)$ в линейном пределе.

2.2 Обратное нелинейное преобразование Фурье

BNFT отображает рассеянные данные $\sum$ на поле $q(t)$. Данное отображение можно получить путем решения уравнения Гельфанда-Левитана-Марченко (GLME) для неизвестной функции $K(t,x)$ [5–7, 9]. GLME без солитонов можно записать в виде:

$$K(t,x) + F(t+x) + \int_{-\infty}^{t} ds \int_{-\infty}^{t} dr \quad K(t,r)\bar{F}(s+y)F(r+s) = 0. \tag{15}$$

Здесь $F(x)$ является линейным FT для $r(\xi)$:

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} r(\xi) e^{-i\xi x} d\xi \tag{16}$$

Решая уравнение (26) для K (t,x), обратное NFT можно получить в виде

$$q(t) = 2 \lim_{x \to t-0} K(t,x) \tag{17}$$

3 Численные методы

3.1 Вычисление непрерывного спектра

Непрерывный спектр (т.е. $a(\xi)$, $b(\xi)$ и $r(\xi)$) может быть вычислен путем непосредственного интегрирования ZSP (7), а затем оценивая пределы для соответствующих компонентов функции Йоста как:

$$\begin{aligned} a(\xi) &= \lim_{t \to +\infty} \phi_1(t,\xi) e^{i\xi t} \\ b(\xi) &= \lim_{t \to +\infty} \phi_2(t,\xi) e^{-i\xi t} \end{aligned} \tag{18}$$

Для решения уравнения ZSP можно использовать легко распараллеливающийся метод PCA (7). При этом потенциал $q(t)$ отсекается вне диапазона $(-T_0; T_0)$. Внутри рассматриваемой области $q(t)$ берется равным константе $q_n = q(t_n)$ на каждом элементарном подинтервале $(t_n - \Delta t/2; t_n + \Delta t/2)$, где $t_n = -T_0 + n\Delta t$, $\Delta t = T_0/M$ — это шаг по времени и $2M+1$ — общее число разбиений рассматриваемой области. Идея метода PCA основана на том, что уравнение (7) может быть решено точно внутри каждого элементарного интервала сетки для произвольного значения спектрального параметра ξ:

$$\Phi(t_n + \Delta t/2, \xi) = T(q_n, \xi)\Phi(t_n - \Delta t/2, \xi), \tag{19}$$

где матрица перехода $T(q_n, \xi)$ задается в виде

$$T(q_n, \xi) = \exp\left[\Delta t \begin{pmatrix} -i\xi & q_n \\ -\bar{q}_n & i\xi \end{pmatrix}\right] \tag{20}$$

Задачу рассеяния можно решить путем "продолжения" решения итеративно, начиная от $-T_0$ до правой границы T_0, используя набор матриц переноса $T(q_n, \xi)$ заданных уравнением (20). Итоговый результат может быть выражен в виде

$$\begin{array}{l} \Phi(T_0 - \Delta t/2, \xi) = \prod(\xi)\Phi(-T_0 + \Delta t/2, \xi) \\ \prod(\xi) = \prod\limits_{n=1}^{2M} T(q_n, \xi) \end{array} \tag{21}$$

Начальное условие, определенное на правом конце временного интервала, может быть представлено в ввиде

$$\Phi(-T_0 - \Delta t/2, \xi) = (1, 0)^T e^{-i\xi(-T_0 - \Delta t/2)}, \tag{22}$$

на левом конце интервала имеет место следующее равенство

$$\Phi(T_0 - \Delta t/2, \xi) = \begin{array}{l} a(\xi)e^{-i\xi(T_0 - \Delta t/2)} \\ b(\xi)e^{i\xi(T_0 - \Delta t/2)}, \end{array} \tag{23}$$

Коэффициенты Йоста могут быть выражены в виде

$$\begin{array}{l} a(\xi) = \prod_{11}(\xi)e^{2i\xi T_0} \\ b(\xi) = \prod_{21}(\xi)e^{-i\xi\Delta t}, \end{array} \tag{24}$$

где $\prod_{11}(\xi)$ и $\prod_{21}(\xi)$ — соответствуюшие компоненты матрицы $\prod(\xi)$. В общем случае, когда потенциал q(t) обрывается вне интервала (T_{min}, T_{max}) с произвольными границами, выражение (24) может быть представлено в виде

$$\begin{array}{l} a(\xi) = \prod_{11}(\xi)e^{i\xi(T_{max} - T_{min})} \\ b(\xi) = \prod_{21}(\xi)e^{-2i\xi(T_{max} + T_{min} - \Delta t)}, \end{array} \tag{25}$$

Из (21) видно, что можно вычислить матрицы переноса $T(q_n,\ xi)$ независимо друг от друга. В результате алгоритм PCA может быть легко реализован параллельно, что позволяет существенно сократить время вычислений.

В качестве примера рассмотрим прямоугольный импульс

$$q(t) = \begin{cases} A, t \in [T_1, T_2] \\ 0, \end{cases}$$

Непрерывный спектр этого прямоугольного импульса определяется следующим образом:

$$r(\xi) = \frac{\bar{A}}{j\xi} e^{-2j\xi t} \left(1 - \frac{\sqrt{\xi^2 + |A|^2}}{j\xi} \cot\left(\sqrt{\xi^2 + |A|^2}(T_2 - T_1) \right) \right)$$

Определим нормированную среднеквадратичную ошибку (RMSE) для PCA следующим образом:

$$RMSE = \frac{1}{A} \sqrt{\frac{1}{N} \sum_{k=1}^{N} |r_{exact}(\xi_k) - r_{numeric}(\xi_k)|^2}$$

3.2 Вычисление BNFT

Для удобства в численных расчетах GLME можно переписать в виде двух связанных интегральных уравнений:

$$\begin{array}{l} A_1(x,t) + \int\limits_{-\infty}^{x} F(t+x)\bar{A}_2(x,y)dy = 0 \\ A_2(x,t) - \int\limits_{-\infty}^{x} F(t+x)\bar{A}_1(x,y)dy = F(x+t), x > t, \end{array} \tag{26}$$

где $F(t)$ это обратное линейное преобразование Фурье для $r(\xi)$. Для численного анализа далее используется замена

$$u(x,s) = \bar{A}_1(x, x-s), \quad v(x,\tau) = \bar{A}_2(x, \tau - x), \tag{27}$$

так, что GLME (26) можно переписать в виде

$$\begin{aligned} u(x,s) + \int_{\tau}^{2x} F(t-s)v(x,\tau)d\tau = 0 \\ v(x,\tau) - \int_{0}^{s} F(t-s)u(x,s)ds = F(\tau) \end{aligned} \tag{28}$$

Функции $u(x,s)$ и $v(x,s)$ определены внутри интервала $0 \le \tau \le 2x \le 2T_0$. BNFT отображение $r(\xi)$ на временную область задается формулой

$$q(x) = 2v(x, 2x-0) \tag{29}$$

Следуя процедуре дискретизации, представленной в [10], разделим интервал $0 \le \tau \le 2T_0$, где функция $F(\tau)$ известна, в отрезки длины $h = 2T_0/N$. Дискретные переменные τ_n, s_k и x_m определяются по формулам

$$\begin{array}{ll} s_k = h(k-1)/2, & k = 1,2,\ldots,m \\ \tau_n = h(n-1)/2, & n = 1,2,\ldots,m \\ x_m = mh/2, & m = 1,2,\ldots,N \end{array} \tag{30}$$

Определим сеточные функции

$$u_n^{(m)} = u(x_n, \tau_m), \quad v_n^{(m)} = v(x_n, \tau_m), \quad F_n = F(nh) \tag{31}$$

Используя прямоугольную квадратурную схему для приближения интегралов в уравнении (28), получаем следующую дискретную форму GLME:

$$\begin{aligned} u_k^{(m)} + h \sum_{n=k}^{m} \bar{F}_{n-k} v_n^{(m)} = 0 \\ v_n^{(m)} - h \sum_{k=1}^{m} F_{n-k} u_k^{(m)} = F_n \end{aligned} \tag{32}$$

m-ый сеточный элемент BNFT (во временной области) затем задается следующим образом:

$$q^{(m)} = v_m^{(m)} \tag{33}$$

Уравнения (32) могут быть записаны в матричной форме

$$\mathbf{G}^m \begin{pmatrix} u^{(m)} \\ v^{(m)} \end{pmatrix} = \mathbf{b}^{(m)} \tag{34}$$

где $\mathbf{b}^{(m)}$ формируется из нулевого вектора размерности m и вектора размерности m с компонентами F_n; $\mathbf{G}^m$ в (34) является квадратной матрицей размерностей $2m \times 2m$, которая имеет следующий вид:

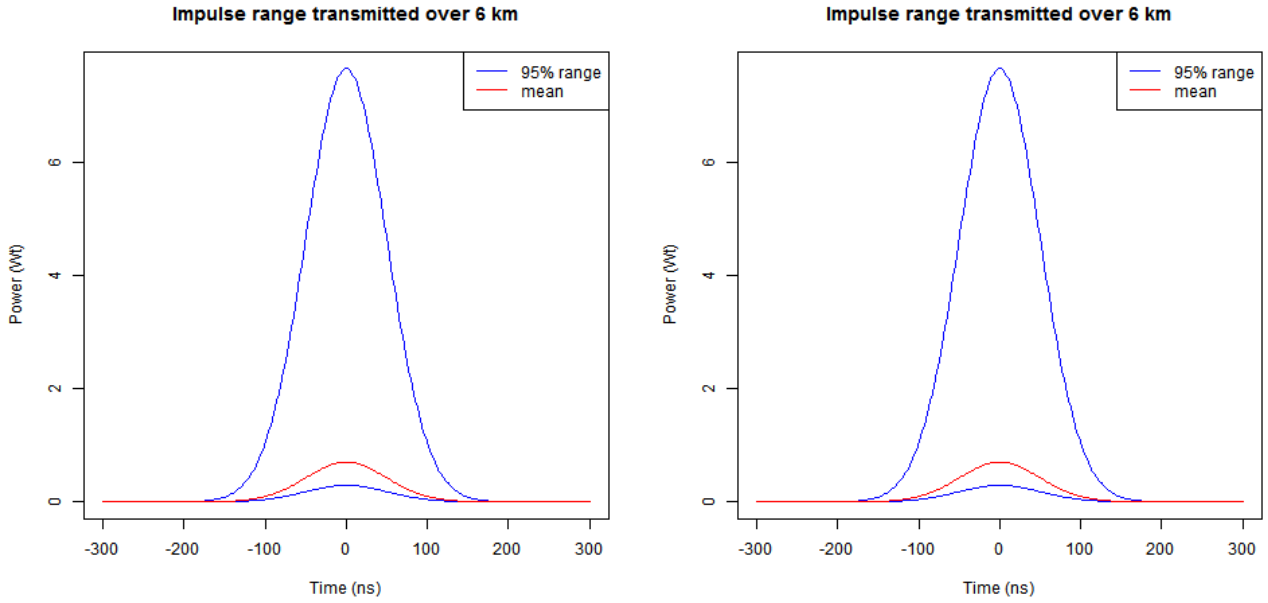
$$\mathbf{G}^m = \begin{pmatrix} \mathbf{E}^{(m)} & h\bar{\mathbf{F}}^{(m)} \\ -h\mathbf{F}^{(m)} & \mathbf{E}^{(m)} \end{pmatrix} \tag{35}$$

4 Результаты

Целью работы является изучение нелинейных эффектов при распространении и взаимодействии лазерных импульсов в длинном оптическом волокне, возникающих из-за модуляции неустойчивости. Распространение длинных импульсов в оптическом волокне описывается NLS [1]:

$$\frac{\partial A}{\partial z} + \frac{1}{c}\frac{\partial A}{\partial t} = i\frac{|\beta_2|}{2}\frac{\partial^2}{\partial t^2}A + i\gamma|A|^2A - \alpha A, \tag{36}$$

где $A(t,z)$ — амплитуда электромагнитного поля на несущей частоте $\omega_0 = 2\pi n_0/\lambda_0$, амплитуда нормируется степенью $P = |A|^2$; c — скорость света в оптическом волокне с показателем преломления n_0; $\beta_2 = -10\mathrm{km}^{-1} \cdot \mathrm{nm}^{-2}$ — дисперсия на длине волны 1550nm; $\gamma = 3\mathrm{W}^{-1} \cdot \mathrm{km}^{-1}$ — коэффициент нелинейности;



$\alpha = 0.17\text{dB/km}$ — коэффициент оптических потерь. Граничное условие на входе в оптическое волокно дается как сумма гауссовского импульса с шириной в половине максимума $T_0 = 100\text{ns}$ и амплитудой шума $A_n(t)$:

$$A(t,0) = \sqrt{P_0}\exp\left(-\frac{t^2}{T_0^2}2\ln 2\right)\sqrt{\frac{4\ln 2}{\pi}} + A_n(t).$$

Здесь $P_0 = \varepsilon_0/T_0$ это средняя мощность энергии импульса ε_0. Случайный шум имеет ненулевую спектральную плотность мощности $P_n(\omega) = |A_n(\omega)|^2$ ($P_n(\omega) = \varepsilon_n/\Delta$) в спектральной области для $\Delta = 1$ nm относительно центральной несущей частоты ω_0; ε — средняя энергия шума за период $\tau = 2\text{ms}$ Отношение энергий импульса и шума $\varepsilon_0/\varepsilon_n$ принималось равным 0.7 [3]

Для численного решения уравнения (36) использовалась описанная выше модифицированная NIS схема со временем $T = 1\text{ms}$, при этом количество точек разбиения бралось равным $N = 2^{20}$. Пространственный шаг полагался равным $\Delta z = 0.1\text{mm}$. Начальное распределение было задано гауссовским импульсом с длительность в половине максимальной интенсивности 100ns и переменной пиковой мощности P_p, а именно $A(0,t) = \sqrt{P_p}\exp(-t^2/2T_0^2)$, и к нему добавлен белый гауссовский шум со спектральной полосой 1nm. Количество смоделированных импульсов составляло 10^6.

Список литературы

[1] G. Agrawal, Nonlinear fiber optics, Academic Press, New York, 1996

[2] В. Е. Захаров и А.Б. Шабат, Точная теория двумерной самофокусировки и одномерной самомодуляции волн в нелинейных средах // Сов. Физ. 34, 6269, 1972.

[3] S.T. Lee, J.E. Prilepsky, S.K. Turitsyn, Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers // OPTICS EXPRESS, Vol. 22, No 22, 2014.

[4] Z. Dong, S. Hari, G. Tao, Z. Kangping, M. I. Yousefi, L. Chao, et al., Nonlinear Frequency Division Multiplexed Transmissions Based on NFT // PTL, IEEE, vol. 27, pp. 1621–1623, 2015.

[5] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, The inverse scattering transform-Fourier analysis for nonlinear problems // Stud. Appl. Math. 53, 249315, 1974.

[6] V. E. Zakharov, S. V. Manakov, S. P. Novikov, and L. P. Pitaevskii, Theory of Solitons. The Inverse Scattering Method. Colsultants Bureau, New York, 1984.

[7] M. J. Ablowitz and H. Segur, Solitons and the Inverse Scattering Transform. SIAM, Philadelphia, 1981.

[8] Crank, J.; Nicolson, P., "A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type" // Proc. Camb. Phil. Soc. 43 (1): 50–67, 1947. doi:10.1017/S0305004100023197

[9] A. C. Newell, Solitons in mathematics and physics. *SIAM, Philadelphia*, 1985.

[10] O. V. Belai, L. L. Frumin, E. V. Podivilov, and D. A. Shapiro, "Efficient numerical method of the fiber Bragg grating synthesis", // J. Opt. Soc. Am. B 24(7), 1451–1457, 2007.

Лукинов Виталий Леонидович — к.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: vitaliy.lukinov@sscc.ru.
Дата поступления — 30 апреля 2019 г.

ЧИСЛЕННО УСТОЙЧИВЫЙ ВЕРОЯТНОСТНЫЙ КЛАССИФИКАТОР ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ

В. Л. Лукинов

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск
Новосибирский государственный университет, 630090, Новосибирск
Сибирский государственный университет телекоммуникаций и информатики, 630102, Новосибирск

УДК 519.254
DOI: 10.24411/9999-016A-2019-10047

При обработке реальных данных методом логистической регрессии требуется определить и устранить негативное влияние на качество оценок коэффициентов модели, связанное с неполнотой данных, существованием "выбросов" в данных и особых ковариантов, коллинеарностью, критериями отбора ковариантов в многофакторную модель. В данной работе предложен и реализован алгоритм, позволяющий контролировать и устранить перечисленные выше негативные факторы. Другой задачей является разработка параллельных алгоритмов расчета коэффициентов логистической регрессии и характеристик качества. При этом необходимо решить оптимизационную задачу, реализующую критерий максимального правдоподобия. Известный подход, основанный на итерационном методе Ньютона–Рафсона, является численно неустойчивым и может привести к неправильному нахождению коэффициентов регрессии [9]. В данной работе предлагается новый параллельный численно-устойчивый итерационный алгоритм решения задачи минимизации на основе случайного поиска. Проведено сравнение с базовым алгоритмом *glm* из языка статистической обработки данных R [8].
Ключевые слова: задача классификации, бинарная логистическая регрессия.

Введение

Одним из наиболее распространенных и широко используемых в статистическом анализе больших данных является разработанный в прошлом веке метод логистической регрессии [1]. В биологии и медицине логистическая регрессия применяется в самых разных областях: выявлении и исследовании степени влияния предикторов заболеваний и послеоперационных осложнений, фундаментальном для ретроспективных групповых сравнительных исследований методе Propensity Score Matching, методах автоматического и полуавтоматического распознавания медицинских изображений, методах статистического кластерного анализа [2–5].

1 Задача бинарной логистической регрессии

Модель логистической регрессии строится на обучающей выборке, определяемой парами вида

$$(\mathbf{x}_i, y_i), \quad i = 1, 2, \ldots, n, \tag{1}$$

где

$$\mathbf{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \cdots \\ x_{im} \end{pmatrix} = \begin{pmatrix} 1 \\ x_{i2} \\ \cdots \\ x_{im} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \cdots \\ \mathbf{x}_n^T \end{pmatrix} = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1m} \\ 1 & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n2} & \cdots & x_{nm} \end{pmatrix} \tag{2}$$

является набором значений i-го объекта (значения ковариант), бинарные переменные y_i определяют принадлежность одному из двух классов, например, $y_i == -1$ — для первого класса и $y_i == 1$ — для второго класса, $m - 1$ — количество ковариант (признаков) у каждого объекта, n — число наблюдений.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-01-00698 А) и Института вычислительной математики и математической геофизики СО РАН (гос. задание 0315-2016-0002).

ISBN 978-5-901548-42-4

Конечной целью является построение апостериорной функции принадлежности произвольного объекта x классу $y == -1$, называемой функцией логистической регрессии, имеющей вид [1]

$$p(x) = \frac{1}{1 + \exp(-\mathbf{b}^T \mathbf{x})},$$

и принимающей значения в интервале $(0; 1)$. Принадлежность объекта x классу $y == 1$ определяется через дополнение вероятности как $1 - p(x)$. Объекты с полученным значением $p(x) > 0.5$ относят к первому классу. Оптимальные значения отсечения $p(x)$ для максимизации чувствительности и специфичности классификации на априорных данных можно получить методами ROC анализа [6].

Известно [7], что поиск функции логистической регресии можно представить в виде задачи минимизации функционала

$$Q(\mathbf{b}) = \sum_{\mathbf{i=1}}^{\mathbf{n}} \ln\left(\mathbf{1} + \mathbf{e}^{-\mathbf{y_i}\mathbf{b^T}\mathbf{x_i}}\right) \to \min_{\mathbf{b}\in\mathbf{R^m}}, \tag{3}$$

где $\mathbf{b^T} = (b_1, b_2, \ldots, b_m)$ — искомый вектор коэффициентов регрессии, представляющий границу классификации объектов в виде гиперплоскости

$$b_1 + \sum_{i=2}^{m} b_i x_i = 0$$

2 Итерационный метод поиска минимума

По причине отсутствия аналитического решения задачи (3), для численного поиска коэффициентов регрессии $\mathbf{b}$ применяются итерационные методы спуска для поиска экстремума. Многие программные пакеты (в том числе и метод соответствия обобщенных линейных моделей *glm* базового пакета *stats* в языке R [8]) использует метод Ньютона–Рафсона. Далее приводится описание основных шагов этого алгоритма. В качестве нулевого приближения берется "наивное" решение задачи классификации методом многомерной линейной регрессии:

$$\mathbf{b^{(0)}} = \left(\mathbf{X^T X}\right)^{-1} \left(\mathbf{X^T Y}\right). \tag{4}$$

На k-ом итерационном шаге происходит уточнение вектора коэффициентов $\mathbf{b^{(k)}}$:

$$\mathbf{b^{(k)}} = \mathbf{b^{(k-1)}} - h_{\mathbf{k}} \left(Q''(\mathbf{b^{(k-1)}})\right)^{-1} Q'(\mathbf{b^{(k-1)}}), \tag{5}$$

где $Q'(\mathbf{b^{(k-1)}})$ — градиент функционала Q в точке $\mathbf{b^{(k-1)}}$; $Q''(\mathbf{b^{(k-1)}})$ — матрица вторых производных функционала Q в точке $\mathbf{b^{(k-1)}}$; h_k — шаг итерации, который может, в тривиальном случае, быть единичным, "оптимальным" фиксированным или изменяться на каждой итерации [7].

Для исследования свойств алгоритма (5) выразим градиент функционала $Q'(\mathbf{b})$ и матрицу вторых производных $Q''(\mathbf{b})$. Обозначим значение так называемой сигмоидной функции как $\sigma = \sigma_i = \sigma_i(\mathbf{y}_i \mathbf{b}^T \mathbf{x}_i) = 1/(1 + \exp(-\mathbf{y}_i \mathbf{b}^T \mathbf{x}_i))$. Заметим, что для производной сигмоидной функции справедливо следующее соотношение $\sigma_i'(b) = \sigma_i(b)(1 - \sigma_i(b))$. Выразим первые частные производные функционала $Q(\mathbf{b})$:

$$\frac{\partial Q(\mathbf{b})}{\partial b_k} = -\sum_{i=1}^{n} (1 - \sigma_i) y_i x_k, \qquad k = 1, \ldots, m \tag{6}$$

Вторые частные производные функционала $Q(\mathbf{b})$:

$$\frac{\partial^2 Q(\mathbf{b})}{\partial b_j \partial b_k} = -\frac{\partial}{\partial b_j} \sum_{i=1}^{n} (1 - \sigma_i) y_i x_k = \sum_{i=1}^{n} \sigma_i (1 - \sigma_i) x_j x_k \qquad j = 1, \ldots, m, \qquad k = 1, \ldots, m \tag{7}$$

Положим:

$\mathbf{W}_{n\times n} = \operatorname{diag}\left\{\sqrt{\sigma_i(1 - \sigma_i)}\right\}$ — диагональная матрица весов наблюдений;

$\widetilde{\mathbf{X}}_{n\times m} = \mathbf{W}_{n\times n}\mathbf{X}_{n\times m}$ — матрица взвешенных ковариант;

$\widetilde{y}_i = y_i\sqrt{(1-\sigma_i)/\sigma_i}$, $\widetilde{y} = (\widetilde{y}_i)_{i=1,\ldots,n}$ — взвешенные отклики наблюдений.

Тогда произведение обратной матрицы вторых производных на вектор градиента из (5) в новых обозначениях можно записать в виде:

$$\left(Q''(\mathbf{b})\right)^{-1} Q'(\mathbf{b}) = -\left(\mathbf{X}^T\mathbf{W}^2\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}\widetilde{y} = -\left(\widetilde{\mathbf{X}}^T\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}^T\widetilde{y} \tag{8}$$

3 Причины неустойчивости и медленной сходимости

Существует несколько причин возможной численной неустойчивости и медленной сходимости метода Ньютона–Рафсона (4), (5). В случае медленной сходимости, итерации в программе могут прерваться из-за принудительного ограничения количества итераций и не достичь минимума функционала с заданной точностью. Рассмотрим некоторые причины.

Коррелируемость ковариант. Наихудшим случаем является линейная зависимость признаков, тогда матрица $Q''(\mathbf{b}) = \widetilde{\mathbf{X}}^T\widetilde{\mathbf{X}}$ является вырожденной и не может быть обратимой. В статистических расчетах на реальных данных такая ситуация редко возникает при формальных расчетах без исследования свойств наблюдений. Намного чаще возможен случай коррелируемых или так называемые мультиколлинеарных ковариант, при котором матрица $Q''(\mathbf{b})$ обратима, но близка к вырожденной. Столбцы такой матрицы приближенно линейно зависимы в пределах малой точности, а сама матрица является плохо обусловленной. Причиной плохой обусловленности является наличие собственных значений близких к нулю. Действительно, число обусловленности матрицы есть

$$\mu(Q''(\mathbf{b})) = \frac{\lambda_{max}}{\lambda_{min}},$$

где λ_{max} и λ_{min} соответственно максимальное и минимальное собственное значение матрицы. Известно [10], что обращение плохо обусловленной матрицы численно неустойчиво и умножение обратной матрицы на вектор может увеличивать относительную погрешность в μ раз.

Из (4) следует, что мультиколлинеарность может являться причиной плохого начального приближения, что в свою очередь влияет на число итераций.

Кроме того, плохая обусловленность матрицы $Q''(\mathbf{b})$ повышает дисперсии коэффициентов логистической регрессии, что приводит к неверной оценки вероятности нулевой гипотезы с помощью Z-статистик и статистик Вальда.

Резюмируя, можно сказать, что коррелируемость ковариант является причиной:

— неправильной оценки параметров $\mathbf{b}$ в следствии численной неустойчивости;

— неверного вычисления уровня значимости p для коэффициентов регрессии;

— увеличения вычислительных итераций вплоть до принудительной остановки подпрограммы *glm*.

Плохое начальное приближение. На первом шаге (4) начальное приближение выбирается из решения задачи линейной регрессии с помощью метода наименьших квадратов. Полученная разделяющая плоскость может совсем не разделять наблюдения и быть на большом расстоянии, что крайне затрудняет сходимость итераций на следующих шагах алгоритма. В стандартном алгоритме *glm* есть возможность начать итерации с предоставленных пользователем значений, но при этом варианты вычисления начального приближения не предоставляются.

Наличие локальных минимумов целевой функции. Функция $Q(\mathbf{b})$ может иметь несколько локальных минимумов [11]. В этом случае возможен случай, когда, попав в точку локального минимума, алгоритм завершает свою работу, предоставляя значения коэффициентов $\mathbf{b}$ для этой точки, вместо глобального минимума.

4 Исходный метод построения однофакторных и многофакторных моделей логистической регрессии

Во многих клинических исследованиях логистическая регрессия применяется для выявления отдельных предикторов послеоперационных осложнений (однофакторные модели) и мультипликативных предикторов (многофакторные модели) [3, 12, 13]. Краткая схема метода:

1. построение для каждой коварианты однофакторной модели логистической регресси;
2. отбор в многофакторную модель ковариант с достигнутым уровнем значимости коэффициентов в однофакторных моделях меньше фиксированного порога значений;
3. при необходимости — устранение части ковариант из отобранных для устранения мультиколлинеарности;
4. построение многофакторной модели логистической регрессии;
5. итеративное построение оптимальной по информационному критерию Акаике модели многофакторной логистической регрессии одним из алгоритмов прямого, обратного или смешанного шага;
6. верификация построенных моделей на использованных для обучения данных.

В таблице 2 представлен пример оформления результатов расчетов. Включение возраста в многофакторную модель привело к значительному увеличению достигнутого уровня значимости, что является признаком неустойчивого счета в следствии коррелируемости возраста и стеноза.

Таблица 1: Логистические регрессии для предикторов осложнений

	Однофакторные модели		**Полная многоф. модель**		**Оптим. многоф. модель**	
Коварианты	**ОШ [95 % ДИ]**	**p**	**ОШ [95 % ДИ]**	**p**	**ОШ [95% ДИ]**	**p**
Возраст	0.96 [0.897; 1.02]	0.162	0.991 [0.92; 1.07]	0.812		
ХНМК	1.5 [0.959; 2.48]	0.074	1.51 [0.902; 2.51]	0.117	1.62 [0.977; 2.68]	0.062
ИК	1 [1; 1.02]	0.003	1.01 [1; 1.02]	0.008	1.01 [1; 1.02]	0.003
Стеноз	0.94 [0.889; 0.998]	0.042	0.969[0.915; 1.02]	0.266		

5 Численный метод

Представим модифицированный устойчивый параллельный метод, выполняющий расчеты для исходного метода:

1. параллельно рассчитываются однофакторные модели логистических регрессий для каждой i-ой коварианты, при этом сохранаяем полученные значения свободного коэффициента b_i^1, и коэффициента коварианты b_i^2; с полученной оценкой стандартного отклонения коэффициента d_i^2;
2. отбираем для включения в полную модель коварианты с достигнутым уровнем значимости меньше порогового уровня;
3. при необходимости — удаляем часть ковариант из отобранных для устранения мультиколлинеарности;
4. устраняем коварианты с процентом пропущенных значений больше заданного порогового значения;
5. формируем вектор начального приближения из коэффициентов b_i^2 и среднего арифметического свободных коэффициентов b_i^1;
6. на каждом k-ом шаге итерации следим, чтобы значение коэффициента b^{k_i} не превосходило область значений $b_i^2 \pm c * d_i^2$, где c — заранее фиксированный пороговый уровень; при достижении порогового уровня делаем случайный отскок от граничного значения внутрь области;
7. используя предложенную выше модификацию строим оптимальную многофакторную модель.

6 Результаты

Для численной проверки генерировались наборы значений $(\mathbf{x}_i, y_i)$, $\mathbf{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \end{pmatrix} = \begin{pmatrix} 1 \\ x_{i2} \\ x_{i3} \end{pmatrix}$, $i = 1, \ldots, n$. Объем выборки $n = 40$, первая часть выборки моделируется для первого класса $y_i = -1$, $i = 1, \ldots, 20$, для второго класса $y_i = 1$, $i = 21, \ldots, 40$. Признаки моделируются как независимые гауссовские $\mathbf{X}^1 = (X_2^1, X_3^1)$, $\mathbf{X}^2 = (X_2^2, X_3^2)$ с одинаковой дисперсией равной 1.5 и математическими ожиданиями $\mathbf{E}X_2^1 = 6$, $\mathbf{E}X_3^1 = 3$, $\mathbf{E}X_2^2 = -6$, $\mathbf{E}X_3^2 = -3$. Всего проведено 1000 моделирований.

Таблица 2: Сравнение методов логистической регрессии

		Новый алгоритм			**Алгоритм** *glm*		
E	Оценка	b_1	b_2	b_3	b_1	b_2	b_3
$X_2^1 = 6$	β_j	0.00000	–0.89443	–0.44721	0.00000	–0.89443	–0.44721
$X_3^1 = 3$	$\overline{b}_j$	–0.00063	–0,89737	–0.44578	–0.09851	–0.80379	–0.39406
$X_2^2 = -6$	$SD(b_j$	0.05367	0.00859	0.02984	0.24789	0.14301	0.23874
$X_3^2 = -3$	Δb_j	–0.00082	0.00506	0.00677	0.08537	0,08189	0.07853

В таблице 2 используются следующие обозначения: β_j — аналитические рассчитанные величины коэффициентов логистической регрессии; $\overline{b}_j$ — математические средние оценок коэффициентов логистической регрессии; $SD(b_j$ — средние квадратические значения ошибок оценок коэффициентов; Δb_j — средние значения ошибок оценок коэффициентов. Нетрудно заметить, что разработанный алгоритм выигрывает по сравнению с *glm*.

Заключение

В работе исследованы причины возникновения численных неустойчивостей в работе алгоритма поиска коэффициентов логистической регрессии, представлен общий алгоритм построения моделей логистической регрессии для всестороннего изучения ковариант моделей и разработан экономичный параллельный устойчивый вариант общего алгоритма построения моделей.

Список литературы

[1] Cox D.R., "The regression analysis of binary sequences (with discussion)"// J Roy Stat Soc B. 20 (2): 215–242, 1958.

[2] Walker S.H., Duncan D.B. “Estimation of the probability of an event as a function of several independent variables” // Biometrika. 54 (1/2): 167–178, 1967 doi:10.2307/2333860. JSTOR 2333860.

[3] Truett J., Cornfield J., Kannel W, “A multivariate analysis of the risk of coronary heart disease in Framingham” // Journal of Chronic Diseases. 20 (7): 511–24, 1967 doi:10.1016/0021–9681(67)90082–3

[4] Rosenbaum P.R., Rubin D.B."The Central Role of the Propensity Score in Observational Studies for Causal Effects"// Biometrika. 70 (1): 41–55, 1983 doi:10.1093/biomet/70.1.41

[5] Venkatesan R., Meng J.E., “A novel progressive learning technique for multi-class classification” // Neurocomputing. 207: 310–321, 2016. doi:10.1016/j.neucom.2016.05.006

[6] Hanley, James A., McNeil, Barbara J., “The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve”// Radiology. 143 (1): 29–36, 1982. doi:10.1148/radiology.143.1.7063747

[7] Cleveland W. S.Robust locally weighted regression and smoothing scatter plots //Journal of the American Statistical Association. 1979. Vol. 74, no. 368. Pp. 829–836.

[8] R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[9] Н. П. Васильев, А. А. Егоров, "Опыт расчета параметров логистической регрессии методом Ньютона–Рафсона для оценки зимостойкости растений" // Матем. биология и биоинформ., 6:2 (2011), 190–199

[10] Бахвалов Н. С., Жидков Н. П., Кобельков Г. М., Численные методы, — 8-е изд., 90 ЭЛ. — М.:БИНОМ. Лаб. знаний, 2015. — 639с.

[11] Тырсин А. Н., Костин К. К. Оценивание логистической регрессии как экстремальная задача // Вестн. Том. гос. ун-та. Управление, вычислительная техника и информатика. 2017. № 40. С. 52–60.

[12] Kareva Yulia E., Efendiev Vidady U., et al, Risk factors for the return of mitral insufficiency after coronary artery bypass graft and reconstruction of the mitral valve in patients with ischemic cardiomyopathy // Circulation Pathology and Cardiac Surgery, Vol.23, Issue 2, P. 31–42, 2019 DOI: 10.21688/1681-3472-2019-2-31-42

[13] Simakova M.A., Ryzhkov A.V., et al, Perspectives of using pulmonary arterial stiffness indicators to evaluate the prognosis of patients with pulmonary arterial hypertension // Terapevticheskii arkhiv, V. 90 Issue 1, P. 86–92. DOI: 10.26442/terarkh201890186-92

Лукинов Виталий Леонидович — к.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: vitaliy.lukinov@sscc.ru.
Дата поступления — 30 апреля 2019 г.

ОБРАТНАЯ ЗАДАЧА С ФИНАЛЬНЫМ ПЕРЕОПРЕДЕЛЕНИЕМ ДЛЯ СВЕРХУСТОЙЧИВОЙ ГИПЕРБОЛИЧЕСКОЙ СИСТЕМЫ

Н. А. Люлько

Институт математики имени С.Л.Соболева, Новосибирский государственный университет, 630090, Новосибирск

УДК 517.956.3
DOI: 10.24411/9999-016A-2019-10048

Для линейной автономной распавшейся гиперболической системы первого порядка с двумя независимыми переменными рассматривается класс смешанных задач, которые являются сверхустойчивыми. Показано, что соответствующие полугруппы класса C_0 являются нильпотентными. Это позволяет для рассматриваемых задач поставить корректную обратную задачу с финальным переопределением.

Ключевые слова: распавшиеся гиперболические системы, граничные условия отражения, сверхустойчивость, обратная задача с финальным переопределением.

Введение

В теории управления динамическими системами перевод начального состояния системы в наперед заданное состояние за конечное время T осуществляется зачастую либо в виде граничного управления, либо в виде управления, входящего в саму дифференциальную систему. В случае смешанной задачи в полуполосе $\Pi = \{(x,t) : 0 < x < 1,\, 0 < t < \infty\}$ для линейной автономной гиперболической в узком смысле по Петровскому системы первого порядка в [16] было построено граничное управление, переводящее систему в произвольное заданное состояние, только в случае обратимости задачи (т. е. когда при замене $t \to -t$ задача остается корректной). Если смешанная задача для гиперболической системы обладает свойством стабилизации всех решений к нулю за конечное время, то, очевидно, это свойство обратимости не выполняется. Но в этом случае перевод системы в заданное состояние может быть осуществлен за счет построения источника, входящего в дифференциальную систему.

Рассмотрим в банаховом пространстве X *задачу с финальным переопределением* [15]

$$\frac{d}{dt}x(t) = \mathbf{A}x(t) + f \quad (0 \le t \le T), \tag{1}$$

$$x(0) = x_0, \quad x(T) = x_T.$$

Здесь $\mathbf{A}$ — линейный замкнутый оператор в X с плотной областью определения $D(\mathbf{A}) \subset X$, порождающий полугруппу $V(t)$ класса C_0 [14,19]. Элементы x_0, x_T принадлежат $D(\mathbf{A})$. Решением задачи (1) назовем пару (x, f), где $x \in C^1([0,T]; X)$, $x(t) \in D(\mathbf{A})$ для всех $t \in [0,T]$, и $f \in X$.

В [17,18] задача (1) решена в случае, когда полугруппа $V(t)$ является *сверхустойчивой* (superstable), т.е. экспоненциальный тип полугруппы

$$\omega_0 = \lim_{t\to+\infty} \frac{\log \|V(t)\|}{t}$$

равен $-\infty$ [2,3,6]. Это означает, что для любого числа $\alpha > 0$ существует число $M = M(\alpha) \ge 1$ такое, что

$$||V(t)|| \le Me^{-\alpha t}, \tag{2}$$

ISBN 978-5-901548-42-4

т.е. все решения соответствующей задачи Коши

$$\frac{d}{dt}x(t) = \mathbf{A}x(t) \quad (0 \le t < \infty), \qquad x(0) = x_0 \in X, \tag{3}$$

убывают быстрее экспоненты в любой степени. Полугруппа $V(t)$ называется *нильпотентной*, если существует $t_0 > 0$ такое, что

$$V(t) = 0, \quad t \ge t_0. \tag{4}$$

В этом случае все решения задачи Коши (3) стабилизируются к нулю за конечное время t_0.

Для резольвенты $R(\lambda; \mathbf{A})$ оператора $\mathbf{A}$ справедливо представление

$$R(\lambda; \mathbf{A})x = \int_0^\infty e^{-\lambda t} V(t)x\, dt, \quad x \in X.$$

Отсюда в силу (2) очевидно, что резольвента $R(\lambda; \mathbf{A})$ есть целая функция параметра λ, следовательно, спектр инфинитезимального оператора $\mathbf{A}$ сверхустойчивой полугруппы $V(t)$ пуст. Поэтому примеры таких полугрупп существуют только в бесконечномерных банаховых пространствах. В [5] приведены примеры сверхустойчивых полугрупп в весовом банаховом пространстве $X = L_1(R_+, e^{-\gamma(x)}dx)$ (здесь $\gamma(x) \ge 0$ монотонно неубывающая функция для $x \ge 0$), которые не являются нильпотентными. Инфинитезимальный оператор $\mathbf{A} : X \longrightarrow X$ этих полугрупп имеет вид

$$\mathbf{A}u = -\frac{du}{dx} - \sigma(x)u, \quad D(\mathbf{A}) = \{u \in AC_{loc(R_+)} : \mathbf{A}u \in X,\, u(0) = 0\},$$

где $\sigma(x) \ge 0$ измеримая и локально ограниченная для $x \ge 0$ функция. Нетрудно видеть, что для этого оператора $\mathbf{A}$ задача Коши (3) представляет собой краевую задачу в неограниченной области по пространственной переменной ($x \ge 0$) для уравнения переноса. Если мы рассмотрим уравнение переноса в ограниченной области ($0 \le x \le 1$), то соответствующая полугруппа будет нильпотента. Для простоты [19] положим $\sigma(x) = 0$, $X = L^2(0,1)$, $\mathbf{A} : X \longrightarrow X$,

$$\mathbf{A}u = -\frac{du}{dx}, \quad D(\mathbf{A}) = \{u \in W_2^1(0,1) : u(0) = 0\},$$

тогда соответствующая полугруппа $V(t)$ имеет вид

$$V(t)u_0(x) = u(x,t) = \begin{cases} u_0(x-t), & \text{для} \quad x \ge t, \\ 0, & \text{для} \quad t > x. \end{cases}$$

В работе автора [7] была рассмотрена смешанная задача в полуполосе П для линейной автономной гиперболической в узком смысле по Петровскому системы первого порядка. Было показано, что при определенных распавшихся граничных условиях, рассматриваемые задачи порождают нильпотентные полугруппы.

В настоящей работе мы рассмотрим смешанные задачи в полуполосе П для невырожденных линейных автономных гиперболических систем с граничными условиями отражения. Нами будет сформулирован критерий сверхустойчивости полугрупп, порождаемых этими задачами, и мы покажем, что рассматриваемые сверхустойчивые полугруппы являются нильпотентными. Поэтому для инфинитезимальных операторов этих полугрупп будет корректна обратная задача (1), при решении которой можно воспользоваться следующей теоремой.

Теорема 1. ([17], теорема 4) *Пусть линейный замкнутый оператор $\boldsymbol{A}$ порождает нильпотентную полугруппу $V(t)$ класса C_0, удовлетворяющую условию (4) с фиксированным значением $t_0 > 0$. Тогда при любом выборе $x_0, x_T \in D(\boldsymbol{A})$ элемент $f \in X$, разрешающий обратную задачу (1), находится по формуле*

$$f = \begin{cases} -\mathbf{A}x_T, & 0 < t_0 \le T, \\ -\mathbf{A}x_T + \mathbf{A}(\sum_{k=1}^{N_0} V(kT)(x_0 - x_T)), & 0 < T < t_0, \end{cases}$$

где $N_0 \equiv [t_0/T] - 1$. При этом функция $x(t)$ восстанавливается по формуле

$$x(t) = V(t)x_0 + \int_0^t V(s)f\, ds, \quad 0 \le t \le T.$$

1 Постановка задачи и основные результаты

Рассмотрим в полуполосе Π смешанную задачу для распавшейся автономной гиперболической системы

$$\partial_t u + A(x)\partial_x u + B(x)u = 0, \quad (x,t) \in \Pi, \tag{5}$$

с граничными условиями отражения

$$\begin{aligned} u_j(0,t) &= \sum_{k=1}^{m} p_{jk} u_k(1,t) + \sum_{k=m+1}^{n} p_{jk} u_k(0,t), \quad 1 \le j \le m, \\ u_j(1,t) &= \sum_{k=1}^{m} p_{jk} u_k(1,t) + \sum_{k=m+1}^{n} p_{jk} u_k(0,t), \quad m < j \le n, \end{aligned} \tag{6}$$

и начальными данными

$$u(x,0) = \varphi(x), \quad x \in [0,1]. \tag{7}$$

Здесь неизвестная функция $u(x,t) = (u_1, \dots, u_n)$ и начальная функция $\varphi(x) = (\varphi_1, \dots, \varphi_n)$ — это n-мерные векторы вещественнозначных функций, $A = diag(a_1, \dots, a_n)$ и $B = diag(b_1, \dots, b_n)$ — диагональные матрицы из вещественнозначных функций, константы $n \ge 2$ и $0 \le m \le n$ — фиксированные целые числа, и коэффициенты p_{jk} в граничных условиях (6)– вещественные константы.

Предположим, что в (5) коэффициенты $a_j, b_j \in C^1[0,1]$ $(1 \le j \le n)$, при этом выполняются следующие соотношения:

$$\begin{aligned} &\min\{a_j(x) \,:\, x \in [0,1],\, 1 \le j \le m\} \ge a_0, \\ &\max\{a_j(x) \,:\, x \in [0,1],\, m+1 \le j \le n\} \le -a_0 \end{aligned} \tag{8}$$

для некоторого $a_0 > 0$. Условие (8) означает, что гиперболическая система (5) не вырождена и все ее характеристики равномерно ограничены в $\overline{\Pi}$. Случай $m = 0$ $(m = n)$ соответствует тому, что в системе присутствуют только характеристики с отрицательным (положительным) наклоном.

Следуя [9,10], запишем систему (5)–(7) на гладких решениях в эквивалентном интегральном виде. Для фиксированных $j \le n$, $x \in [0,1]$ и $t > 0$, j-я характеристика уравнения (5), выходящая из точки $(x,t) \in \overline{\Pi}$, определяется как решение $\xi \in [0,1] \mapsto \omega_j(\xi) = \omega_j(\xi, x, t) \in R_+$ задачи

$$\partial_\xi \omega_j(\xi, x, t) = \frac{1}{a_j(\xi)}, \;\; \omega_j(x, x, t) = t.$$

Характеристическая кривая $\tau = \omega_j(\xi, x, t)$ достигает границу Π в двух точках с различными ординатами. Пусть $x_j(x,t)$ обозначает абсциссу той точки, чья ордината меньше. Заметим, что значение $x_j(x,t)$ не зависит от точки x, t, если $t > \frac{1}{a_0}$, а именно:

$$x_j(x,t) = \begin{cases} 0, & \text{если } 1 \le j \le m, \\ 1, & \text{если } m < j \le n, \end{cases} \qquad \text{для } t > \frac{1}{a_0}.$$

Введем в рассмотрение пространство Y, состоящее из вектор-функций $y = (y_1(x,t), ..., y_n(x,t))$, где y_i — непрерывные и ограниченные в замыкании полуполосы Π функции, с нормой

$$||y|| = \max_i \sup_{(x,t)\in\overline{\Pi}} |y_i|,$$

и пусть Z–пространство вектор-функций $z = (z_1(t), ..., z_n(t))$, где z_i — непрерывные и ограниченные для всех $t \ge 0$ функции, с аналогичной равномерной нормой. Обозначим

$$c_j(\xi, x) = \exp \int_x^\xi \left(\frac{b_j}{a_j}\right)(\eta)\, d\eta, \quad 1 \le j \le n,$$

и определим линейный ограниченный оператор $R \in L(Y, Z)$

$$(Ru)_j(t) = \sum_{k=1}^{m} p_{jk} u_k(1,t) + \sum_{k=m+1}^{n} p_{jk} u_k(0,t) \quad 1 \le j \le n. \tag{9}$$

Непосредственные вычисления показывают, что гладкая в $\overline{\Pi}$ функция u есть решение задачи (5)–(7) тогда и только тогда, когда она удовлетворяет следующей системе интегральных уравнений:

$$u_j(x,t) = (Qu)_j(x,t), \quad 1 \le j \le n, \tag{10}$$

где оператор Q, заданный по формуле

$$(Qu)_j(x,t) = \begin{cases} c_j(x_j(x,t),x)\,(Ru)_j\,(\omega_j(x_j(x,t),x,t)), & \text{если} \quad x_j(x,t) = 0 \text{ или } x_j(x,t) = 1, \\ c_j(x_j(x,t),x)\varphi_j(x_j(x,t)), & \text{если} \quad x_j(x,t) \in (0,1), \end{cases}$$

определен на множестве непрерывных в $\overline{\Pi}$ функций, удовлетворяющих начальному условию (7). Непрерывная функция u, удовлетворяющая (10) в $\overline{\Pi}$, называется *непрерывным решением* задачи (5)–(7). Корректность задачи (5)–(7) в пространствах непрерывных и непрерывно диференцируемых функций доказана в [1, 9].

Введем в рассмотрение линейный ограниченный оператор $S \in L(Z,Y)$

$$(Sv)_j(x,t) = c_j(x_j,x)v_j(\omega_j(x_j,x,t)), \quad j \le n, \tag{11}$$

где

$$x_j = \begin{cases} 0, & \text{если } 1 \le j \le m, \\ 1, & \text{если } m < j \le n. \end{cases}$$

Оператор $SR \in L(Y)$ является оператором переноса граничных условий вдоль характеристических кривых внутрь полуполосы Π. Поэтому [9,10,12] каждое непрерывное решение задачи (5)–(7) стабилизируется к нулю за конечное время $d > 0$ (не зависящее от начальных данных), если найдется такое $k \in N$, что $(SR)^k = 0$.

Для формулировки наших результатов запишем систему (5)–(7) в виде абстрактной задачи Коши (3). Для этого введем в рассмотрение оператор $\mathbb{A}: L^2(0,1) \to L^2(0,1)$,

$$(\mathbb{A}v)(x) = -(A(x)\frac{dv}{dx} + B(x)v),$$

с областью определения

$$D(\mathbb{A}) = \{v \in L^2(0,1) : \partial_x v \in L^2(0,1),\ v_{out} = Pv_{in}\}.$$

Здесь $v_{out} = (v_1(0),...,v_m(0),v_{m+1}(1),...,v_n(1))$ и $v_{in} = (v_1(1),...,v_m(1),v_{m+1}(0),...,v_n(0))$, матрица P состоит из коэффициентов отражения p_{ij} в граничных условиях (6). Из результатов, полученных в [10] для неавтономной задачи (5)–(7), вытекает что для любой функции $\varphi \in L^2(0,1)$ задача (5)–(7) имеет единственное L^2-обобщенное решение u, причем для этого решения справедлива оценка

$$||u(\cdot,t)||_{L^2(0,1)} \le Me^{\beta t}||\varphi||_{L^2(0,1)}, \quad t \ge 0. \tag{12}$$

Здесь константы $M \ge 1$ и $\beta \ge 0$ не зависят от t и начальных данных.

Определение 2. ([10, определение 4.3]) Пусть $\varphi \in L^2(0,1)$. Тогда функция $u \in C\left([0,\infty), L^2(0,1)\right)$ называется L^2-обобщенным решением задачи (5)–(7), если для любой последовательности $\varphi^l \in C_0^\infty([0,1])^n$, для которой $\varphi^l \to \varphi$ в $L^2(0,1)$, соответствующая последовательность непрерывно дифференцируемых решений $u^l(x,t)$ задачи (5)–(7), в которой φ заменено на φ^l, обладает сходимостью

$$\|u(\cdot,\theta) - u^l(\cdot,\theta)\|_{L^2(0,1)^n} \to 0 \quad \text{при} \quad l \to \infty,$$

равномерно по θ из отрезка $0 \le \theta \le t$, для любого $t > 0$.

Существование L^2-обобщенного решения задачи (5)–(7) и наличие оценки (12) равносильно тому, что замкнутый линейный оператор $\mathbb{A}$ порождает в пространстве $L^2(0,1)$ полугруппу $W(t)$ класса C_0. Действительно, для любой функции $\varphi \in L^2(0,1)$ L^2-обобщенное решение $u(x,t)$ задачи (5)–(7) является L^2-решением задачи Коши (3) с оператором $\mathbb{A}$. Поэтому [14] справедливо, что

$$u(x,t) = W(t)\varphi(x), \quad t \ge 0,$$

где $W(t)$ — полугруппа, инфинитезимальным оператором которой является оператор $\mathbb{A}$. Тогда справедлива следующая теорема [9, 10, 12].

Теорема 3. *Полугруппа $W(t)$ нильпотентна, когда оператор SR, задаваемый соотношениями (9), (11), нильпотентен.*

В [11] доказаны следующие свойства рассматриваемой полугруппы $W(t)$.

Теорема 4. *Полугруппа $W(t)$ нильпотентна тогда и только тогда, когда спектр оператора $\mathbb{A}$ пустой.*

Доказательство этой теоремы основано на исследовании аналитических свойств резольвенты оператора $\mathbb{A}$. Из теоремы 4 и из свойства пустоты спектра инфинитезимального оператора сверхустойчивой полугруппы вытекает следующее утверждение.

Следствие 5. *Полугруппа $W(t)$ сверхустойчива тогда и только тогда, когда она нильпотентна.*

1.1 Спектр оператора $\mathbb{A}$

Известно [8], что спектр $\sigma(\mathbb{A})$ замкнутого оператора $\mathbb{A}$ состоит из не более чем счетного множества собственных чисел λ конечной кратности, которые предельной точкой могут иметь только $\lambda = \infty$. С другой стороны, из оценки (12) вытекает, что $\sigma(\mathbb{A})$ содержится в полуплоскости $Re\lambda \le \beta$. Найдем характеристическое уравнение для точек спектра оператора $\mathbb{A}$, для чего введем следующие обозначения:

$$\tau_j = |\int_0^1 \frac{1}{a_j(\xi)}\, d\xi|, \quad j = 1, ..., n,$$

где τ_j — время, за которое j-я характеристика системы переходит с одной боковой стороны полуполосы Π на другую сторону Π. Непосредственными вычислениями доказывается следующая лемма.

Лемма 6. *Комплексное число λ есть собственное число оператора $\mathbb{A}$ тогда и только тогда, когда λ удовлетворяет характеристическому уравнению*

$$\Delta(\lambda) = 0, \tag{13}$$

где

$$\Delta(\lambda) = \det(I_n - diag(e^{-\lambda\tau_1}, ..., e^{-\lambda\tau_n})P_1). \tag{14}$$

Здесь I_n — единичная матрица размерности n, а

$$P_1 = diag(\underbrace{1, ..., 1}_{m}, e^{\int_0^1 \frac{b_{m+1}(\xi)}{a_{m+1}(\xi)}\, d\xi}, ..., e^{\int_0^1 \frac{b_n(\xi)}{a_n(\xi)}\, d\xi}) P diag(e^{-\int_0^1 \frac{b_1(\xi)}{a_1(\xi)}\, d\xi}, ..., e^{-\int_0^1 \frac{b_m(\xi)}{a_m(\xi)}\, d\xi}, \underbrace{1, ..., 1}_{n-m}).$$

Выражение $\Delta(\lambda)$ в (14) есть полином Дирихле

$$\Delta(\lambda) = 1 + \sum_{k=1}^{M} E_k e^{-\lambda r_k}, \tag{15}$$

где числа r_k определяются через числа τ_i $(i = 1, ..., n)$, а коэффициенты E_k через элементы матриц A, B и P. При этом $r_1 < r_2 < ... < r_M$ и $M \ge 1$. Из свойств полиномов Дирихле [13] следует, что если хотя бы один из коэффициентов E_k в (15) отличен от нуля, то $\Delta(\lambda)$ имеет счетное число нулей, заключенных в полосе, параллельной мнимой оси, причем в любой ограниченной области их конечное число. Поэтому полином Дирихле $\Delta(\lambda)$ вида (15) не имеет нулей в комплексной плоскости тогда и только тогда, когда $\Delta(\lambda) \equiv 1$, т.е.

$$E_k = 0, \quad k = 1, ..., M. \tag{16}$$

Итак, условие (16) равносильно тому, что спектр оператора $\mathbb{A}$ пустой. При фиксированной матрице P коэффициенты E_k в (16) зависят от элементов матриц A и B. Поэтому спектральный критерий нильпотентности полугруппы $W(t)$ является неустойчивым относительно изменения этих матриц. В [11] доказана следующая теорема.

Теорема 7. *Полугруппа $W(t)$ нильпотентна для любых матриц A и B тогда и только тогда, когда матрица, составленная из абсолютных значений матрицы P, нильпотентна.*

2 Примеры

1. Квазилинейные гиперболические системы вида (5), в котором матрицы A и B могут быть недиагональными и зависеть, в общем случае, от x, t, u, являются объектом исследований в одномерной газовой динамике, в магнитной гидродинамике, в механике деформируемых тел и во многих других областях естествознания.

Записанные через инварианты Римана эти системы имеют вид системы уравнений, в каждом из которых дифференцирование осуществляется вдоль одной неизвестной функции. В [4, глава 3] имеются многочисленные примеры различных физических явлений, математические модели для которых имеют вид смешанных задач (5)–(7).

2. *Пример неустойчивости спектрального критерия нильпотентности относительно изменения матрицы A.* Рассмотрим в Π смешанную задачу для гиперболической системы

$$u_t + u_x = 0, \qquad v_t - v_x = 0, \tag{17}$$

$$u(0,t) = u(1,t) - v(0,t), \qquad v(1,t) = u(1,t) - v(0,t), \tag{18}$$

$$u(x,0) = \varphi_1(x), \qquad v(x,0) = \varphi_2(x). \tag{19}$$

В этом случае

$$\det(I_2 - diag(e^{-\lambda}, e^{-\lambda})P) = 1 - e^{-2\lambda} + e^{-2\lambda},$$

т. е. характеристическое уравнение (13) для соответствующего оператора $\mathbb{A}$ имеет вид

$$1 = 0.$$

Поэтому спектр оператора $\mathbb{A}$ пуст и соответствующая полугруппа нильпотентна, т.е. все решения задачи (17)-(19) становятся равными нулю при $t > 2$.

Рассмотрим для возмущенной ($\varepsilon > 0$) гиперболической системы

$$u_t + (1+\varepsilon)u_x = 0, \qquad v_t - v_x = 0, \tag{20}$$

смешанную задачу с теми же граничными условиями (18) и начальными данными (19). В этом случае

$$\det(I_2 - diag(e^{\frac{-\lambda}{1+\varepsilon}}, e^{-\lambda})P) = 1 + e^{-\lambda} - e^{\frac{-\lambda}{1+\varepsilon}},$$

т.е. характеристическое уравнение (13) примет вид

$$1 + e^{-\lambda} - e^{\frac{-\lambda}{1+\varepsilon}} = 0.$$

Поэтому оператор $\mathbb{A}$, соответствующий возмущенной задаче, имеет бесконечное число собственных значений. Следовательно, соответствующая полугруппа не является нильпотентной, т.е. задача (20), (18), (19) не обладает свойством стабилизации всех решений к нулю за конечное время.

3. *Пример неустойчивости спектрального критерия нильпотентности относительно изменения матрицы B.* Рассмотрим в качестве задачи, обладающей свойством стабилизации всех решений к нулю, задачу (17)-(19). Для возмущенной ($\varepsilon > 0$) гиперболической системы

$$u_t + u_x + \varepsilon u = 0, \qquad v_t - v_x = 0,$$

рассмотрим смешанную задачу с теми же граничными условиями (18) и начальными данными (19). Характеристическое уравнение (13) в данном случае имеет вид

$$1 + e^{-\lambda} - e^{-(\lambda+\varepsilon)} = 0.$$

Поэтому оператор $\mathbb{A}$, соответствующий возмущенной задаче, имеет бесконечное число собственных значений. Следовательно, соответствующая полугруппа не является нильпотентной.

4. *Пример матрицы P размерности (2×2), порождающей нильпотентную полугруппу. При этом полугруппа остается нильпотентной для любых матриц A и B.*

Матрица P, соответствующая граничным условиям в задаче (17)-(19), не удовлетворяла условиям теоремы 7, что позволило построить примеры 2, 3. Рассмотрим матрицу P следующего вида:

$$P = \begin{pmatrix} 0 & p_{12} \\ p_{21} & 0 \end{pmatrix}, \quad p_{12}p_{21} = 0.$$

Очевидно, что $P^2 = 0$ для любых значений p_{12}, p_{21} : $p_{12}p_{21} = 0$, т. е. матрица P удовлетворяет условиям теоремы 7. Поэтому соответствующая смешанная задача (5)–(7) в случае $n = 2$ будет обладать свойством стабилизации всех решений к нулю за конечное время для любых гладких матриц A и B, где элементы A удовлетворяют условиям (8).

Список литературы

[1] Аболиня В. Э., Мышкис А. Д. Смешанная задача для почти линейной гиперболической системы на плоскости // Математический сборник. 1960. Т 50, № 4. с. 423–442.

[2] Balakrishnan A. V. On superstability of semigroups // Systems modelling and optimization, in: M.P.Polis et al. (Eds.), Proceedings of the 18th IFIP TC7 Conference on System Modelling and Optimization, CRC, Research Notes in Mathematics, Chapman and Hall. 1999. P. 12–19.

[3] Balakrishnan A. V. Superstability of systems // Applied Mathematics and Computation. 2005. V. 164, iss. 2. P. 321–326.

[4] Bastin G., Coron J.-M. Stability and boundary stabilization of 1-D hyperbolic systems, 2016.

[5] Ву Нгуен Шон Тунг. Примеры суперустойчивых полугрупп в теории обратных задач для эволюционных уравнений // Современные проблемы теории функций и их применения, ООО Изд-во Научная книга, Саратов. 2018. с. 95–98.

[6] Creutz D., Mazo M. Jr., Preda C. Superstability and finite time extinction for C_0-semigroups. 2013. E-print: https://arxiv.org/abs/0907.4812.

[7] Елтышева Н.А. О качественных свойствах решений некоторых гиперболических систем на плоскости // Математический сборник. 1988. Т. 135, № 2. с. 186–209.

[8] Като Т. Теория возмущений линейных операторов. М.: Мир, 1972.

[9] Kmit I. Smoothing solutions to initial-boundary problems for first-order hyperbolic systems // Applicable Analysis. 2011. V. 90, iss. 11. P. 1609–1634.

[10] Kmit I., Lyul'ko N. Perturbations of superstable linear hyperbolic systems // J. Math. Anal. Appl. 2018. V. 460, iss. 2. P. 838–862.

[11] Kmit I., Lyul'ko N. Finite time stabilization criteria for initial-boundary value hyperbolic problems. In preparation.

[12] Kmit I., Recke L., Tkachenko V., Robustness of the exponential dichotomies of boundary-value problems for the general first-order hyperbolic system // Ukrainian Math. J. 2013. V. 65, iss. 2. P. 260–276.

[13] Левин Б.Я. Распределение корней целых функций. М: Гостехиздат, 1956.

[14] Pazy A. Semigroups of operators and applications to partial differential equations. Springer-Verlag, Berlin, 1983.

[15] Prilepko A. I., Orlovsky D. G., Vasin I. A. Methods for solving inverse problems in matthematical physics. N.Y.; Basel; Marsel Dekker, 2000.

[16] D. Russell, Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions // SIAM Review. 1978. V. 20, P. 639-739.

[17] Тихонов И.В., Ву Нгуен Шон Тунг. Разрешимость линейной обратной задачи для эволюционного уравнения с суперустойчивой полугруппой // Вестник РУДН, Серия МИФ. 2018. Т. 26, № 2. с. 103–118.

[18] Тихонов И.В., Эйдельман Ю.С. Вопросы корректности прямых и обратных задач для эволюционного уравнения специального вида // Математические заметки. 1994. Т. 56, № 2. с. 99–113.

[19] Хилле Э., Филлипс Р. Функциональный анализ и полугруппы. М.: Изд-во иностр. литер., 1962.

Люлько Наталья Альбертовна — к.ф.-м.н., ст. науч.сотр. Института Математики имени С. Л. Соболева СОРАН;
e-mail: natlyl@mail.ru.
Дата поступления — 30 апреля 2019 г.

АЛГОРИТМЫ ПОСТРОЕНИЯ НАЧАЛЬНОГО ПЛАНА В ЗАДАЧАХ ТРАНСПОРТНОЙ ЛОГИСТИКИ

О. А. Ляхов

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск

УДК 519.854.2
DOI: 10.24411/9999-016A-2019-10049

В алгоритмах решения задач транспортной логистики (vehicle routing problem — VRP) для практических целей исходная задача разбивается на две части: построение начального решения; улучшение решения с применением приближенных алгоритмов. Использование моделей автоматической классификации (таксономии) позволяет разделить множество потребителей на близкие подмножества (таксоны) для задач нескольких коммивояжеров и доставки продуктов потребителям в необходимом количестве ограниченными по грузоподъемности транспортными средствами.

Ключевые слова: задача нескольких коммивояжеров, модели маршрутизации транспорта, таксономический анализ.

Введение

Рассматриваются задачи нескольких коммивояжеров и доставки однородных продуктов клиентам. Исходные формулировки включают необходимость посещения заранее заданного числа объектов и возвращением в исходный пункт всех коммивояжеров, а для транспортной логистики каждый коммивояжер (транспортное средство) имеет ограниченную грузоподъемность и для каждого клиента известен размер потребностей. Цель решения задач общем случае формулируется как поиск оптимальных маршрутов передвижения нескольких транспортных средств (коммивояжеров) с доставкой товаров потребителям с минимальными затратами и возвратом в исходную вершину.

Для решения задач достаточной для практики размерности применяются приближенные поисковые алгоритмы, в которых начальное допустимое решение улучшается с помощью вспомогательных процедур, в том числе эвристических правил. Предлагается использование методов таксономического анализа для построения начального эффективного решения задач маршрутизации и построения априорных оценок целевой функции.

1 Задача нескольких коммивояжеров

Обозначения.
$G = (V, E)$ — неориентированный граф;
$V = (i_0, i_1, ..., i_n)$ — множество вершин, i_0 — исходный пункт (депо), $i_1, ..., i_n$ — потребители в задачах маршрутизации (здесь объекты, которые необходимо посетить);
E — множество ребер $(i, j) \in E, i \in V, j \in V$;
$C = \{c_{ij}\}$ — матрица неотрицательных расстояний (стоимости пути) между объектами по ребрам $(i, j), i \in V, j \in V$;
m — количество коммивояжеров или транспортных средств;
R_k — маршрут k-го транспортного средства $(k = 1, ..., m)$;
$C(R_k) = \sum_{j=0}^{n} \sum_{i=0}^{n} c_{ij} x_{ijk}$ — длина маршрута $R_k, (k = 1, ..., m)$.
Переменные $x_{ijk} = 1$, если и только если в k-й маршрут включено ребро $(i, j), i \in V, j \in V$.

Работа выполнена по плану ПФНИ (проект 0315-2016-006).

ISBN 978-5-901548-42-4

Требуется обойти вершины $i_1, ..., i_n$, начиная с пункта i_0 и заканчивая в нем. Различные варианты постановки [1] содержат условия использования обязательного и необязательного использования всех коммивояжеров. Предполагается: для всех ребер выполнение неравенства треугольника; каждая вершина посещается один раз.

Целочисленная модель задачи расчета количества коммивояжеров и построения маршрутов включает условия 1-6 и целевую функцию 7.

$$\sum_{k=1}^{m}\sum_{i=0}^{n} x_{ijk} = 1 \qquad j = 1, ..., n \tag{1}$$

$$\sum_{k=1}^{m}\sum_{i=0}^{n} x_{jik} = 1 \qquad j = 1, ..., n \tag{2}$$

$$\sum_{i=0}^{n} x_{ijk} = \sum_{i=0}^{n} x_{jik} \qquad j = 0, ..., n, k = 1, ..., m \tag{3}$$

$$u_{ik} - u_{jk} + (n+1)x_{ijk} \le n \;\; i, j = 1, ..., n; k = 1, ..., m \tag{4}$$

$$x_{ijk} \in [0, 1], \text{целые} \qquad i, j = 0, ..., n, k = 1, ..., m \tag{5}$$

$$u_{ik} \ge 0 \qquad i = 1, ..., n, k = 1, ..., m \tag{6}$$

$$\sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=1}^{m} c_{ij}x_{ijk} \to \min \tag{7}$$

Условия (1) и (2) определяют однократное посещение каждой внутренней вершины во всех маршрутах и возможность посещения одного потребителя не более одного раза разными коммивояжерами. (3) гарантируют связность маршрутов. Группы условий (4) (заимствованы из [4]) гарантируют включение депо в каждый маршрут. Целевая функция (7) минимизирует суммарные транспортные затраты. В задаче (1-7) необходимо выбрать количество коммивояжеров, обеспечивающее минимальные затраты (7) и построить их маршруты. Если $x_{ijp} = 0,\ i, j = 0, ..., n,\ k = 1, ..., m$, то маршрут p не входит в план (фиктивный).

Варианту этой задачи с обязательным использованием всех коммивояжеров (задача неленивых коммивояжеров [1, с.21]) соответствует модель (1-8). Условия (8) гарантируют отсутствие фиктивных маршрутов, т.е. включение в план всех m коммивояжеров (каждый коммивояжер обязательно должен один раз выйти из депо).

$$\sum_{j=1}^{n} x_{i_0jk} = 1 \quad k = 1, ..., m \tag{8}$$

2 Задачи маршрутизации однородных транспортных средств

Первая постановка задачи маршрутизации транспортных средств при перевозке грузов — VRP (Vehicle Routing Problem) приведена в [2], обзор — в [3]. Задача является развитием задачи коммивояжера. Отличие заключается возможности использования нескольких средств передвижения (больше одного коммивояжера), учете грузоподъемности транспорта и поставки груза каждому потребителю в необходимом количестве. Исходные формулировки включают необходимость посещения заранее заданного числа объектов (вершин) и возврат в исходный пункт (депо). На рис.1 приведены маршруты одного ($0 \to 1 \to 2 \to 3 \to 4 \to 0$) и двух коммивояжеров ($0 \to 1 \to 2 \to 0$ и $0 \to 3 \to 4 \to 0$). Если каждой вершине сопоставить потребность в продукте и ввести грузоподъемность транспортных средств, то задача нескольких коммивояжеров превращается в задачу транспортной логистики по доставке однородных продуктов.

Маршрутизация транспорта в общем случае формулируется как поиск оптимальных маршрутов передвижения нескольких транспортных средств (коммивояжеров) с целью доставки товаров потребителям с минимальными затратами и возвратом в исходную вершину.

В задаче маршрутизации требуется построить минимальные по суммарной длине маршруты, обеспечивающие доставку грузов всем потребителям в заданном количестве. От задачи коммивояжера наследуются требования выполнения для каждого ребра неравенства треугольника и не более чем однократное посещение

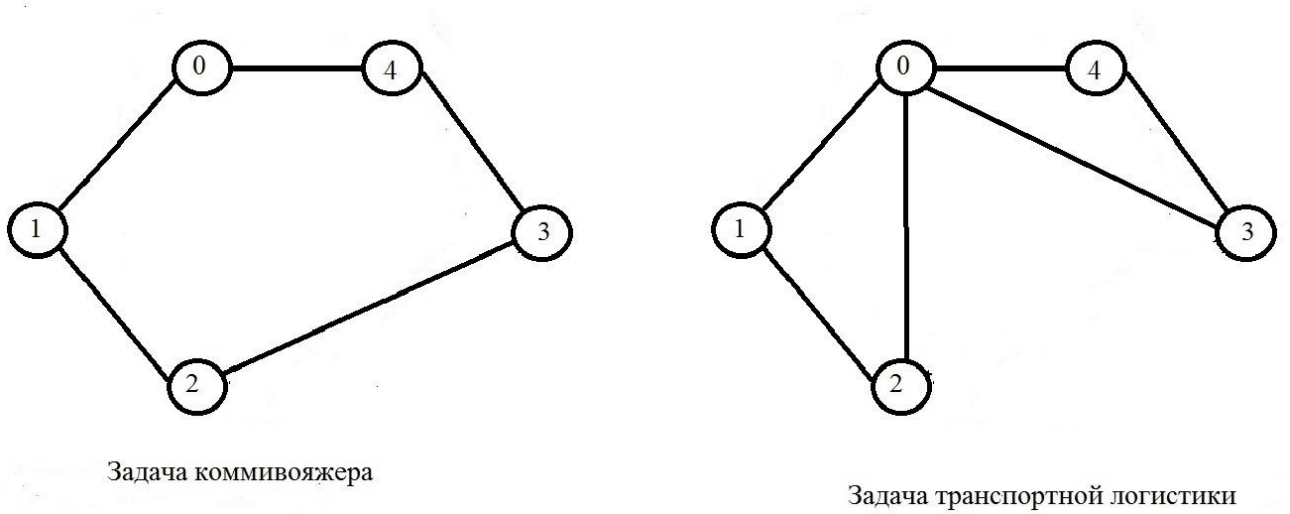


Рис. 1: Задачи коммивояжера и транспортной логистики

каждого потребителя. Все маршруты должны начинаться и заканчиваться в вершине i_0 (депо). Все транспортные средства имеют одинаковую грузоподъемность Q. q_i — объем груза, который необходимо доставить i-му потребителю. Вводится переменная y_{ik} — вес (количество единиц) продукта, доставленного клиенту i транспортным средством k. Предполагается, что парк транспортных средств достаточен для перевозки груза: $m * Q \geq \sum_{i=1}^{n} q_i$.

Целочисленная модель VRP включает условия (1)–(6), целевую функцию (7) и ограничения по доставке грузов (9)–(12).

$$y_{ik} \leq q_i \sum_{j=0}^{n} x_{ijk} \qquad i = 1, ..., n,\ k = 1, ..., m \tag{9}$$

$$\sum_{i=1}^{n} y_{ik} \leq Q \qquad k = 1, ..., m \tag{10}$$

$$\sum_{k=1}^{m} y_{ik} = q_i \qquad i = 1, ..., n \tag{11}$$

$$y_{ik} \geq 0 \qquad i = 1, ..., n,\ k = 1, ..., m \tag{12}$$

Группы условий (9)–(12) определяют необходимость доставки груза каждому потребителю в необходимом количестве и ограниченность грузоподъемности транспортных средств.

VRP широко используется при решении различных практических задач. Известные формулировки связаны с передвижением в чрезвычайных ситуациях. Здесь можно выделить задачи построения маршрутов при осуществлении спасательных работ, доставку расходных материалов по опасным дорогам, ликвидация последствий стихийных происшествий (пожары, наводнения, эпидемии), организация взаимодействия тыловых и авангардных армейских подразделений, материальное обеспечение войск боеприпасами, продовольствием, медикаментами, построение маршрутов осуществления операций по охране государственной границы [5], маршруты разведывательных полетов [6, с.234–235] и т.д. Критерием выбора маршрутов в таких задачах является минимизация ущерба при передвижении и выполнении работ на местах. Ущерб оценивается в зависимости от исходных условий в разных единицах: в дозах облучения персонала при передвижении и выполнении ремонтных работ на АЭС [7], в вероятностных показателях поражения в осуществлении военных операций, во временных показателях пребывания в опасных условиях.

3 Алгоритм таксономического анализа в задачах маршрутизации

Таксономический анализ используется для классификации объектов. Целью является построение классов (таксонов) объектов с близкими свойствами. Задача таксономии связана с разделением исходного множества на непересекающиеся группы по критерию близости [8,9]. Каждый элемент характеризуется набором признаков $r_{i_1}, ..., r_{i_m}$ в метрическом пространстве, $i = 1, ..., n$. Близость (сходство) объектов i и j может быть оценена расстоянием между ними:

$$d(i,j) = \sqrt{\sum_{k=1}^{m}(r_{ik} - r_{jk})^2}, i, j = 1, ..., n. \tag{13}$$

Известная постановка задачи таксономии заключается в разбиении исходного множества объектов на p таксонов так, чтобы суммарное расстояние от эталонов до элементов таксонов было минимальным:

$$d_p = \sum_{\mu=1}^{p} \sum_{i_{\eta} \in \eta} d(u_{\eta}, i_{\eta}) \to min \tag{14}$$

u_{η} — объект, наиболее близкий к центру тяжести таксона η(эталон);
i_{μ} — объект из η-го таксона.
Центр тяжести (эталон) таксона — точка, наименее удаленная от всех элементов таксона.

Алгоритмы таксономии включают две разновидности: построение классификации для заранее заданного числа таксонов [10] и разбиение исходного множества на таксоны при фиксированном радиусе (расстояние от центра тяжести до наиболее удаленной точки таксона).

Разделение исходного множества на заданное число таксонов включают следующие процедуры.

1. Разбиение множества объектов на непересекающиеся подмножества (таксоны).
2. Для каждого таксона определение центра тяжести и эталонного объекта.
3. Просмотр всех элементов таксонов: очередной объект переносится в таксон с наименьшим расстоянием от эталона. Пересчитываются эталоны таксонов.
   Пункт 3 выполняется до тех пор, пока возможно перемещение элементов из таксона в таксон.

Алгоритм позволяет получить оптимальное разбиение объектов при заранее заданном числе таксонов по критерию (14).

Для задач маршрутизации, если положить $c_{ij} = d(i,j)$, применимы методы таксономического анализа.

3.1 Алгоритм таксономии для задачи нескольких коммивояжеров

В задаче таксономии для p коммивояжеров будем считать объектами вершины графа $G = (V, E)$, включая депо. Определим расстояние между объектами $d(i,j) = c_{ij}$. Критериальная функция для разбиения на p таксонов — сумма расстояний от эталонов до всех вершин графа:

$$d_p = \sum_{\mu=1}^{p} \sum_{i_{\eta} \in \eta} c_{u_{\eta} i_{\eta}} x_{i_{\eta}} \to min \tag{15}$$

u_{η} — вершина, наиболее близкая к центру тяжести таксона η(эталон);
i_{μ} — вершина из η-го таксона;
$x_{i_{\eta}}$ — переменная: $x_{i_{\eta}} = 1$, если вершина i принадлежит таксону i_{η}, иначе $x_{i_{\eta}} = 0$.
В отличие от стандартной постановки таксономии будем считать, что вершина i_0 (депо) принадлежит каждому таксону. В алгоритме (раздел 3) изменяется п.3: в каждом таксоне присутствует непереносимая вершина i_0.

Для варианта с обязательным использованием всех p коммивояжеров построенная классификация задает исходные множества для построения маршрутов. Для каждого таксона в отдельности решается задача одного коммивояжера. Не сложно показать, что для любого таксона маршрут из i_0 в i_0, содержащий все внутренние точки таксона, по длине не превосходит удвоенного минимума суммарного расстояния от эталона

до всех элементов таксона. $2d_p$ *является верхней оценкой для решения задачи неленивых коммивояжеров (модель (1)–(8) при выполнении для всех ребер графа* $G = (V, E)$ *неравенства треугольника.*

Для задачи, в которой требуется построить маршруты не более, чем p коммивояжеров (модель (1)–(7) начальное приближение содержит p вариантов разбиения исходного графа на таксоны. Каждому варианту $l = 1, ..., p$ соответствует решение задачи разбиения на l таксонов. *Оценка сверху для задачи (1)–(7) при выполнении неравенства треугольника для всех ребер исходного графа* $D = \min 2d_l,\ l = 1, ..., p$.

В качестве исходного может быть выбран любой из p вариантов деления на таксоны. Для каждого варианта решается задача одного коммивояжера, затем выбирается лучшее решение. Целесообразно начинать с варианта с минимальной оценкой D.

3.2 Алгоритм таксономии для задачи VRP

В построении исходного решения задачи (1)–(7), (9)–(12) дополнительно должны быть отражены ограничения по грузоподъемности транспортных средств и необходимость доставки грузов потребителям в необходимом количестве для вершин каждого таксона. При этом предполагается, что парк транспортных средств достаточен для удовлетворения потребностей всех клиентов. Задача таксономии формулируется следующим образом: необходимо разбить множество вершин графа G на $1 \le l \le m$ таксонов. Каждый таксон соответствует исходной задаче построения маршрута одного транспортного средства. Этой задаче соответствует модель (16)–(21).

$$x_{i_l} = 1, \quad \text{если} \quad i_l \in l, \quad l = 1, ..., m \tag{16}$$

$$x_{i_l} \in [0, 1], \quad \text{целые} \quad i_l \in E \tag{17}$$

$$\textstyle\sum_{i_l \in E} x_{i_l} = n \tag{18}$$

$$x_{0_l} = 1, \quad l = 1, ..., m \tag{19}$$

$$\textstyle\sum_{i_l \in E} x_{i_l} q_{i_l} \ge Q \quad l = 1, ..., m \tag{20}$$

$$d_l = \textstyle\sum_{i_l \in E} c_{i_l u_i} x_{i_l} \to min \tag{21}$$

x_{i_l} — переменная, определяющая принадлежность вершины таксону, $i_l \in l$.
$x_{i_l} = 1$, если вершина i принадлежит таксону l, иначе $x_{i_l} = 0$. $c_{i_l u_i}$ — расстояние между эталоном u_l и вершиной i_l.
Ограничения (16-18) определяют однократное присутствие внутренних вершин в таксонах, (19) показывает, что депо входит в каждый таксон. Группа условий (20) гарантирует перевозки груза в необходимом количестве с без нарушения требований грузоподъемности каждого транспортного средства. Критериальная функция (21) определяет качество классификации вершин.

Любое допустимое решение задачи (16)–(21) соответствует допустимому решению задачи (1)–(7), (9)–(12) с оценкой сверху $D = \min 2d_l,\ l = 1, ..., p$ *при выполнении для ребер исходного графа неравенства треугольника.*

Целесообразность предварительного таксономического анализа для построения начальных допустимых решений VRP задач подтверждается значительным сокращением размерности исходных данных: число переменных во вспомогательной задаче (16)–(21) равно $p*m$; количество переменных в исходной задаче (1)–(7), (9)–(12) равно $2 * m^2 * p + m * p$.

Решая l задач (16)–(21) получим начальные решения для всех вариантов выбора средств передвижения от 1 до p. Для продолжения расчетов в качестве исходного может быть выбран любой из p вариантов деления на таксоны. Для каждого варианта $l = 1, ..., p$ решается задача l коммивояжеров (см. п. 3.1), затем выбирается лучшее решение. Целесообразно начинать с варианта с минимальным значением целевой функции.

Если $p = 1$ (задача одного коммивояжера), то алгоритм таксономии (раздел 3) позволяет получить априорную оценку целевой функции задачи одного коммивояжера.

3.3 Условия неоднократного посещения вершин в модели VRP

В VRP предполагается, что каждая внутренняя вершина маршрута $R_k (k = 1, ..., m)$ посещается один раз. Это соответствует поиску гамильтоновых циклов, но не позволяет отображать возможность посещения вершин более одного раза. Условия (1), (2) гарантируют однократное посещение внутренних вершин каждого из k путей с возвратом в начальную вершину. При невыполненном неравенстве треугольника могут существовать более эффективные решения относительно начальной формулировки задачи, чем оптимальные по

модели (1)–(7), (9)–(12) [11]. Потеря эффективных маршрутов в связи с условием "однократности" может возникнуть в каждом пути данной задачи. Самый короткий путь с однократным посещением вершин может оказаться длиннее пути с повторным заходом в вершины. Если отказаться от однократности и искать решение не только в гамильтоновых циклах, то условия (1)–(7) не соответствуют исходной формулировке задачи, т.к. возможно существование не гамильтоновых более коротких путей, построение которых блокируется ограничениями (2).

Пусть $C' = \{c'_{ij}\}$ — матрица, в которой $c'_{ij} \geq 0$ — минимальное расстояние путей из i в $j, i, j = \overline{0,n}$. Более эффективные решения задачи VRP без условия однократности могут быть получены по модели (1-8) с матрицей C' и восстановлению маршрутов по дугам оптимального пути, таким что $c_{ij} > c'_{ij}$ [11]. При этом в новых вершинах $y_{ik} = 0$ (i- добавленная в маршрут k вершина). В этом случае в построении начального решения используется C'.

Заключение

Для построения начальных решений задач транспортной логистики предложены алгоритмы таксономического анализа. Рассмотрен способ преобразования исходной матрицы расстояний, позволяющий сократить транспортные затраты за счет повторного посещения вершин.

Список литературы

[1] Меламед И.И., Сергеев С.И., Сигал И.Х. Задача коммивояжера. Вопросы теории // Автоматика и телемеханика. 1989. № 9. — С. 3–33.

[2] G. Dantzig, J. Ramser. The truck dispatching problem. Management Science, 6: P. 80–91, 1959.

[3] C. Archetti, M. Speranza. The split delivery vehicle routing problem: A survey. B.Golden, S. Raghavan, and E. Wasil, eds. The Vehicle Routing Problem: Latest Advances and NewChal-lenges. Springer, New York, 103–122, 2008.

[4] Khmelev A., Kochetov Y. A hybrid VND method for the split delivery vehicle routing problem. // Elec. Notes Disc. Math. 2015. V. 47, P. 5–12.

[5] Кучков А.Ф., Лукашевич Н.Ф., Попов Г.П., Шумов В.В. Математическое моделирование служебно-боевых действий пограничных войск. — М.: ФПС России, 1996.

[6] Чуев Ю.В. Иследование операций в военном деле. М.: Воениздат, 1970. 256 с.

[7] Сесекин А.Н., Ташлыков О.Л., Щеклеин С.Е., Куклин М.Ю., Ченцов А.Г., Кадников А.А. Использование метода динамического программирования для оптимизации траектории перемещения работников в радиационно опасных зонах с целью минимизации облучения // Ядерная энергетика. 2006. — № 2. — С. 41–48.

[8] Загоруйко Н.Г. Методы распознавание и их применение. М.:Сов. Радио, 1972. 208 с.

[9] Загоруйко Н.Г. , Елкина В.Н., Лбов Г.С., Емельянов С.В. Пакет прикладных программ ОТЭКС. Изд."Финансы и статистика". М., 1986.

[10] Ильюшонок С.Е. Оптимизация темпов и пропорций развития аграрно-промышленного комплекса. Новосибирск, Наука, 1980. С. 261–267.

[11] Ляхов О. А. Задача минимизации доз облучения при техническом обслуживании АЭС // Проблемы информатики, 2016, № 1. С.19–25

Ляхов Олег Алексеевич — к.э.н., н.с. Институт вычислительной математики и математической геофизики СО РАН;
e-mail: loa@rav.sscc.ru.
Дата поступления — 2 апреля 2019 г.

МОДЕЛИРОВАНИЕ ПОЛЕЙ ДАВЛЕНИЯ В ТРЕЩИНОВАТЫХ КОЛЛЕКТОРАХ

А. А. Мазитов[1], Ю. О. Бобренева[1], И. М. Губайдуллин[1,2]

[1] *Уфимский государственный нефтяной технический университет, ул.Космонавтов,1, 450062, г.Уфа*
[2] *Институт нефтихимии и катализа УФИЦ РАН, Проспект Октября, 141, 450075, г.Уфа*

УДК 519.63
DOI: 10.24411/9999-016A-2019-10050

Рассматривается процесс фильтрации жидкости в пласте в трещиновато-поровом коллекторе, который осуществляется по сети трещин, а матрица является емкостью, непрерывно подпитывающая сеть естественных трещин. Распределение давления в системе «сеть трещин–матрица» описывается уравнениями пьезопроводности. Для решения дифференциальных уравнений использовался метод конечных разностей. Поставленная задача аппроксимировалась неявной разностной схемой. Для решения системы линейных алгебраических уравнений использовался метод матричной прогонки. В результате расчета смоделированы поля давлений при различных входных параметрах.

Ключевые слова: математическая модель, дифференциальные уравнения, разностная схема, трещиноватый коллектор, естественная трещиноватость, метод прогонки.

Достоверная информация о фильтрационно-емкостных свойствах пласта важна на разных этапах нефтяного инжиниринга. Инженер-разработчик должен располагать достаточной информацией о пласте для корректного анализа показателей разработки и прогнозирования добычи при различных вариантах разработки. Инженер по добыче обязан знать состояние добывающих и нагнетательных скважин для установления оптимальной их производительности. Один из основных способов получить информацию можно по результатам гидродинамических исследований скважин.

Математическим фундаментом для анализа гидродинамических исследований скважин на неустановившихся режимах фильтрации является уравнение пьезопроводности в радиальных координатах. Уравнение пьезопроводности выражает связь между пластовым давлением, временем и расстоянием от скважины до точки наблюдения. Уравнение имеет аналитическое решение при условиях, что пласт однородный, изотропный; эффективная толщина его постоянна, сжимаемость жидкости мала и также является постоянной величиной. В данной работе рассматривается модель для карбонатного коллектора, где процесс фильтрации значительно отличается от фильтрации однородного коллектора [1, 2].

Карбонатные коллекторы формируют сложную микроструктуру пустотного пространства, вследствие физико-химических свойств. Карбонатный коллектор характеризуется трещиноватостью и кавернозностью [3, 4].

Процесс фильтрации флюида в трещиноватых коллекторах также значительно изменяется, так как присутствуют две поровые системы–система матриц и система естественных трещин с различными значениями геометрических размеров и фильтрационно-емкостных свойств.

Расчетом характеристик течения в особых условиях резкой неоднородности коллектора занимались разные авторы. В результате литературного анализа была выбрана модель Уоррена–Рута, так как она является универсальной среди всех представляемых моделей фильтрации жидкости в пласте. В модели Уоррена–Рута трещиноватый пласт представлен одинаковыми прямоугольными параллелепипедами, обладающие низкой проницаемостью и высокой пористостью. Низкопроницаемая матрица разделена сетью естественных трещин, они же в свою очередь, наоборот, обладают высокой проницаемостью и низкой пористостью. Движение флюида к скважине происходит только по системе трещин, а матрица непрерывно подпитывает всю систему естественных трещин. В модели матрица и трещины имеют индивидуальные свойства и характеризуются собственными значениями пористости, проницаемости, и сжимаемости [1,5]. Распределение давления в системе «сеть трещин–матрица» описывается следующими уравнениями:

Работа выполнена при частичной финансовой поддержке грантов РФФИ (коды проектов 16-29-15116 и 19-37-50025).

ISBN 978-5-901548-42-4

$$\begin{gathered}\phi_m c_m \frac{\partial P}{\partial t} + S\frac{k_m}{\mu}(P_m - P_f) = 0 \\ \phi_f c_f \frac{\partial P}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(\frac{k_f}{\mu} r \frac{\partial P_f}{\partial r}) - S\frac{k_m}{\mu}(P_m - P_f) = 0\end{gathered} \tag{1}$$

Для уравнения (1), ставятся начальные и граничные условия.

$$\begin{gathered}P_m|_{t=0} = P_0; P_f|_{t=0} = P_0; P_m|_{r=re} = P_e; P_f|_{r=re} = P_e \\ 2\pi h \frac{k_f}{\mu}(r\frac{\partial P_f}{\partial r})_{r=r_w} = -q; P_f|_{r=0} = P_0 - \Delta P\end{gathered} \tag{2}$$

С помощью параметра S количественно характеризуется перераспределение флюида между матрицей и трещинами, и главным образом зависит от двух параметров: формы и размера блоков.

Форма учитывается с помощью параметра n.

$$\alpha = \frac{4n(n+2)}{l_m^2} \tag{3}$$

Этот параметр определяет, в каких направлениях (X, Y, Z) возможен обмен флюида между матрицей и трещинами. Чем выше n, тем легче идет перераспределение флюида. l_m^2 — характеристический размер блока матрицы, равный:

$$l_m = \frac{3abc}{ab + bc + ac} \tag{4}$$

Таким образом, замкнув задачу граничными и начальными условиями, мы получаем модель, характеризующую перераспределение давления в пласте и сети трещин.

Рассмотрим численное решение задачи [6]. Для решения дифференциальных уравнений в частных производных будем использовать метод конечных разностей [7,8].

Для этого введем равномерные сетки по пространству и по времени:

$$\overline{W}_h = \{r_i = ih, i = 0, 1...N, Nh = R_e\} \tag{5}$$

$$\overline{W}_\tau = \{t_j = j\tau, j = 0, 1...M, M\tau = t_e\} \tag{6}$$

При применении явной схемы, необходимо накладывать дополнительное условие, называемое условием Куранта, на шаг по времени, что значительно увеличивает объем вычислений. Поэтому в работе используется неявная схема с первым порядком точности по времени t и вторым по пространственной координате h. При этом неявная разностная схема является абсолютно устойчивой, то есть можно проводить интегрирование краевой задачи с любым разностным шагом по времени:

$$\phi_m c_{tm} \frac{P_{mi}^{j+1} - P_{mi}^{j}}{\tau} = S\frac{k_m}{\mu}(P_{mi}^{j+1} - P_{fi}^{j+1}) \tag{7}$$

$$\phi_f c_{tf} \frac{P_{fi}^{j+1} - P_{fi}^{j}}{\tau} = \frac{1}{r}\frac{k_f}{\mu}\frac{1}{h^2}(r_{i+\frac{1}{2}} P_{fi+1}^{j+1} - (r_{i+\frac{1}{2}} - r_{i-\frac{1}{2}}) + r_{i-\frac{1}{2}} P_{fi-1}^{j+1}) + S\frac{k_m}{\mu}(P_{mi}^{j+1} - P_{fi}^{j+1}) \tag{8}$$

В результате аппроксимации частных производных соответствующими конечными разностями получаем систему линейных алгебраических уравнений. Для решения системы линейных алгебраических уравнений с блочной трехдиагональной матрицей использовался метод матричной прогонки. Матричная прогонка относится к прямым методам решения разностных уравнений и по сравнению с другими прямыми методами решения разностных задач более универсальна, так как позволяет решать уравнения с переменными коэффициентами и не накладывает сильных ограничений на вид граничных условий.

Для написания программного модуля в качестве языка программирования выбран язык C#.

Результатом моделирования фильтрации жидкости в трещиновато-пористых коллекторах является кривая, представляющая собой зависимость забойного давления от времени. График состоит из двух частей: период работы скважины (или кривая падения давления) и период остановки на исследование (или кривая

Таблица 1: Начальные параметры

Параметр	Значение	Единица измерения
Проницаемость матрицы	1E-16	$м^2$
Сжимаемость матрицы	3E-9	$Па^{-1}$
Пористость матрицы	0,1	-
Проницаемость трещин	1E-14	$м^2$
Сжимаемость трещин	5E-9	$Па^{-1}$
Пористость трещин	0,01	-
Сжимаемость нефти	3E-9	$Па^{-1}$
Сжимаемость воды	2E-9	$Па^{-1}$
Вязкость жидкости	0,0008	Па·с
Мощность пласта	10	$м^2$
Длина блока	10	$м^2$
Ширина блока	10	$м^2$
Высота блока	1	$м^2$
Радиус исследования	100	$м^2$
Радиус скважины	0,102	$м^2$
Время отбора	10	сут
Дебит	250	$м^3$/сут
Начальное давление	25	МПа
Конечное давление	22	МПа

восстановления давления). Входными данными являются фильтрационные свойства продуктивного пласта, которые были получены ранее при испытании разведочных скважин. Расчет производится при заданных начальных параметрах. Рассмотрим конкретный пример: скважина простаивала продолжительное время, затем была запущена в работу с постоянным дебитом, работала определенный период и была остановлена на гидродинамическое исследование методом кривой восстановления давления. Считается, что на работу скважины не воздействуют соседние скважины [9]. Давления до запуска и перед остановкой известны. Необходимо построить динамику изменения давления в пласте, то есть КПД и КВД. КПД отражает период от запуска скважины до ее остановки, КВД — период от остановки скважины до восстановления давления до первоначального значения. Исходные данные представлены в таблице 1.

По исходным данным из таблицы 1 получен график зависимости забойного давления от времени, представленный на рисунке 1.

На рисунке четко определяются границы КПД и КВД, соответствующие условиям задачи. Давление пласта до запуска составляло 25 МПа, после запуска происходит резкое снижение давления в течение короткого периода, затем наблюдается стабилизация давления и его плавное снижение до заданного значения на забое скважины. После остановки скважины наблюдается восстановление давления, в очень короткий период оно увеличивается, а затем плавно стабилизируется до исходного значения. Полученный результат отражает реальное поведение скважины при обозначенных условиях.

В результате расчета было выявлено, что при шаге равным 0.001, численное решение согласуется с аналитическим, поэтому все расчеты проводились при данном шаге.

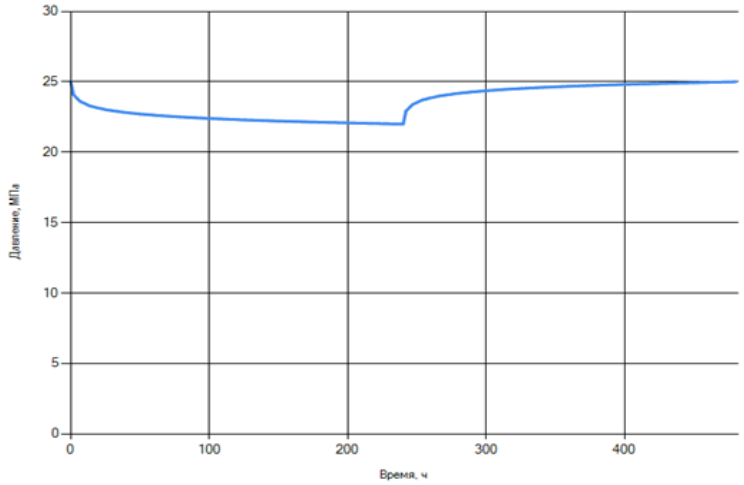


Рис. 1: Динамика изменения забойного давления во времени

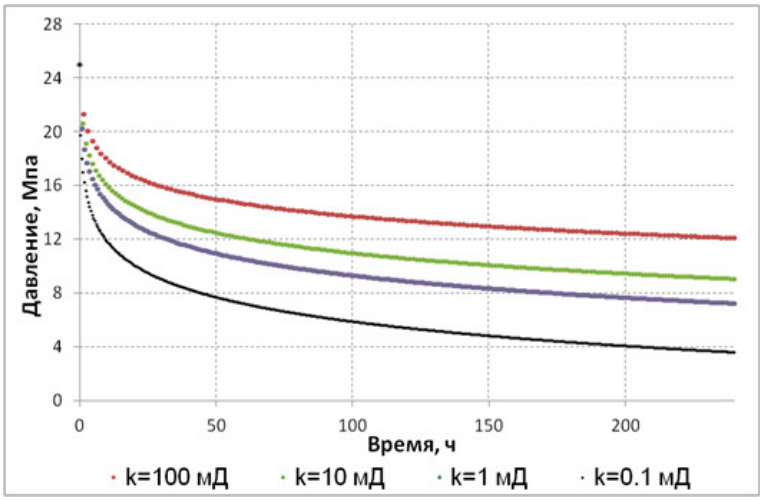


Рис. 2: Динамика изменения забойного давления во времени при различных значениях проницаемости

На рисунке 2 представлены результаты моделирования для разных значений проницаемостей. Чем проницаемость выше, тем быстрее протекают процессы в пласте, значит, проявление псевдорадиального режима наступит раньше, и подключение матрицы в работу системы, тоже наступит быстрее.

Список литературы

[1] Голф-Рахт Т.Д. Основы нефтепромысловой геологии и разработки трещиноватых коллекторов. [ред.] Ковалева А.Г. [перев.] Голованова П.К., Власенова В.В., Покровский В.В. Бардина Н.А. М.: Недра, 1986.

[2] Баренблатт Г.И., Ентов В.М., Рыжик В.М. Движение жидкостей и газов в природных пластах. М.: Недра, 1984.

[3] Warren J.E., Root P.J. The behaviour of naturally fractured reservoirs. б.м.: Soc.Petrol.Eng.J., 1963, pp. 245-255.

[4] Азиз Х., Сеттари Э. Математическое моделирование пластовых систем. [ред.] Максимов М.М. [перев.] Кестнер В.П. Королев А.В. Москва-Ижевск : Институт компьютерных исследований, 2004 г.

[5] Abdullah AI-Ghamdi, Iraj Ershaghi. Pressure Transient Analysis of Dually Fractured Reservoirs. [SPE 26959] б.м.: SPE Journal, 1996.

[6] Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. М.: Наука, 1978 г.

[7] Самарский А.А., Гулин А.В. Численные методы. М.: Наука, 1989 г.

[8] Тихонов А.Н., Самарский А.А. Уравнения математической физики. М.: Наука, 1972 г.

[9] Aguilera R., Ng M.e. Decline-curve analysis of hydraulically fractured wells in dual-porosity reservoirs. [SPE-22938] Dallas, TX : 66th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers held, 1991.

Мазитов Айнур Асгатович — магистрант Уфимского государственного нефтяного технического университета;
e-mail: mazitov.ainur13@gmail.com;
Бобренева Юлия Олеговна — техник-проектировщик управления научных исследований и разработок; Уфимский государственный нефтяной технический университет;
e-mail: yu.o.bobreneva@gmail.com;
Губайдуллин Ирек Марсович — д.ф.-м.н., с.н.с. Института нефтехимии и катализа УФИЦ РАН; профессор Уфимского государственного нефтяного технического университета;
e-mail: irekmars@mail.ru.

ПОДХОД К МОДЕЛИРОВАНИЮ ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМ НА БАЗЕ РАСПРЕДЕЛЕННЫХ СЕТЕЙ ИНТЕЛЛЕКТУАЛЬНЫХ ГЕОСЕНСОРОВ

А. А. Майоров[1], А. В. Матерухин[1], О. Г. Гвоздев[1]

[1] *Московский государственный университет геодезии и картографии (МИИГАиК), 105064, Москва*

УДК 52.08:004.62:681.518.3
DOI: 10.24411/9999-016A-2019-10051

Современные требования к обработке пространственно-временных данных, поступающих от распределенных сетей интеллектуальных геосенсоров, предполагают, что непрерывные запросы на определение динамически меняющихся пространственных отношений между объектами наблюдения должны работать с потенциально неограниченными потоками входящих данных и учитывать временной порядок в этих входящих данных. Требуется обеспечить низкую задержку обработки и стабильность этой задержки при возможных изменениях в интенсивности входных потоков данных. В настоящем докладе представлен разработанный авторами подход к моделированию информационно-измерительных систем, который позволяет реализовать компьютерное моделирование обобщенной информационно-измерительной системы на базе распределенных сетей интеллектуальных геосенсоров.

Ключевые слова: имитационное статистическое моделирование, информационно-измерительные системы, геосенсоры.

Как показано в монографии [1] моделирование информационно-измерительных систем всегда было и остается важным этапом процесса их проектирования. Одним из перспективных классов информационно-измерительных систем являются информационно-измерительные системы на базе распределенных сетей интеллектуальных сенсоров (датчиков). Эти датчики могут представлять данные измерений цифровыми кодами, отслеживать собственную работоспособность и поддерживать ее при аппаратурных отказах, а также осуществлять обмен данными через сетевые каналы. Важным классом таких интеллектуальных сенсоров являются геосенсоры — устройства с возможностью определения своего местоположения и возможностями передачи данных в централизованную систему обработки, оборудованные датчиком (сенсором), которые предназначены для получения данных о событиях, для которых пространственный аспект собранных данных имеет существенное значение [2].

В [3] показано, что при обработке потоков пространственно-временных данных, поступающих от распределенных сетей интеллектуальных геосенсоров в информационно-измерительных системах необходимо эффективно определять значение пространственных предикатов вида: «Есть ли пересечение между областями g и p?», где g — область неопределенности положения геосенсора в момент измерения, а p — область, определенная задачей предметной области, например буферная зона вокруг транспортной магистрали или зона, ограниченная административной границей муниципального образования. Область неопределенности положения геосенсора в момент измерения g в случае использования глобальных навигационных спутниковых систем (например, GPS или ГЛОНАСС) может определяться, к примеру координатами центра (широтой и долготой), полученными приемником, и радиусом, определяемым на основе показателя «горизонтальное снижение точности», который обозначается в англоязычной научной литературе аббревиатурой HDOP (horizontal dilution of precision).

Результаты были получены в рамках выполнения государственного задания Минобрнауки России (номер для публикаций: 5.6972.2017/8.9).

ISBN 978-5-901548-42-4

Необходимость оценки эффективности выполнения различных алгоритмов для определения значения пространственных предикатов к потокам пространственно-временных данных требует проведения соответствующего моделирования работы информационно-измерительной системы. Поскольку в рамках создаваемой модели необходимо воспроизвести различные случайные факторы, события и процессы, связанные функционированием этой системы, то представляется достаточно логичным использовать для этой цели имитационное статистическое моделирование. В результате такого моделирования можно получить вероятностные оценки для различных параметров, характеризующих функционирование и эффективность моделируемой информационно-измерительной системы. Современные требования к обработке пространственно-временных данных, поступающих от распределенных сетей интеллектуальных геосенсоров, предполагают, что непрерывные запросы на определение динамически меняющихся пространственных отношений между объектами наблюдения должны работать с потенциально неограниченными потоками входящих данных и учитывать временной порядок в этих входящих данных. Требуется обеспечить низкую задержку обработки и стабильность этой задержки при возможных изменениях в интенсивности входных потоков данных [2].

В настоящем докладе представлен разработанный авторами подход к моделированию информационно-измерительных систем, который позволяет реализовать компьютерное моделирование обобщенной информационно-измерительной системы на базе распределенных сетей интеллектуальных геосенсоров, концептуальная модель которой была представлена в [4].

Базовая идея этого подхода состоит в том, что моделируемая система рассматривается как стохастическая, поскольку на ее функционирование существенным образом влияют различные случайные факторы — местоположение каждого геосенсора в момент проведения измерения; время проведения каждого измерения каждым геосенсором; задержки, связанные с передачей данных от геосенсора в систему обработки, и т.д. Таким образом, математической моделью входных данных является некоторая многомерная случайная величина X. Выходные данные системы Y, также представляющие собой некоторый вектор, однозначно определяются реализацией этой случайной величины. Связь X и Y определяется оператором системы $L : Y = L(X)$. Поскольку X случайная величина, то Y также является случайной величиной. Задавая статистические характеристики X и зная оператор системы $L(X)$ можно определить статистические характеристики Y, используя имитационное статистическое моделирование.

В качестве оператора $L(X)$ обобщённой информационно-измерительной системы на базе распределенных сетей интеллектуальных геосенсоров предлагается использовать статистическую модель разработанной авторами настоящего доклада обобщенной высокопроизводительной схемы применения алгоритмов определения значений пространственных предикатов (ОВСПП), разработанной с применением авторской композиции методов обобщенного и метапрограммирования [5]

ОВСПП представляет собой фреймворк для выполнения различных алгоритмов определения пространственных предикатов (называемых далее в тексте доклада также ядерными алгоритмами, или для краткости ядрами) в мультипроцессорных окружениях с неоднородным доступом к оперативной памяти. Мультипроцессорные системы представляют собой компьютерные системы, где есть несколько процессоров и одно адресное пространство, видимое для всех процессоров. Под мультипроцессорным окружением с неоднородным доступом к оперативной памяти понимается вычислительная среда мультипроцессорной системы, в которой время доступа к памяти определяется её расположением по отношению к процессору. За системами с такой архитектурой распределенной разделяемой памяти в русскоязычной литературе закрепилось название NUMA-системы [6]. Название происходит от англоязычной аббревиатуры NUMA (Non-Uniform Memory Access). В NUMA-системах каждый процессор имеет локальную память, которая рассматривается как часть общей памяти системы и которой выделен отдельный диапазон адресов в едином адресном пространстве. Помимо возможности обратиться к своей локальной памяти, каждый процессор имеет возможность обратиться к локальной памяти остальных процессоров. В докладе далее будет использоваться термин NUMA-узел для обозначения отдельного узла такой системы.

Принципиальная схема ОВСПП представлена на рис. 1. Zygote-процессы, обозначенные на рис. 1, представляют собой аналоги Zygote-процессов в операционной системе Android или web-обозревателе Chromium.

Первичный запуск программной среды, реализующей ОВСПП, происходит следующим образом:

1. Запускается диспетчер рабочих процессов (далее в тексте доклада он обозначается аббревиатурой ДРП).

2. ДРП порождает необходимое (в зависимости от настроек и количества NUMA-узлов) количество Zygote-процессов, привязанных к отдельным NUMA-узлам, являющихся шаблонами для будущих рабочих процессов.

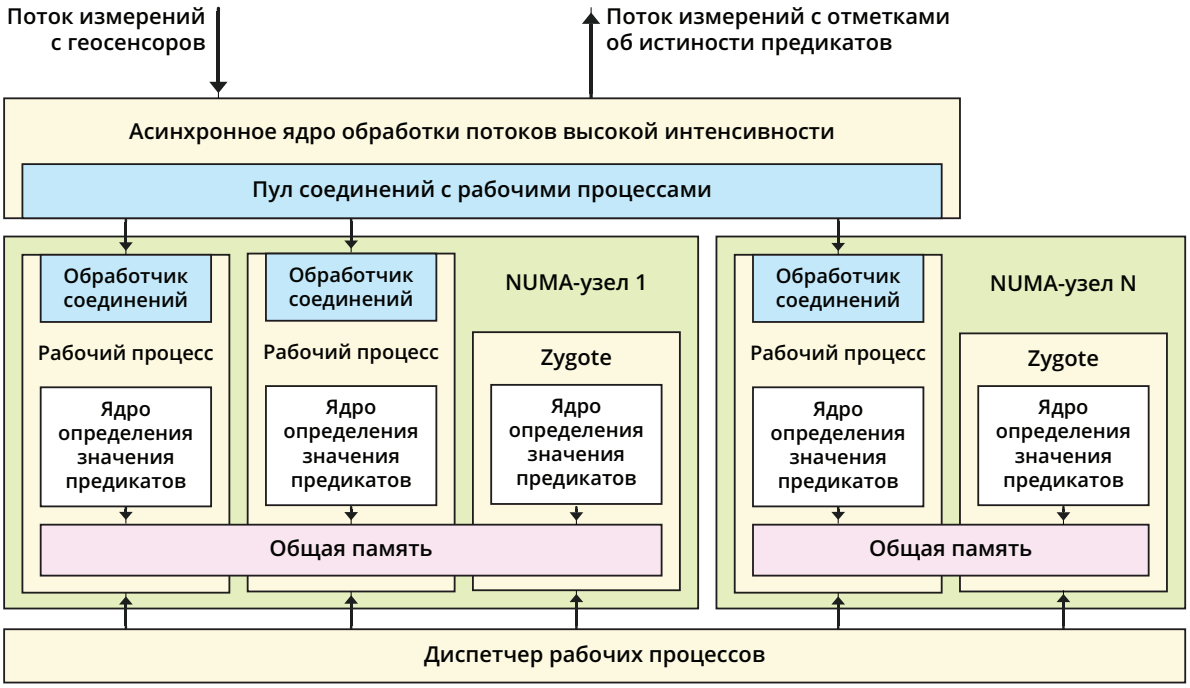


Рис. 1: Обобщенная высокопроизводительная схема применения алгоритмов определения значений пространственных предикатов

3. Каждый Zygote-процесс загружает ядро определения значения предикатов (конкретная реализация зависит от настроек) и запускает процедуру его первичной инициализации.
4. По окончании инициализации всех Zygote-процессов, ДРП передает им команды на порождение необходимого количества рабочих процессов. Данные, необходимые для вычисления предикатов остаются общими между рабочими процессами в рамках одного NUMA-узла, но при этом каждый NUMA-узел хранит свою копию этих данных. Это обеспечивает наилучший баланс между производительностью и потреблением оперативной памяти.
5. Каждый запущенный рабочий процесс переводит ядро определения значения предикатов в рабочий режим и запускает обработчик входящих соединений.
6. Запускается асинхронное ядро обработки потоков высокой интенсивности (далее в тексте доклада - АЯ).
7. АЯ формирует пул соединений с рабочими процессами.
8. Система переходит в рабочее состояние.

Процесс модификации программной среды, реализующей ОВСПП, необходимый при изменении данных областей, определенных задачами предметной области, происходит следующим образом:

1. ДРП порождает необходимое (в зависимости от настроек и количества NUMA-узлов) количество Zygote-процессов, привязанных к отдельным NUMA-узлам, являющихся шаблонами для будущих рабочих процессов. В этот момент новые Zygote-процессы работают независимо от старых.
2. Каждый новый Zygote-процесс загружает ядро определения значения предикатов (конкретная реализация зависит от настроек) и запускает процедуру его первичной инициализации.
3. По окончании инициализации всех Zygote-процессов, ДРП передает им команды на порождение необходимого количества рабочих процессов.
4. Каждый запущенный рабочий процесс переводит ядро определения значения предикатов в рабочий режим и запускает обработчик входящих соединений.

5. ДРП передает старым рабочим процессам команды на прекращение принятия новых задач.
6. АЯ обновляет пул соединений с рабочими процессами.

Процесс обработки одного элемента потока пространственно-временных данных в программной среде, реализующей ОВСПП происходит следующим образом:

1. АЯ считывает данные измерения из входящего потока.
2. АЯ присваивает измерению идентификатор.
3. АЯ извлекает из данных измерения географические координаты и параметр HDOP.
4. АЯ извлекает из пула соединений с рабочими процессами готовое соединение и отправляет рабочему процессу кортеж вида: (Идентификатор измерения, широта, долгота, HDOP).
5. Рабочий процесс получает данные измерения и передает их в ядро определения значения предиката.
6. Рабочий процесс передает результаты определения предикатов в АЯ.
7. АЯ получает результат обработки, сопоставляет его с одним из измерений по присвоенному идентификатору.
8. АЯ снабжает измерение необходимыми отметками о значения предикатов.
9. АЯ передает данные измерения в выходной поток.

Обоснование использования ОВСПП как базового фреймворка для выполнения различных алгоритмов определения пространственных предикатов в информационно-измерительных систем на базе распределенных сетей интеллектуальных геосенсоров содержится в [3].

Таким образом, наш подход к разработке системы статистического моделирования информационно-измерительных систем на базе распределенных сетей интеллектуальных геосенсоров может быть кратко сформулирован следующим образом:

1. Информационно-измерительная система описывается в виде совокупности аппаратных и программных компонентов, передающих друг другу сообщения, содержащие измерительную информацию.
2. Моделью, описывающей топологию передачи этих сообщений между компонентами и отражающей дискретно-событийное представление функционирования информационно-измерительной системы, является разработанная авторами обобщенная высокопроизводительная схема применения алгоритмов определения значений пространственных предикатов.
3. Моделями измеряемой величины являются случайные функции времени (частным случаем таких функций являются детерминированные функции).
4. Моделью, описывающей последовательность значений моментов измерений, является случайный поток событий.
5. Требуется по заданным статистическим характеристикам входных данных и оператору системы $L(X)$ определить численные значения показателей эффективности различных алгоритмов, выполняемых в моделируемой системе.

Подход, предложенный авторами доклада, позволяет уменьшить априорную неопределенность эффективности использования различных алгоритмов определения значений пространственных предикатов в разрабатываемой системе и исследовать процессы в информационно-измерительной системе, для которых нет точного аналитического математического описания. Преимуществом моделирования, использующего предложенный подход, является возможность уже на этапе проектирования информационно-измерительной системы проверить какие алгоритмы выполнения запросов на значение предиката к потокам пространственно-временных данных являются наиболее эффективными с точки зрения различных требований, предъявляемых к системе. В частности можно проверить выполнение требования обеспечить минимальную задержку обработки и стабильность этой задержки при возможных изменениях в интенсивности входных потоков данных. Результаты такого моделирования могут быть использованы при проектировании и разработке новых архитектур информационно-измерительных систем.

Список литературы

[1] Демин А.В., Копорский Н.С. Имитационное моделирование информационно-измерительных и управляющих систем. СПб: СПбГУ ИТМО, 2007. 139 с.

[2] Матерухин А.В. Теоретические основы и методология обработки потоков пространственно-временных данных [Текст] : автореф. дис. ... д-ра техн. наук : 25.00.35 / Матерухин Андрей Викторович; Московский государственный университет геодезии и картографии. М., 2018. 48 с.

[3] Майоров А.А, Матерухин А.В., Гвоздев О.Г. Отчет о научно-исследовательской работе «Разработка геоинформационной технологии создания информационно-измерительных систем на базе распределенных сетей интеллектуальных геосенсоров» (промежуточный, этап 2) (регистрационный номер АААА-Б19-219032190147-1 в ЕГИСУ НИОКТР)

[4] Maiorov A.A., Materukhin A.V. A conceptual model of an information measurement system based on distributed networks of smart geosensors //Measurement Techniques. 2018. Volume 61, Issue 5. pp. 452-458 DOI: https://doi.org/10.1007/s11018-018-1451-0

[5] А.А. Майоров, О.Г. Гвоздев, Композиция методологий обобщенного и метапрограммирования как средство увеличение адаптивности процесса разработки специализированных программных систем обработки пространственных данных // Известия вузов. Геодезия и Аэрофотосъемка. — 2015. — №5. — С.77–81

[6] Организация ЭВМ и систем: Учебник для вузов / Б.Я. Цилькер, С.А. Орлов. — 2-е изд. — СПб.: Питер, 2011. — 688 с.

Майоров Андрей Александрович — д.т.н., заведующий кафедрой Информационно-измерительных систем Московского Государственного Университета Геодезии и Картографии (МИИГАиК);
e-mail: maiorov@miigaik.ru;
Матерухин Андрей Викторович — к.т.н., доцент кафедры Информационно-измерительных систем Московского Государственного Университета Геодезии и Картографии (МИИГАиК);
e-mail: a_materuhin@miigaik.ru;
Гвоздев Олег Геннадиевич — к.т.н., доцент кафедры Информационно-измерительных систем Московского Государственного Университета Геодезии и Картографии (МИИГАиК);
e-mail: gvozdev@miigaik.ru.
Дата поступления — 30 апреля 2019 г.

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ РАЗМЕЩЕНИЯ СТОКОВ В БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЯХ С НЕНАДЁЖНЫМИ УЗЛАМИ ДЛЯ ПОВЫШЕНИЯ ВЕРОЯТНОСТИ УСПЕШНОГО МОНИТОРИНГА

Д. А. Мигов, К. А. Волжанкина

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск

УДК 519.17+519.24
DOI: 10.24411/9999-016A-2019-10052

В статье рассмотрена задача оптимального размещения стоков (sink nodes) в беспроводных сенсорных сетях (wireless sensor networks) в предположении, что сенсоры подвержены отказам. Под оптимальным понимается такое расположение стоков, которое максимизирует надёжность функционирования сети, т.е. вероятность успешного мониторинга заданной области. Предлагается метод размещения стоков в сети, основанный на использованием генетических алгоритмов.

Ключевые слова: беспроводная сенсорная сеть, беспроводная самоорганизующаяся сеть, надежность сети, задача размещения, генетические алгоритмы, площадь покрытия.

Введение

Беспроводные сенсорные сети (БСС, wireless sensor networks, WSNs) получают всё большее развитие, что связано, в первую очередь, с многочисленными приложениями в различных областях, простотой их развертывания, и рядом других преимуществ [1, 2]. БСС являются беспроводными самоорганизующимися сетями (ad hoc networks), наряду с транспортными сетями (VANET), беспроводными mesh-сетями, и другими. Способность самоорганизации и самовосстановления БСС позволяет в случае отказа части узлов спонтанно формировать новую структуру сети.

Основная цель при использовании таких сетей заключается в мониторинге различных объектов: зданий, местности, периметра (границы), и многих других. Узлы БСС содержат некоторые датчики, которые могут быть самыми разнообразными. Чаще других используются датчики температуры, давления, влажности, освещенности, вибрации, реже — магнитоэлектрические, химичские, звуковые и некоторые другие. Набор применяемых датчиков зависит от функций, выполняемых сетью.

Для БСС надежность является важным фактором эффективности их работы [1]. Разработанные ранее классические показатели сетевой надежности не являются в данном случае достаточно адекватными, так как не учитывают специфику функционирования беспроводных сенсорных сетей. В первую очередь, это изначально избыточное количество сенсоров для обеспечения функционирования сети. При этом сеть функционирует, как правило, если достаточное количество сенсоров может устанавливать соединение с одни из стоков. Связность стоков между собой тоже, как правило, является обязательной, но для некоторых сетей это требование будет лишним, так как в них стоки связываются с базовой станцией напрямую и независимо друг от друга. Ранее нами был предложен новый показатель надежности, удовлетворяющий этим условиям [2].

В данной работе рассмотрена возможность оптимального размещения стоков, позволяющей максимально повысить надёжность БСС в указанном выше смысле. Предлагается использовать для размещения стоков биоинспирированные методы, а именно, генетические алгоритмы (ГА). Они хорошо зарекомендовали себя в решении различных задач оптимизации, в том числе задач расстановки [5, 6].

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, код проекта 18-07-00460.

ISBN 978-5-901548-42-4

1 Размещение стоков в БСС для максимизации надежности БСС

Пусть структура беспроводной самоорганизующейся сети представлена в виде неориентированного графа $G = (V, E)$, $|V| = N$, $|E| = M$. Для каждой вершины задана вероятность её присутствия в графе, что соответствует надёжности сенсоров. K — выделенное множество вершин, которые соответствуют стокам сети. Предполагается, что K содержит хотя бы один элемент.

Задано также натуральное число T, такое, что $0 < T < N - |K| + 1$. Смысл этого параметра следующий: T — это требуемое для успешного функционирования сети количество работоспособных сенсоров, имеющих связь с каким-либо стоком, т.к. изначально количество сенсоров часто избыточно, и сеть может функционировать, если достаточное количество сенсоров может устанавливать соединение с одни из стоков.

Определим надёжность сети как вероятность такого события, что вершины K образуют связное множество, и количество вершин, связанных с K, не меньше T. Будем обозначать введённый показатель как $R_{K,T}(G)$. Подробнее этот показатель описан в [2].

Рассмотрим также ещё один показатель, а именно, вероятность покрытия сетью заданной площади. Предполагается, что сеть в состоянии оставаться работоспособной при выходе из строя определённого количества узлов при сохранении покрываемой сетью площади, которая не должна опускаться ниже заданного порогового значения. Другими словами, под надёжностью беспроводной самоорганизующийся сети понимается вероятность её безотказной работы за определённый временной промежуток, под безотказным функционированием сети подразумевается возможность стока получать данные (и, возможно, обрабатывать, в зависимости от функционального назначения узла) от работоспособных (неотказавших) сенсоров, которые ведут мониторинг на области ограниченной снизу площади. Будем обозначать введённый показатель как $S_{K,T}(G)$, где K — множество стоков, T — порог площади. Подробнее этот показатель описан в [4].

Вычислять надёжность предлагается с помощью специально модифицированного для этой цели метода факторизации (ветвления) Идея метода ветвления проста — выбрать еще не рассмотренный элемент графа и перейти к рассмотрению двух подграфов, в одном из которых элемент присутствует с вероятностью 1 (ветвь стягивания), в другом — с вероятностью 0 (ветвь удаления). Вероятность первого из этих случаев будет равна вероятности присутствия элемента, вероятность второго — вероятности отсутствия элемента. Выбранный для ветвления элемент графа называется разрешающим. К полученным подграфам рекурсивно применяем эту же процедуру. Формула полной вероятности при этом дает искомое выражение для надежности на каждом шаге ветвления. Например, для $R_K(G)$ — вероятности связности подмножества вершин K в графе с ненадёжными ребрами — она принимает такой вид:

$$R_K(G) = r_e R_{\tilde{K}}(\tilde{G}(e)) + (1 - r_e) R_K(G \backslash \{e\}), \tag{1}$$

где r_e — надежность ребра e, $\tilde{G}(e)$ — граф, стянутый по ребру e, $G\backslash\{e\}$ — граф, получаемый из G удалением ребра e. Рекурсии продолжаются до получения либо несвязного графа, либо до получения графа, для которого вероятность связности можно вычислить непосредственно — графа специального вида или графа малой размерности

Особенности факторизации при расчёте $R_{K,T}(G)$ и $S_{K,T}(G)$ приводятся псевдокод алгоритма с подробным описанием приведен в [2, 4]. В частности, так рассмотрена возможность раннего прекращения процесса ветвления при достижении заведомо неуспешных событий. Кроме того, используется и оценка этого параметра методом Монте–Карло, что позволяет получать значение с высокой точностью за приемлемое время. Для определения скорости сходимости и необходимого количества итераций использован аппарат, приведённый в одной из наших прошлых работ [3].

2 Генетические алгоритмы

Многие задачи оптимизации в исследовании анализа функционирования современных сетей являются $\mathcal{NP}$-трудными из-за большого пространства решений, поэтому для таких задач часто используют бионические методы, например, генетические алгоритмы, алгоритмы клонирования и др. [7, 8].

В классическом генетическом алгоритме задача формализуется таким образом, чтобы её решение могло быть закодировано в виде вектора генов («генотипа»), в нашем случае каждый ген представлен битовой строкой. Случайным образом создаётся множество генотипов исходной популяции и рассчитывается значение «приспособленности» для каждой особи, чаще всего это значение функции пригодности. Из полученного множества возможных решений («поколения») с помощью оператора «селекция» выбираются особи, к которым применяются «операторы генетического алгоритма»: скрещивание (кроссовер) и мутация, результатом

чего является получение новых решений. Для них вычисляется значение приспособленности, и, после этого производится отбор наилучших решений в следующее поколение.

Такой порядок действий итеративно повторяется, и таким образом моделируется «эволюционный процесс», который продолжается несколько жизненных циклов (поколений), пока не будет выполнен один из критериев остановки алгоритма. Этими критериями чаще всего являются:

- нахождение глобального или субоптимального решения;
- достижение заданного количества итераций, отпущенных на эволюцию;
- стагнация решения (чаще всего происходит из-за вырождения популяции);
- истечение время, отпущенного на эволюцию.

Настройка алгоритма для решения задачи расстановки стоков. Для представления решения (генотипа) будем использовать битовые строки, каждый бит кодирует наличие стока в узлах заданного графа.

Функцией пригодности (приспособленности) является $R_{K,T}(G)$ для некоторой расстановки стоков в вершинах $V = (v_1, \ldots, v_N)$.

Есть несколько типов *селекции* решений для последующего скрещивания, в этой работе использован метод турнирной селекции, так как он имеет преимущество перед методом колеса рулетки из-за более обширного разнообразия генотипов популяции, что позволяет расширить область поиска решений и предотвратить преждевременное вырождение популяции. Оператор селекции отбирает пары родителей $A_i, B_i, i = 1, \ldots, PopSize$, где $PopSize$ — размер популяции, для последующего скрещивания.

Скрещивание проводится следующим образом: у пары родительских генотипов A и B все совпадающие биты переносятся в новый генотип, оставшиеся биты случайно заполняются 1, не превышая количество стоков.

Оператор *мутации* реализован двумя разными способами, сохраняющими количество стоков:

- *мутация перестановкой* меняет два случайных бита в генотипе, имеющие разные значения;
- *мутация инверсией* выбираются два случайных бита, все биты между которыми переставляются в обратной последовательности.

3 Эксперименты

Расчеты проводились на беспроводной сенсорной сети, представленной в виде решётки 5×5, при фиксированном количестве стоков равном 3. В силу небольшого пространства решений, при расчётах было выявлено, что для нахождения оптимального решения достаточно взять размер популяции равным 10 и количество итераций равное 5, вероятность мутации — 0,01. Результаты подсчётов сравнивались с алгоритмом полного перебора.

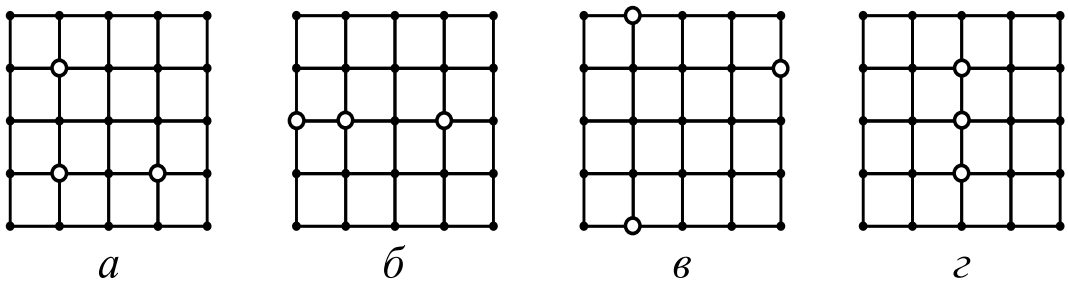


Рис. 1: Оптимальные варианты расстановки стоков

На рис. 1 представлены оптимальные варианты расстановки для $T = 15$ при $p = 0,1$ и $0,5$ (*а*), при $p = 0,9$ (*б*), для $T = 20$ (*в*) и $T = 10$ (*г*) при любых значениях p, полученные в [2]. Все эти варианты были найдены ГА за время, указанное в таблице 1.

При расстановке стоков для максимизации площади покрытия была рассмотрена структура сети, приведённая на рис. 2. Найденными узлами для размещения в них трёх стоков оказались узлы 12, 14, 23.

Таблица 1: Время расчёта на решётке 5×5

$p = 0.5$	$T = 10$	$T = 15$	$T = 20$
Генетический алгоритм	33с 613мс	48с 820мс	94мс
Полный перебор	9м 24с 650мс	11м 39с 200мс	2с 91мс

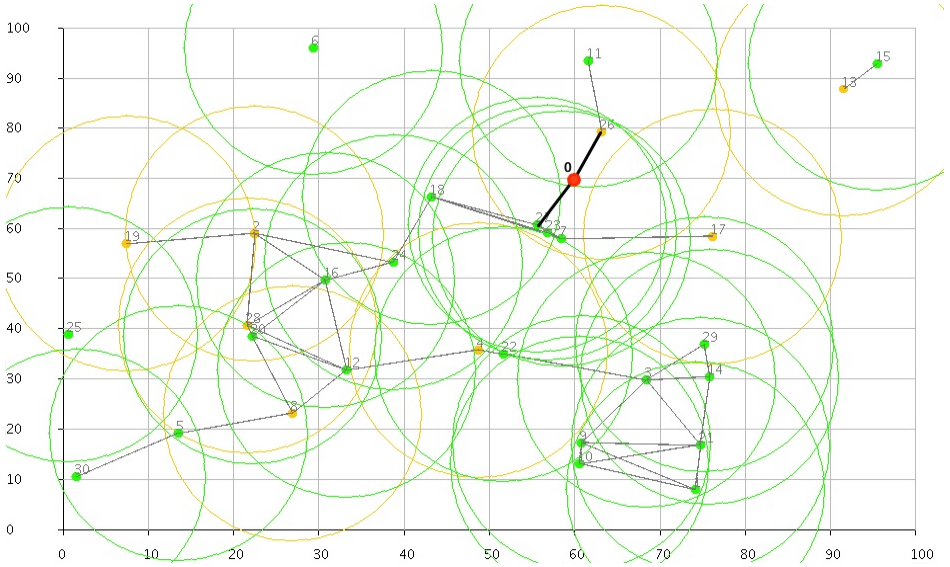


Рис. 2: Тестовая беспроводная сенсорная сеть

При использовании приближенного расчёта методом Монте–Карло удалось значительно сократить время оптимизационного процесса и получить в большинстве случаев те же варианты расстановок. Расчёты велись до обеспечения точности надёжности с погрешностью не более 0.01, чему соответствует 2250 итераций метода Монте-Карло [3].

Заключение

В данной работе мы предлагаем способ размещения стоков в БСС, делающий БСС более надёжной. В качестве надёжности рассматриваются два показателя: вероятность покрытия сетью заданной площади и вероятность доступа к стокам определённого количества сенсоров. Для каждого показателя были использованы точный и приближённый методы расчёта, эксперименты показали целесообразность использования приближённых методов.

Список литературы

[1] Мочалов В.А. Алгоритм оценки надежности структуры сенсорной сети // Информационно-управляющие системы. 2009. № 5. С. 61–66.

[2] Мигов Д.А. Показатель надежности для беспроводных самоорганизующихся сетей // Вестник СибГУТИ. 2014. № 3. С. 3–12.

[3] Д.А. Мигов, Д.В. Винс. Параллельная реализация и имитационное моделирование оценки надёжности сети методом Монте-Карло // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. Т. 47, 2019. С. 66–74.

[4] Кучеров А.В., Мигов Д.А. Расчёт ожидаемой площади покрытия беспроводной сенсорной сети с ненадёжными узлами // Проблемы информатики. № 3 , 2018, с. 21–33.

[5] Tkachev K.V., Volzhankina K.A., Sokolova, O.D. On a Problem of the Monitoring Devices Placement on Transport Networks // 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering, APEIE 2018 — Proceedings, 2018, p. 232–235.

[6] J. Sanner, Y. Hadjadj-Aoul, M. Ouzzif and G. Rubino, "An evolutionary controllers' placement algorithm for reliable SDN networks,"2017 13th International Conference on Network and Service Management (CNSM), Tokyo, 2017, pp. 1-6.

[7] Volzhankina K.A., Rodionov A.S. A Parallel Genetic Algorithm Approach for Monitoring Devices Placement // Труды 13-ой Международной Азиатской школы-семинара в рамках международной мультиконференции IEEE SIBIRCON 2017. Под редакцией А.С. Родионова. 2017. С. 186–189.

[8] Ulutas, B.H., Islier, A.A. Parameter Setting for Clonal Selection Algorithm in Facility Layout Problems // In Proc. of the ICCSA-2007, LNCS 4705, Part I, pp. 886–899.

Мигов Денис Александрович — к.ф.-м.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: mdinka@rav.sscc.ru;
Волжанкина Ксения Александровна — мл.науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: ksu.nech@rav.sscc.ru.
Дата поступления — 30 апреля 2019 г.

АЛГОРИТМЫ ВЕРОЯТНОСТНОГО АНАЛИЗА СЕЙСМИЧЕСКОЙ ОПАСНОСТИ

В. А. Миронов[1], С. А. Перетокин[2], К. В. Симонов[1], М. А. Курако[3,*]

[1] *Институт вычислительного моделирования СО РАН, 660036, Красноярск*
[2] *СКТБ «Наука» ИВТ СО РАН, 660049, Красноярск*
[3] *Сибирский федеральный университет, 660041, Красноярск*

УДК 550.34
DOI: 10.24411/9999-016A-2019-10053

Исследование посвящено разработке вычислительной методики вероятностного анализа сейсмической опасности (ВАСО) как одного из основных этапов в инженерно-сейсмологических изысканиях. От выбора программного обеспечения для проведения ВАСО, понимания его возможностей и ограничений во многом зависит результат исследований. В работе рассматриваются современные подходы к оценке сейсмической опасности и представлен обзор развития программных средств ВАСО.

Ключевые слова: ВАСО, анализ сейсмической опасности, землетрясение, сейсмические параметры опасности, программное обеспечение для ВАСО.

Введение

Оценка сейсмической опасности площадки строительства — это неотъемлемая часть комплекса инженерно-геологических изысканий при проектировании ответственных объектов. В отечественных нормативных документах сейсмическая опасность, как правило, складывается из оценки исходной сейсмичности на некоторых средних грунтах и поправки на реальные грунтовые условия площадки (сейсмическое микрорайонирование — СМР). При этом оценки исходной сейсмичности либо принимаются по картам общего сейсмического районирования (ОСР), либо уточняются по результатам детального сейсмического районирования (ДСР). С 26.11.2018 г. в действие вступил свод правил СП 14.13330.2018 «Строительство в сейсмических районах». В соответствии с пунктом 4.3 для уточнения сейсмичности района строительства объектов повышенного уровня ответственности необходимо проводить специализированные сейсмологические и сейсмотектонические исследования (ДСР).

Состав исследований ДСР существенно зависит от типа и уровня ответственности проектируемого сооружения. Например, для гидротехнических сооружений и объектов атомной отрасли приняты ведомственные нормативные документы, регламентирующие перечень требований, мероприятий и методик, обязательных при выполнении ДСР. Однако, общее для всех то, что характеристики сейсмических воздействий, получаемые в результате работ ДСР, должны иметь обеспеченность, соответствующую уточняемой карте ОСР.

Как и в большинстве стран мира, в Российской Федерации принят вероятностный подход к оценке сейсмической опасности. Нормативный комплект карт ОСР состоит из трёх карт А, В и С, отражающих интенсивность сотрясений, которая не будет превышена в течение 50 лет с вероятностью 90%, 95% и 99% соответственно. Выбор карты А, В или С, по сути, это выбор уровня приемлемого риска в 10%, 5% или 1% того, что будет превышено сейсмическое воздействие, на которое проектируется сооружение.

1 Детерминистский подход при анализе сейсмической опасности

Детерминистский анализ сейсмической опасности (ДАСО, англ. DSHA — deterministic seismic hazard analysis) — это наиболее очевидный и исторически сложившийся подход. Значение интересующего параметра опасности оценивается для максимально возможной магнитуды землетрясения в каждой зоне возникновения очагов

ISBN 978-5-901548-42-4

землетрясения (ВОЗ) по кратчайшему расстоянию до гипоцентра [1]. Детерминистский подход не учитывает случайную природу, присущую таким параметрам землетрясения, как магнитуда, времена повторяемости, положение гипоцентра [2]. Для расчёта используется один сценарий, описывающий вариант максимально-возможного землетрясения.

Существует несколько процедур выбора максимально возможной магнитуды землетрясения. Первая основана на записях о крупнейших исторических землетрясениях и палеосейсмических данных. Этот подход особенно применим в районах с низкой сейсмичностью, где крупные события имеют долгий период повторяемости [2]. При отсутствии каких-либо тектоно-геологических данных предполагается, что максимально возможная магнитуда равна либо максимально наблюдаемой в заданном районе, либо к ней прибавляется некоторое приращение. Обычно это приращение варьируется от 0,25 до 1 единицы магнитуды [1].

Другая часто используемая детерминистическая процедура для оценки максимальной магнитуды, особенно для площадных сейсмических источников, основана на экстраполяции магнитудно-частотной зависимости Гутенберга-Рихтера. Примером таких процедур экстраполяции являются процедура, изложенная Фрёлихом, и процедура, применяемая Нютли [1]. Также существуют методы, где значение максимальной магнитуды оценивают по скорости накопления напряжения или градиенту скорости долговременной тектонической деформации. Однако в большинстве случаев неопределённость по максимальной магнитуде часто достигает 1 единицы магнитуды.

Таким образом, ДАСО позволяет определить сейсмическую опасность исходя из наиболее негативного сценария — максимально возможного землетрясения, без учёта надёжности определения максимальных магнитуд, без учёта неоднозначности модели затухания и т.д.

2 Классический подход для вероятностного анализа сейсмической опасности

Поскольку одиночный сценарий землетрясения не в состоянии обеспечить высокий уровень надёжности полученных результатов, альтернативным подходом для оценки сейсмичности исследуемых площадок является вероятностный анализ сейсмической опасности (ВАСО, англ. PSHA — probabilistic seismic hazard analysis). Оценка опасности методом ВАСО состоит в определении ожидаемого уровня превышения параметров сейсмической опасности с необходимым уровнем надёжности (достоверности), с заданной вероятностью, для всех возможных землетрясений, которые вероятно произойдут в течение заданного промежутка времени [3]. Сейсмическая опасность по ВАСО тесно связана с понятием сейсмической сотрясаемости, введённым Ю.В. Ризниченко в 1965 г.

Как вычислительная процедура подход ВАСО впервые был представлен Корнеллом для пикового ускорения грунта [3]. Чтобы учесть случайную природу возникновения землетрясений он использовал магнитудно-частотную зависимость Гутенберга-Рихтера. Повторяемость землетрясений по времени Корнелл рассматривал как пуассоновский процесс [4], который соответственно следует пуассоновскому распределению. Эпицентры возможных землетрясений равномерно распределялись вдоль линейного разлома или по площади вокруг исследуемой точки наблюдения, а также рассматривались как точечные источники [3]. Для оценки интенсивности в баллах шкалы Меркалли использовалась модель затухания или так называемое уравнение прогнозирования движения грунта (УПДГ, англ. GMPE — ground motion prediction equation) следующего вида:

$$\mathrm{MMI} = c_1 + c_2 M - c_3 \ln R, \tag{1}$$

где MMI — интенсивность в баллах шкалы Меркалли, R — фокальное расстояние до очага, M — магнитуда, c_1, c_2, c_3 — полуэмпирические константы [5].

Для оценки пикового ускорения грунта (ПУГ, англ. PGA — peak ground acceleration) использовалось следующее УПДГ со своим набором констант:

$$\mathrm{PGA} = b_1 \cdot e^{b_2 M} \cdot R^{-b_2}. \tag{2}$$

Отметим, что уравнение для оценки интенсивности землетрясения по шкале Меркалли имеет очень схожий характер с уравнением макросейсмического поля Блейка-Шебалина для точечного источника. В своих исследованиях Корнелл также отмечал, что полная вероятность превышения параметров сейсмической опасности является суммой вкладов от всех источников землетрясений [4].

Вероятностный подход, который предложил Корнелл для оценки сейсмической опасности, в современной литературе трактуется как *классический* ВАСО. Согласно данному подходу, для того чтобы вычислить вероятность превышения заданной амплитуды движения грунта на исследуемой площадке, вклады опасности интегрируются по всем магнитудам и расстояниям для всех зон ВОЗ по теореме о полной вероятности [2]. Таким образом, вероятность превышения определённого уровня сотрясения y^* на площадке равна:

$$\mathbb{P}(Y > y^*) = \sum_i \alpha_i \int_{M_{MIN}}^{M_{MAX}} \int_{R_0}^{R_{MAX}} f_{m,i}(M) \cdot f_{r,i}(R) \cdot \mathbb{P}(Y > y^* \mid M,\, R)\, \mathrm{d}M\, \mathrm{d}R, \tag{3}$$

где α_i — средняя годовая повторяемость землетрясений в i-м источнике, $f_{m,i}(M)$ — функция плотности вероятности магнитуды в i-м источнике, $f_{r,i}(R)$ — функция плотности вероятности эпицентрального расстояния между различными положениями эпицентров в i-м источнике и площадкой, $\mathbb{P}(Y > y^* \mid M,\, R)$ — вероятность того, что амплитуда движения грунта при землетрясении с магнитудой M и эпицентральным расстоянием R превысит y^* [2].

Оригинальный метод не учитывал вероятностный разброс амплитуд движения грунта вокруг среднего значения в моделях затухания. Многие ранние, а так же некоторые более поздние исследования также не учитывали случайный разброс в параметрах опасности при проведении ВАСО [3]. Отметим, что Корнелл в своей работе 1968 г. использовал неограниченную по магнитуде зависимость Гутенберга-Рихтера (отсутствовала верхняя граница возможной магнитуды), что приводило к линейной зависимости в логарифмическом масштабе кривой вероятности превышения амплитуды колебания грунта [5]. Но в дальнейшем, в работе 1969 г. он признавал, что магнитудно-частотная зависимость должна быть ограниченной [6].

Классический подход для ВАСО, известный как подход Корнелла-Макгвайера, включает четыре шага:

- Идентификация и параметризация сейсмических источников, которые оказывают влияние на исследуемую площадку. Источники могут быть представлены как площади, разломы, точки и др. Площадной источник часто используется, когда мы не можем идентифицировать определённый разлом. Каждому источнику землетрясений присваивается равномерное распределение сейсмичности, допуская, что очаги землетрясений равновероятно будут происходить в любой точке в зоне источника [1];

- Характеристика временного и магнитудного распределения сейсмичности для источника. Классический подход предполагает, что повторяемость землетрясений во времени статистически независима, что землетрясения происходят с постоянной частотой, а повторяемость будущих событий не зависит от последнего произошедшего землетрясения [1]. Наиболее часто используемая модель повторяемости магнитуд — это магнитудно-частотная зависимость Гутенберга-Рихтера и зависимость характеристических землетрясений (обе используются в различных модификациях);

- Вычисление УПДГ и их неопределённостей. УПДГ используется, чтобы прогнозировать движение грунта в точках наблюдения. Вычисляемые параметры в УПДГ включают в себя пиковое ускорение грунта, пиковые скорости грунта, пиковые смещения грунта, спектральные ускорения, интенсивность, продолжительность сильного движения грунта и пр. Большинство УПДГ, доступные сегодня — эмпирические, и зависят от магнитуды, расстояния источник-площадка, типа разлома и локальных условий площадки [1]. Выбор УПДГ является важным моментом, поскольку часто УПДГ — это основной фактор вклада неопределённостей при выполнении ВАСО;

- Реализация и учёт различных видов неопределённостей. Например, в положении землетрясения в зоне ВОЗ, в магнитудах, в графиках повторяемости. Конечным результатом ВАСО являются вероятности превышения заданного параметра движения грунта хотя бы от одного события и спектры реакции грунта на определённом пункте наблюдения [1].

3 Вероятностный анализ сейсмической опасности на основе метода Монте Карло

В последнее время наблюдается тенденция к решению проблемы ВАСО с помощью моделирования методом Монте-Карло. Этот метод также известен как стохастический метод, ранее в 1980-х годах использовался для численной оценки сейсмической опасности. Его использовали Розенхауэр, Шапира, Джонсон и Коянаги, однако широкое распространение метод получил после работ Муссона [7].

Процедуру BACO на основе метода Монте-Карло можно условно разделить на два этапа:

1. На основе модели зон ВОЗ генерируется каталог землетрясений для заданного периода времени в T лет. Каждое землетрясение характеризуется набором параметров: магнитудой; длиной и шириной площадки очага; азимутом её верхней кромки, задающим направление простирания очага; углом её падения (углом между горизонтальной плоскостью и площадкой очага); географическими координатами и глубиной гипоцентра;

2. Вычисляется сейсмический эффект от каждого землетрясения в расчётной точке на основе принятой модели затухания и набирается статистика по количеству сотрясений разной интенсивности. Полученная статистика переводится в кумулятивный вид, из которого рассчитываются кривые опасности.

Использование метода Монте-Карло позволяет получить прозрачную схему расчёта на каждом шаге вычислений и возможность более тонкой настройки моделей зон ВОЗ. Использование весьма длительного синтетического каталога землетрясений позволяет учесть неопределённости в случайном характере параметров возможных землетрясений, что позволяет существенно упрощать логическое дерево в части учёта неопределённостей параметров зон ВОЗ.

Наиболее известным комплексом программных модулей для проведения BACO методом Монте-Карло является программный комплекс EqHaz. В нём реализованы три программных модуля. EqHaz1 генерирует синтетический каталог на основе заданных параметров. EqHaz2 вычисляет кривые опасности на основе сгенерированного каталога, выбранного УПДВ и заданных неопределённостей. EqHaz3 вычисляет значение движения грунта для заданного периода повторяемости и выполняет процесс деагрегации вкладов вероятностей превышения значения движения грунта от землетрясений [8].

Отметим, что для создания отечественного комплекта карт ОСР-97 был использован пакет программ PRB-60, разработанный в 1994-1995 гг. Реализованная в нём процедура BACO базируется на методе Монте-Карло. Методические основы пакета и базовые алгоритмы были разработаны А.А. Гусевым с участием Л.С. Шумилиной и В.М. Павлова. Программная реализация пакета осуществлена В.М. Павловым. Актуализация PRB-60 выполнялась в 2003, 2010 и 2016 годах. Последняя модификация этого ПО получила наименование EAST-2016.

4 Развитие программных средств на основе классического подхода BACO

В 1976 году Макгвайер реализовал алгоритм вычисления BACO на основе метода Корнелла в программе EQRISK [9]. Программа EQRISK выполняла BACO, используя только площадные источники землетрясений и уже ограниченное магнитудно-частотное распределение Гутерберга-Рихтера 1958 года [10]. В заложенном алгоритме площадной источник землетрясений делится на субисточники окружностями, центр которых находится в точке исследования с заданным шагом по радиусу. Так как повторяемость равномерно распределена по всей площади исходной зоны ВОЗ, повторяемость каждого субисточника соотносится с повторяемостью исходного также как соотносятся их площади. Далее по методу Корнелла производится расчёт сейсмической опасности от каждого субисточника. Однако вычисление BACO выполнялись только для средних значений движения грунта, без учёта неопределённостей [10].

В 1978 году Макгвайер реализовал метод Корнелла в программе FRISK [11]. Программа FRISK выполняла BACO для прямолинейных разломов или разломов, состоящих из нескольких прямолинейных сегментов. В FRISK использовались ограниченное кусочно-экспоненциальное магнитудно-частотное распределение, логнормальное распределение для длины разрыва и равномерное распределение положения разрыва на разломе. Если разлом (или линеамент) — это линия, вдоль которой равномерно распределены эпицентры землетрясений, то разрыв — это только та область разлома, на которой происходит процесс смещения грунтов при землетрясении (в данном случае разрывы представлены линиями).

Для вычисления использовались ближайшие расстояния от разрыва на разломе до точки наблюдения. В программе учитывались неопределённости для следующих параметров: магнитуды будущих землетрясений; длины разрыва (которая является функцией от магнитуды); положения разрыва на разломе; значений движения грунта.

Кроме введения неопределённостей и типа рассматриваемых источников FRISK отличался от EQRISK тем, что рассчитываемое значение движения грунта зависело и от длины разрыва на разломе. Также имелась возможность задания сетки точек наблюдения на исследуемой площадке. Впоследствии FRISK и EQRISK

были взяты за основу для разработки программы FRISK88M с более расширенными возможностями для расчёта BACO [9]. FRISK88M активно использовался в прошлом, например, для построения вероятностных карт сейсмической опасности Южно-Американской субдукционной зоны [12].

Андерсон и Трифунак в 1977 году в своей работе по вычислению Фурье спектра амплитуд ускорений провели BACO моделированием сейсмичности с более реалистичным подходом [13]. Они использовали пять различных видов источника: точечный; линейный; площадной с произвольной границей, произвольно падающую плоскость разлома и объёмный источник произвольной формы. В разработанном алгоритме расчёта сейсмической опасности использовались два метода. В первом методе модель затухания зависела от эпицентрального расстояния до источника землетрясения (как в работах Макгвайера и Корнелла), а во втором — от расстояния до ближайшей точки разрыва, происходящего в зонах ВОЗ, с учётом длины и направления разрывов. При этом разрывы были представлены кругами заданного радиуса.

Одним из наиболее распространенных инструментов вычисления сейсмической опасности являлась программа SEISRISK III — последняя версия семейства программ SEISRISK. Данная программа и её предыдущие версии использовали классический подход BACO и весьма схожий алгоритм расчёта, который использовали Корнелл и Макгвайер.

Таким образом, можно отметить, что метод, который разработал Корнелл, претерпевал изменения, реализовывался и развивался в различных программах для решения задачи оценки сейсмической опасности. Модели зон ВОЗ становились всё более физическими, а расчёты сейсмической опасности выполнялись на ЭВМ.

Заключение

Поскольку получение корректных оценок сейсмической опасности территорий под строительство и эксплуатацию объектов разной степени ответственности является актуальной задачей, выбор программы и алгоритмов, на основе которых выполняется BACO, является очень важным моментом в исследованиях. Можно утверждать, что выбор инструмента BACO сказывается на корректности и надёжности полученных результатов.

На протяжении более полувека методы расчета BACO улучшаются и претерпевают изменения. Простые точечные источники землетрясений стали более сложными по геометрии и параметрам, стали физически более обоснованными. УПДГ становятся более сложными, с большим набором влияющих параметров. В настоящее время, в связи с развитием ЭВМ, развитием ГИС технологий, накоплением базы данных записанных землетрясений и определением их параметров, существует необходимость постоянно развивать и дополнять возможности программных модулей.

Список литературы

[1] Kijko A. Introduction to Probabilistic Seismic Hazard Analysis // Encyclopedia of Solid Earth Geophysics. Springer, 2011. p. 1–27.

[2] Atkinson G.M. The Integration of Emerging Trends in Engineering Seismology // 13th World Conference on Earthquake Engineering, Lisbon, Portugal (September 24–28, 2012). Lisbon, 2012.

[3] Gupta I.D. Probabilistic seismic hazard analyses method for mapping of spectral amplitudes and other design-specific quantities to estimate the earthquake effects on manmade structures // ISET Journal of Earthquake Technology, Paper No. 480. 2007. V. 44. No. 1, March. p. 127–167.

[4] McGuire R.K. Probabilistic seismic hazard analysis: Early history // Earthquake Engng Struct. Dyn. 2008, 37. p. 329–338.

[5] Cornell C.A. Engineering seismic risk analysis // Bulletin of the Seismological Society of America. 1968, 58 (5). p. 1583–1606.

[6] Cornell C.A., Vanmarcke E.H. The major influences on seismic risk // Proceedings of the 4th World Conference on Earthquake Engineering. Santiago, Chile, 1969.

[7] Musson R.M.W. The use of Monte Carlo simulations for seismic hazard assessment in the U.K. // Annally Di Geofisica. 2000, No. 43. p. 1–9.

[8] Atkinson G.M., Assatourians K. EqHaz: An Open-Source Probabilistic Seismic-Hazard Code Based on the Monte Carlo Simulation Approach // Seismological Research Letters, 2013. V. 84, No. 3. p. 516–524.

[9] Monelli P.D., Weatherill G., et al. OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model // Seismological Research Letters. 2014. V. 85, No. 3. p. 692–702.

[10] McGuire R. Fortran program for seismic risk analysis // U.S. Geol. Surv. Open-File Rept. 76-67.

[11] McGuire R. FRISK: Computer program for seismic risk analysis using faults as earthquake sources // U.S. Geol. Surv. Open-File Rept. 78-1007.

[12] Dimate C., Drake L., Yepez H., et al. Seismic hazard assessment in the Northern Andes (PILOTO Project) // Annally Di Geofisica. 1999. V. 42. No. 6. p. 1039–1055.

[13] Anderson J.G., and Trifunac M.D. On Uniform Risk Functionals Which Describe Strong Earthquake Ground Motion: Definition, Numerical Estimation, and an Application to the Fourier Amplitude Spectrum of Acceleration // Report CE 77-02, University of Southern California, Los Angeles, U.S.A, 1977.

Симонов Константин Васильевич — д.т.н., в.н.с. Института вычислительного моделирования СО РАН; e-mail: simonovkv@icm.krasn.ru;

Перетокин Сергей Анатольевич — к.т.н., с.н.с. Института вычислительных технологий СО РАН; e-mail: saperetokin@yandex.ru;

Миронов Василий Александрович — аспирант Института вычислительного моделирования СО РАН; e-mail: vasya-kun@mail.ru;

Курако Михаил Александрович — ст. преп. Института космических и информационных технологий Сибирского федерального университета;

e-mail: mkurako@sfu-kras.ru.

Дата поступления — 30 апреля 2019 г.

DYNAMIC APPROACH TO CLASSICAL MOMENT PROBLEM

A. S. Mikhaylov[1,2], V. S. Mikahaylov [1,2],

[1] *St. Petersburg Department of V.A. Steklov Institute of Mathematics of the Russian Academy of Sciences, 191023, St. Petersburg*

[2] *Saint Petersburg State University, 199034, St. Petersburg*

UDC 517.984.54, 517.977.1, 517.927.25
DOI: 10.24411/9999-016A-2019-10054

We consider the problem of the construction of a measure supported on a real line from prescribed moments. The main idea is to use the auxiliary dynamical system with the discrete time associated with a semi-infinite Jacobi matrix. Then the inverse dynamic data for this system, so called response operator (discrete analog of a dynamic Dirichet-to-Neumann map) is given in terms of moments, and we can use ideas of the Boundary Control method to recover the spectral data, i.e. the measure of a truncated moments problem, from dynamic one. The remarkable fact is that in our procedure we do not use the Jacobi matrix itself. We also formulate the results on the uniqueness of the solution of Hamburger and Stieltjes moments problems.

Keywords: moment problem, Boundary control method, Jacobi matrices.

Introduction

For a given a sequence of numbers $s_0, s_1, s_2, \dots$ called moments, a solution of a Hamburger moment problem [1,13] is a Borel measure $d\rho(\lambda)$ on $\mathbb{R}$ such that

$$s_k = \int_{-\infty}^{\infty} \lambda^k \, d\rho(\lambda), \quad k = 0, 1, 2, \dots \tag{1}$$

In the present paper we offered an approach to this problem based on the ideas of the Boundary control (BC) method [5, 6]. We consider an initial boundary value problem (IBVP) for an auxiliary dynamical system with discrete time for a Jacobi matrix. The associated with this system is so-called response operator (discrete analog of a dynamic Dirichlet-to-Neumann map), which usually plays the role of inverse data in dynamic inverse problems [5, 8]. The dynamical system of this type and inverse problems for this system were studied in [9–12], where the authors have shown that the kernel of response operator, which is called response vector, admits a spectral representation in terms of a spectral measure of the Jacobi matrix. Using the ideas of the BC method we extract the spectral data (the spectral measure, which is a solution of the moment problem!) from the dynamical one, i.e. from the response vector.

In the first section we set up the IBVP for a dynamical system associated with a finite Jacobi matrix, derive special representations of its solution, introduce operators of the BC method. In the second section we solve a truncated Hamburger moment problem by extracting spectral data from the dynamic one and derive the generalized spectral problem, whose solution is exactly the spectrum of truncated Jacobi matrix and controls of the dynamical system with special properties.

The work has been supported by the Russian foundation of basic research (grant numbers 18-01-00269, 17-01-00529, 17-01-00099).

ISBN 978-5-901548-42-4

1 Initial boundary value problem for a dynamical system associated with Jacobi matrix. Operators of the Boundary control method

For a given sequence of positive numbers $\{a_0, a_1, \ldots\}$ (in what follows we assume $a_0 = 1$) and real numbers $\{b_1, b_2, \ldots\}$, we denote by A the Jacobi operator, defined on l_2, which has a matrix form:

$$A = \begin{pmatrix} b_1 & a_1 & 0 & 0 & 0 & \ldots \\ a_1 & b_2 & a_2 & 0 & 0 & \ldots \\ 0 & a_2 & b_3 & a_3 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \tag{2}$$

For $N \in \mathbb{N}$, by A^N we denote the $N \times N$ Jacobi matrix which is a block of (2) consisting of the intersection of first N columns with first N rows of A. Fixing $N \in \mathbb{N}$ we consider the dynamical system with discrete time associated with a finite Jacobi marix:

$$\begin{cases} v_{n,t+1} + v_{n,t-1} - a_n v_{n+1,t} - a_{n-1} v_{n-1,t} - b_n v_{n,t} = 0, \ t \in \mathbb{N} \cup \{0\}, \ n \in 1, \ldots, N, \\ v_{n,-1} = v_{n,0} = 0, \quad n = 1, 2, \ldots, N+1, \\ v_{0,t} = f_t, \quad v_{N+1,t} = 0, \quad t \in \mathbb{N}_0, \end{cases} \tag{3}$$

where $f = (f_0, f_1, \ldots)$ is a *boundary control*. The solution to (3) is denoted by v^f. Note that (3) is a discrete analog of dynamical system with boundary control for a wave equation on an interval [2, 7].

The operator corresponding to a finite Jacobi matrix we also denote by A^N. Operator A^N is defined on $\mathbb{R}^N$, is given by

$$\begin{aligned} (A\psi)_n &= a_n \psi_{n+1} + a_{n-1} \psi_{n-1} + b_n \psi_n, \quad 2 \leqslant n \leqslant N-1, \\ (A\psi)_1 &= b_1 \psi_1 + a_1 \psi_2, \quad n = 1, \end{aligned}$$

and the Dirichlet condition at the "right end":

$$\psi_{N+1} = 0.$$

We fix some positive integer T and denote by $\mathcal{F}^T$ the *outer space* of the system (3), the space of controls: $\mathcal{F}^T := \mathbb{R}^T$, $f \in \mathcal{F}^T$, $f = (f_0, \ldots, f_{T-1})$, we use the notation $\mathcal{F}^\infty = \mathbb{R}^\infty$ when control acts for all $t \geqslant 1$.

Definition 1. *For $f, g \in \mathcal{F}^\infty$ we define the convolution $c = f * g \in \mathcal{F}^\infty$ by the formula*

$$c_t = \sum_{s=0}^{t} f_s g_{t-s}, \quad t \in \mathbb{N} \cup \{0\}.$$

The input $\longmapsto$ output correspondence in the system (3) is realized by a *response operator*: $R_N^T : \mathcal{F}^T \mapsto \mathbb{R}^T$ defined by the rule

$$\left(R_N^T f\right)_t = v^f_{1,t}, \quad t = 1, \ldots, T.$$

This operator plays the role of inverse data, the corresponding IP were considered in [9, 11]. The *response vector* is a convolution kernel of a response operator, $r^N = (r_0^N, r_1^N, \ldots, r_{T-1}^N)$:

$$\left(R_N^T f\right) = r^N * f_{\cdot -1}.$$

By choosing the special control $f = \delta = (1, 0, 0, \ldots)$, the kernel of a response operator can be determined as

$$\left(R_N^T \delta\right)_t = v^\delta_{1,t} = r^N_{t-1}.$$

Let $\phi_n(\lambda)$ be a solution to

$$\begin{cases} a_n \phi_{n+1} + a_{n-1} \phi_{n-1} + b_n \phi_n = \lambda \phi_n, \\ \phi_0 = 0, \ \phi_1 = 1. \end{cases} \tag{4}$$

Denote by $\{\lambda_k\}_{k=1}^N$ the roots of the equation $\phi_{N+1}(\lambda) = 0$, it is known [1, 13] that they are real and distinct. We introduce the vectors $\phi^n \in \mathbb{R}^N$ by the rule $\phi_i^n := \phi_i(\lambda_n)$, $n, i = 1, \ldots, N$, and define the numbers ρ_k by

$$(\phi^k, \phi^l) = \delta_{kl} \rho_k,$$

where $(\cdot, \cdot)$ is a scalar product in $\mathbb{R}^N$.

Definition 2. *The set of pairs*

$$\{\lambda_k, \rho_k\}_{k=1}^N$$

is called spectral data of operator A^N.

Let $\mathcal{T}_k(2\lambda)$ be the Chebyshev polynomials of the second kind: i.e. they satisfy

$$\begin{cases} \mathcal{T}_{t+1} + \mathcal{T}_{t-1} - \lambda \mathcal{T}_t = 0, \\ \mathcal{T}_0 = 0,\ \mathcal{T}_1 = 1. \end{cases}$$

In [9, 11] the following representation for the solution v^f was proved:

Proposition 1. *The solution to (3) admits a representation*

$$v^f_{n,t} = \begin{cases} \sum_{k=1}^N c_t^k \phi_n^k, & n = 1, \dots, N, \\ f_t, & n = 0. \end{cases} \quad , \quad c^k = \frac{1}{\rho_k} \mathcal{T}(\lambda_k) * f \tag{5}$$

The *inner space* of dynamical system (3) is $\mathcal{H}^N := \mathbb{R}^N$, $h \in \mathcal{H}^T$, $h = (h_1, \dots, h_N)$. By (5) we have that $v^f_{\cdot,T} \in \mathcal{H}^N$. For the system (3) the *control operator* $W_N^T : \mathcal{F}^T \mapsto \mathcal{H}^N$ is defined by the rule

$$W_N^T f := v^f_{n,T}, \quad n = 1, \dots, N.$$

The *connecting operator* $C_N^T : \mathcal{F}^T \mapsto \mathcal{F}^T$ is defined via the quadratic form: for arbitrary $f, g \in \mathcal{F}^T$ we have that

$$\left(C_N^T f, g\right)_{\mathcal{F}^T} = \left(v^f_{\cdot,T}, v^g_{\cdot,T}\right)_{\mathcal{H}^N} = \left(W_N^T f, W_N^T g\right)_{\mathcal{H}^N}.$$

The spectral function of A^N is introduced by the rule

$$\rho^N(\lambda) = \sum_{\{k \mid \lambda_k < \lambda\}} \frac{1}{\rho_k}, \tag{6}$$

Then (5) implies the following representation formulas:

$$v^f_{n,t} = \int_{-\infty}^{\infty} \sum_{k=1}^{t} \mathcal{T}_k(\lambda) f_{t-k} \phi_n(\lambda)\, d\rho^N(\lambda), \tag{7}$$

$$r^N_{t-1} = \int_{-\infty}^{\infty} \mathcal{T}_t(\lambda)\, d\rho^N(\lambda), \quad t \in \mathbb{N}, \tag{8}$$

$$\{C_N^T\}_{l+1,\,m+1} = \int_{-\infty}^{\infty} \mathcal{T}_{T-l}(\lambda) \mathcal{T}_{T-m}(\lambda)\, d\rho^N(\lambda), \quad l, m = 0, \dots, T-1. \tag{9}$$

Another formula for C^T which is valid if $T \leqslant N$ (see [9, 11]) says that:

$$C^T = \begin{pmatrix} r_0 + r_2 + \ldots + r_{2T-2} & r_1 + \ldots + r_{2T-3} & \ldots & r_T + r_{T-2} & r_{T-1} \\ r_1 + r_3 + \ldots + r_{2T-3} & r_0 + \ldots + r_{2T-4} & \ldots & \ldots & r_{T-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{T-3} + r_{T-1} + r_{T+1} & \ldots & r_0 + r_2 + r_4 & r_1 + r_3 & r_2 \\ r_T + r_{T-2} & \ldots & r_1 + r_3 & r_0 + r_2 & r_1 \\ r_{T-1} & r_{T-2} & \ldots & r_1 & r_0 \end{pmatrix} \tag{10}$$

The speed of a wave propagation in (3) is finite, which implies the following dependence of inverse data on coefficients $\{a_n, b_n\}$: for $M \in \mathbb{N}$, $M \leqslant N$, the element $v^f_{1,2M-1}$ depends on $\{a_0, a_1, \dots, a_{M-1}\}$, $\{b_1, \dots, b_M\}$, on observing this we can make the following

Remark 1. *The entries of response vector* $(r_0^N, r_1^N, \dots, r_{2N-2}^N))$ *depends on* $\{a_0, \dots, a_{N-1}\}$, $\{b_1, \dots, b_N\}$, *and does not depend on the boundary condition at* $n = N+1$, *the entries starting from* r_{2N-1}^N *does "feel" the boundary condition at* $n = N+1$.

2 Truncated moment problem. Recovering Dirichlet spectral data

We observe the following: in the moment problem we are given the sequence of moments (1), and in the dynamic IP for the system (3) we are given a response vector [9, 11], whose spectral representation has a form (8). Thus the knowledge of moments $\{s_0, s_1, \ldots\}$ implies a possibility to calculate the response vector $\{r_0, r_1, \ldots\}$ by (8).

Definition 3. *By a solution of a truncated moment problem of order N we call a Borel measure $d\rho(\lambda)$ on $\mathbb{R}$ such that equalities (1) with this measure hold for $k = 0, 1, \ldots, 2N$.*

Remark 2. *The results from the previous section implies that from finite set of moments $\{s_0, s_1, \ldots, s_{2N-2}\}$, which is equivalent to the knowledge of $\{r_0, r_1, \ldots, r_{2N-2}\}$ it is possible to recover Jacobi matrix $A^N \in \mathbb{R}^{N\times N}$ whose elements can be thought of as a coefficients in dynamical system (3) with Dirichlet boundary condition at $n = N + 1$, or $N \times N$ block in semi-infinite Jacobi matrix in (3) with no condition at the right end.*

In [11] the authors proved the following

Theorem 1. *The vector $(r_0, r_1, r_2, \ldots, r_{2N-2})$ is a response vector for the dynamical system (3) if and only if the matrix C^T (with $T = N$) defined by (9), (10) is positive definite.*

This theorem and formulas for the entries of Jacobi matrix obtained in [11] implies the following procedure of solving the truncated moment problem:

1) Calculate $(r_0, r_1, r_2, \ldots, r_{2N-2})$ from $\{s_0, s_1, \ldots, s_{2N-2}\}$ by using (8).

2) Recover $N \times N$ Jacobi matrix A^N using formulas for a_k, b_k from [11]

3) Recover spectral measure for finite Jacobi matrix A^N prescribing arbitrary selfadjoint condition at $n = N+1$. Or one can do

3') Extend Jacobi matrix A^N to *finite* Jacobi matrix A^M, $M > N$, prescribe arbitrary selfadjoint condition at $n = M + 1$ and recover spectral measure of A^M.

3") Extend Jacobi matrix A^N to *infinite* Jacobi matrix A, and recover spectral measure of A^M.

Every measure obtained in 3), 3′), 3″) gives a solution to the truncated moment problem. Below we propose a different approach: we recover the spectral measure corresponding to Jacobi matrix directly from moments (from the operator C^N), without recovering the Jacoi matrix itself.

Agreement 1. *We assume that controls $f \in \mathcal{F}^N$, $f = (f_0, \ldots, f_{N-1})$ are extended: $f = (f_{-1}, f_0, \ldots, f_{N-1}, f_N)$, where $f_1 = f_N = 0$.*

We introduce the special space of controls $\mathcal{F}_0^N = \{f \in \mathcal{F}^T \mid f_0 = 0\}$ and the operator $D : \mathcal{F}^T \mapsto \mathcal{F}^T$ acting by

$$(Df)_t = f_{t+1} + f_{t-1}.$$

The following statements can be easily proved using arguments from [11]:

Proposition 2. *The operator W^N maps $\mathcal{F}^N$ isomorphically onto $\mathcal{H}^N$ and $\mathcal{F}_0^N$ maps isomorphically onto $\mathcal{H}^{N-1}$.*

Proposition 3. *On the set $\mathcal{F}_0^T$ the following relation holds:*

$$W^N D f = D W^N f, \quad f \in \mathcal{F}_0^N. \tag{11}$$

Taking $f, g \in \mathcal{F}_0^T$ we can evaluate the quadratic form, bearing in mind (11):

$$\big(C^N Df, g\big)_{\mathcal{F}^N} = \big(W^N Df, W^N g\big)_{\mathcal{H}^N} = \big(DW^N f, W^N g\big)_{\mathcal{H}^{N-1}} = \big(A^{N-1} v^f, v^g\big)_{\mathcal{H}^{N-1}}. \tag{12}$$

The latter expression in (12) means that only A^{N-1} block from the whole matrix A^N is in use. Then it is possible to perform the spectral analysis of A^{N-1} using the classical variational approach, the controllability of the system (3) (see Proposition 2) and the representation (12), see also [4]. The spectral data of Jacobi matrix A^{N-1} with Dirichlet boundary condition at $n = N$ can be recovered by the following procedure:

1) The first eigenvalue is given by

$$\lambda_1^{N-1} = \min_{f\in\mathcal{F}_0^N,\,(C^N f,f)_{\mathcal{F}^N}=1} \left(C^N Df, f\right)_{\mathcal{F}^N}. \tag{13}$$

2) Let f^1, be the minimizer of (13), then

$$\rho_1 = \left(C^N f^1, f^1\right)_{\mathcal{F}^N}.$$

3) The second eigenvalue is given by

$$\lambda_2^{N-1} = \min_{\substack{f\in\mathcal{F}_0^N,(C^N f,f)_{\mathcal{F}^N}=1\\ (C^N f,f_l)_{\mathcal{F}^N}=0}} \left(C^N Df, f\right)_{\mathcal{F}^N}. \tag{14}$$

4) Let f^2, be the minimizer of (14), then

$$\rho_2 = \left(C^T f^2, f^2\right)_{\mathcal{F}^T}.$$

Continuing this procedure, we recover the set $\{\lambda_k^{N-1}, \rho_k\}_{k=1}^{N-1}$ and construct the measure $d\rho^{N-1}(\lambda)$ by (6).

Remark 3. *The measure, constructed by the above procedure solves the truncated moment problem for the set of moments* $\{s_0, s_1, \ldots, s_{2N-4}\}$.

2.1 Euler-Lagrange equations

In this section we derive equations which can be thought of as a Euler-Lagrange equations for the problem of the minimization of a functional $\left(C^N Df, f\right)_{\mathcal{F}^N}$ in $\mathcal{F}_0^N$ with the constrain $\left(C^T f, f\right)_{\mathcal{F}^N} = 1$ described in the previous section. Similar method of deriving equations which can be used for recovering of spectral data was used in [3].

By f^k, $k = 1, \ldots, N$ we denote the control that drive system (3) to prescribed state (see (4)):

$$W^T f_k = \phi^k, \quad k = 1, \ldots, N.$$

Due to Proposition 2, such a control exists and is unique for every k. We introduce the operator

$$D_-^N : \mathcal{F}^N \mapsto \mathcal{F}^N,$$
$$\left(D_-^N f\right)_n = f_{n-1}, \quad n = 1, \ldots, N-1, \quad \left(D_-^N f\right)_0 = 0,$$

and denote by $P : \mathcal{F}^{N+1} \mapsto \mathcal{F}^N$ the embedding. Then $P^* : \mathcal{F}^N \mapsto \mathcal{F}^{N+1}$ extends vector by zero: $(P^* f)_N = 0$.

Theorem 2. *The spectrum of* A^N *and (non-normalized) controls* f_k, $k = 1, \ldots, N$ *are the spectrum and the eigenvectors of the following generalized spectral problem:*

$$\left(P\left(D_-^{N+1}\right)^* C^{N+1} P^* + C^N D_-^N\right) f_k = \lambda_k C^N f_k, \quad k = 1, \ldots, N. \tag{15}$$

Proof. For $h \in \mathcal{F}^T$ we always assume that $h_{-1} = h_T = 0$ (see Agreement 1). For a fixed $k = 1, \ldots, N$ we take $f_k \in \mathcal{F}^N$ such that $W^N f_k = v_{\cdot,N}^{f_k} = \phi^k$, then for arbitrary $g \in \mathcal{F}^N$ we can evaluate:

$$\begin{aligned}\left(\lambda_k C^N f_k, g\right)_{\mathcal{F}^N} &= \left(\lambda_k v_{\cdot,N}^{f_k}, v_{\cdot,N}^g\right)_{\mathcal{H}^N} = \left(\lambda_k \phi^k, v_{\cdot,N}^g\right)_{\mathcal{H}^N} = \left(A^N \phi^k, v_{\cdot,N}^g\right)_{\mathcal{H}^N} = \left(\left(A^N v^{f_k}\right)_{\cdot,N}, v_{\cdot,N}^g\right)_{\mathcal{H}^N} \\ &= \left(\left(Dv^{f_k}\right)_{\cdot,N}, v_{\cdot,N}^g\right)_{\mathcal{H}^N} = \left(v_{\cdot,N+1}^{f_k}, v_{\cdot,N}^g\right)\Big)_{\mathcal{H}^N} + \left(v_{\cdot,N-1}^{f_k}, v_{\cdot,N}^g\right)_{\mathcal{H}^N}\end{aligned} \tag{16}$$

We note that

$$\mathcal{H}^N \ni v_{\cdot,N}^g = \left(v_{1,N}^g(T), \ldots, v_{N,N}^g\right),$$
$$\mathcal{H}^{N+1} \ni v_{\cdot,N+1}^{D_-^{N+1} g} = \left(v_{1,N}^g, \ldots, v_{N,N}^g, 0\right).$$

That is why we can rewrite the first summand in the right hand side of (16) as

$$\left(v_{\cdot,N+1}^{f_k}, v_{\cdot,N}^g\right)_{\mathcal{H}^N} = \left(v_{\cdot,N+1}^{f_k}, v_{\cdot,N+1}^{D_-^{N+1} g}\right)_{\mathcal{H}^{N+1}} = \left(C^{N+1} f_k, D_-^{N+1} g\right)_{\mathcal{F}^{N+1}} \tag{17}$$

Analogously:

$$\mathcal{H}^N \ni v^{f_k}_{\cdot,\,N-1} = \left(v^{f_k}_{1,\,N-1}, \dots, v^{f_k}_{N-1,N-1}, 0\right),$$
$$\mathcal{H}^N \ni v^{D_-^T f_k}_{\cdot,\,N} = \left(v^{f_k}_{1,\,N-1}, \dots, v^{f_k}_{N-1,N-1}, 0\right).$$

So we can rewrite the second summand in the right hand side of (16) as

$$\left(v^{f_k}_{\cdot,\,N-1}, v^{g}_{\cdot,\,N}\right)_{\mathcal{H}^N} = \left(v^{D_-^T f_k}_{\cdot,\,N}, v^{g}_{\cdot,\,N}\right)_{\mathcal{H}^N} = \left(C^N D_-^N f_k, g\right)_{\mathcal{F}^N}. \tag{18}$$

Finally from (16), (17) and (18) we deduce that

$$\left(\lambda_k C^N f_k, g\right)_{\mathcal{F}^N} = \left(C^{N+1} f_k, D_-^{N+1} g\right)_{\mathcal{F}^{N+1}} + \left(C^N D_-^N f_k, g\right)_{\mathcal{F}^N}. \tag{19}$$

Using operators P, P^* we can rewrite (19) in the form

$$\left(P\left(D_-^{T+1}\right)^* C^{T+1} P^* + C^T D_-^T\right) f_k = \lambda_k C^T f_k.$$

Thus the pair $\{f_k, \lambda_k\}$ gives the solution to (15). Now let the pair $\{f, \lambda\}$ be the solution to (15) with $f \in \mathcal{F}^N$ $f \neq f_k$, $\lambda \neq \lambda_k$ for any $k = 1 \dots, N$. Then $W^N f = v^f_{\cdot,\,N} = \sum_{k=1}^N a_k \phi^k$ for some $a_k \in \mathbb{R}$. We can evaluate for arbitrary $g \in \mathcal{F}^N$:

$$0 = \left(\left(P\left(D_-^{N+1}\right)^* C^{N+1} P^* + C^N D_-^N\right) f - \lambda C^N f, g\right)_{\mathcal{F}^N} = \left(C^{N+1} P^* f, D_-^{N+1} P^* g\right)_{\mathcal{F}^{N+1}}$$
$$+ \left(C^N D_-^N f, g\right)_{\mathcal{F}^N} - \lambda \left(v^f_{\cdot,\,N}, v^g_{\cdot,\,N}\right)_{\mathcal{H}^N} = \left(v^{P^* f}_{\cdot,\,N+1}, v^{D_-^{N+1} P^* g}_{\cdot,\,N+1}\right)_{\mathcal{H}^{N+1}} + \left(v^{D_-^N f}_{\cdot,\,N}, v^g_{\cdot,\,N}\right)_{\mathcal{H}^N}$$
$$-\lambda \left(v^f_{\cdot,\,N}, v^g_{\cdot,\,N}\right)_{\mathcal{H}^N} = \left(v^{P^* f}_{\cdot,\,N+1}, v^g_{\cdot,\,N+1}\right)_{\mathcal{H}^N} + \left(v^f_{\cdot,\,N-1}, v^g_{\cdot,\,N-1}\right)_{\mathcal{H}^N} - \lambda \left(v^f_{\cdot,\,N-1}, v^g_{\cdot,\,N-1}\right)_{\mathcal{H}^N}$$
$$= \left(\left(A^N v^f\right)_{\cdot,\,N}, v^g_{\cdot,N}\right)_{\mathcal{H}^N} - \lambda \left(v^f_{\cdot,\,N}, v^g_{\cdot,\,N}\right)_{\mathcal{H}^N} = \left(A^N \sum_{k=1}^N a_k \phi^k - \lambda \sum_{k=1}^N a_k \phi^k, W^N g\right)_{\mathcal{H}^N} =$$
$$\left(\sum_{k=1}^N a_k (\lambda_k - \lambda) \phi^k, W^N g\right)_{\mathcal{H}^N}$$

From the above equality and Proposition 2 it follows that all a_k except one are equal to zero, and for such a_j, $\lambda = \lambda_j$, which completes the proof. □

Having found spectrum and non-nomalized controls from (15) we can recover the measure of A^N with Dirichlet boundary condition at $n = N + 1$ by the following procedure:

1) Normalize controls by choosing $\left(C^N f_k, f_k\right)_{\mathcal{F}^N} = 1$,
2) Observe that $W^N f^k = \alpha_k \phi^k$ for some $\alpha_k \in \mathbb{R}$, where the constant is defined by $\alpha_k = (R f_k)_N$.
3) The norming coefficients are given by $\rho_k = \alpha_k^2$, $k = 1, \dots, N$.
4) Recover the measure by (6).

2.2 Conclusion

Denote by M_N the subset of Borel measures on $\mathbb{R}$, such that $d\nu(\lambda) \in M_N$ is a solution of the truncated moment problem (1) of the order N. We use the Boundary control method to construct the *special solution* of a truncated Hamburger moment problem: for $N \in \mathbb{N}$ the set of moments $\{s_0, s_1, \dots, s_{2N}\}$ determines the measure $d\rho^N(\lambda) \in M_N$, where the constructed measure is a spectral measure of a finite Jacobi operator with Dirichlet condition at the right end. We point out that in our procedure we do not use the Jacobi matrix, but rather special matrix, constructed from moments. It is not hard to see (from (9) and results from [11]) that M_N is a convex set, and $M_{N_1} \subseteq M_{N_2}$ when $N_1 > N_2$. Taking N to infinity we deduce that the set of solutions M_∞ of the moment problem (1) either convex, or it consists of one element $d\rho^N(\lambda) \to_{N\to\infty} d\rho(\lambda)$ (see also [1, 13]).

It is possible to apply the approach proposed to studying the special cases: Stieltjes and Hausdorff moment problems, and also to the questions on the uniqueness of solution to moment problems, that will be the subjects of forthcoming publications.

References

[1] Akhiezer N.I. The classical moment problem and some related questions in analysis. Edinburgh: Oliver and Boyd, 1965.

[2] Avdonin S., Mikhaylov V. The boundary control approach to inverse spectral theory// Inverse Problems. 2010, V. 26, iss. 4. 045009, 19pp.

[3] Avdonin S., Mikhaylov A., Mikhaylov V. On some applications of the Boundary Control method to spectral estimation and inverse problems// Nanosystems: Physics, Chemistry, Math. 2015. V. 6, iss. 1. P. 63–78.

[4] Belishev M. On relations between spectral and dynamical inverse data// J. Inverse Ill-Posed Probl. 2001. V. 9, iss. 6. P. 547–565.

[5] Belishev M. Recent progress in the boundary control method // Inverse Problems. 2007, V. 23.

[6] Belishev M. Boundary control and tomography of Riemannian manifolds (the BC-method)// Uspekhi Matem. Nauk. 2017. V. 72, iss. 4. P. 3–66, (in Russian).

[7] Belishev M., Mikhaylov V. Unified approach to classical equations of inverse problem theory// Journal of Inverse and Ill-Posed Problems. 2012. V. 20, iss. 4. P. 461–488.

[8] Kabanikhin S. Inverse and ill-posed problems. Theory and applications.//Inverse and Ill-posed Problems Series, 55. Berlin: Walter de Gruyter GmbH & Co. KG, 2012.

[9] Mikhaylov A., Mikhaylov V. Dynamical inverse problem for the discrete Schrödinger operator// Nanosystems: Physics, Chemistry, Mathematics. 2016. V. 7, iss. 5. P. 842-854.

[10] Mikhaylov A., Mikhaylov V. Boundary Control method and de Branges spaces. Schrödinger operator, Dirac system, discrete Schrödinger operator// Journal of Mathematical Analysis and Applications. 2018. V. 460, iss. 2. P. 927–953.

[11] Mikhaylov A., Mikhaylov V. Dynamic inverse problem for Jacobi matrices// Inverse Problems and Imaging. 2019. V. 13, iss. 3. P. 431–447.

[12] Mikhaylov A., Mikhaylov V., Simonov S. On the relationship between Weyl functions of Jacobi matrices and response vectors for special dynamical systems with discrete time// Mathematical Methods in the Applied Sciences. 2018. V. 41, iss. 16. P. 6401–6408.

[13] Simon B. The classical moment problem as a self-adjoint finite difference operator// Advances in Math. 1998. V. 137, P. 82–203.

Mikhaylov Alexander Sergeevich — PhD, senior researcher of the St. Petersburg Department of V.A. Steklov Institute of Mathematics RAS
e-mail: mikhaylov@pdmi.ras.ru;
Mikhaylov Victor Sergeevich — PhD, senior researcher of the St. Petersburg Department of V.A. Steklov Institute of Mathematics RAS
e-mail: ftvsm78@gmail.com;
Received — April 30, 2019

АЛГОРИТМЫ МЕТОДА МОНТЕ–КАРЛО ДЛЯ ИССЛЕДОВАНИЯ ВРЕМЕННОЙ АСИМПТОТИКИ ПОТОКА ЧАСТИЦ С РАЗМНОЖЕНИЕМ В СЛУЧАЙНОЙ СРЕДЕ

Г. А. Михайлов[1,2], Г. З. Лотова[1,2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.676
DOI: 10.24411/9999-016A-2019-10055

Статья посвящена исследованию плотности числа частиц, рассеивающихся и размножающихся в случайной среде, в зависимости от времени. С этой целью построены и численно апробированы рандомизированные алгоритмы метода Монте–Карло для оценки вероятностных моментов параметра соответствующей экспоненциальной асимптотики.

Ключевые слова: метод Монте–Карло, весовые модификации, теория переноса, временная постоянная размножения частиц.

Вводная информация

Рассматривается односкоростной процесс переноса частиц с изотропным рассеянием. В качестве математической модели процесса используется цепь Маркова $\{x_n\}$, $n = 0, 1, \ldots, N$, соответствующих "столкновений" в фазовом пространстве $X = R \times V \times T$ координат, скоростей (с фиксированным значением v модуля) и времени, т. е. $x_n = (\mathbf{r}_n, \mathbf{v}_n, t_n)$, причем $\mathbf{v}_n = v(\mathbf{r}_n - \mathbf{r}_{n-1})/|\mathbf{r}_n - \mathbf{r}_{n-1}|$, $t_n = t_{n-1} + \left|\mathbf{r}_{n-1} - \mathbf{r}_n\right|/v$. Эта модель определяется плотностью $f(x)$ точки x_0 и плотностью $k(x', x)$ перехода из состояния x' в x, причем

$$\int\limits_X k(x', x)\mathrm{d}x = q(x') \le 1 - \delta, \qquad \delta > 0, \tag{1.1}$$

и тем самым среднее число переходов конечно, например, в ограниченной системе [1]. Заданы $\sigma_s(x)$ и $\sigma_f(x)$ — коэффициенты рассеяния и деления, среднее число ν частиц в точке деления, $\sigma_c(x)$ — коэффициент поглощения. Полный коэффициент ослабления потока частиц $\sigma(x) = \sigma_s(x) + \sigma_f(x) + \sigma_c(x)$ (см. [1]).

Методы Монте–Карло используют для оценки линейных функционалов вида $J_h = (\varphi, h)$, $h \in L_\infty$, где $\varphi(x)$ — плотность столкновений, удовлетворяющая уравнению $\varphi = K\varphi + f$, в котором K — интегральный оператор с ядром $k(x', x)$. При выполнении условия (1.1) имеем $||K||_{L_1} < 1$ и, следовательно, спектральный радиус оператора $\rho(K) < 1$ (см.[1]). Для построения весовой модификации алгоритма используют цепь Маркова с начальной плотностью $f_0(x)$ и плотностью перехода $p(x', x)$, вспомогательные веса $Q_0 = f(x_0)/f_0(x_0)$, $Q_n = Q_{n-1}k(x_{n-1}, x_n)/p(x_{n-1}, x_n)$ и "оценку по столкновениям" $\xi = \sum\limits_{n=0}^{N} Q_n h(x_n)$. Если выполняются "условия несмещенности" [1], то $\mathrm{E}\,\xi = (\varphi, h)$. Если при этом $\rho(K_p) < 1$, где K_p — оператор с ядром $k^2(x', x)\Big/p(x', x)$, и $f^2/f_0 \in L_1(X)$, то $\mathrm{D}\xi < +\infty$ [1].

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (коды проектов 18-01-00599, 17-01-00823, 18-01-00356).

ISBN 978-5-901548-42-4

Пусть $\varphi_0(x;\mathbf{r}_0,\mathbf{v}_0)$ — плотность столкновений (по аргументу x) от одного столкновения в точке $(\mathbf{r}_0,\mathbf{v}_0,0)$, т. е. для $f(x)=\delta(\mathbf{r}-\mathbf{r}_0)\delta(\mathbf{v}-\mathbf{v}_0)\delta(t)$.

Функционал

$$J(t)=\int\limits_R\int\limits_V \varphi(\mathbf{r},\mathbf{v},t)h(\mathbf{r},\mathbf{v})\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{v},\quad \forall f\in L_1(X),\quad h\in L_\infty(R\times V)$$

можно представить [2] в виде

$$J(t)=\int\limits_R\int\limits_V\int\limits_0^\infty f(\mathbf{r}_0,\mathbf{v}_0,t-\tau)F(\mathbf{r}_0,\mathbf{v}_0,\tau)\mathrm{d}\mathbf{r}_0\,\mathrm{d}\mathbf{v}_0\,\mathrm{d}\tau,$$

где

$$F(\mathbf{r}_0,\mathbf{v}_0,t)=\int\limits_R\int\limits_V \varphi_0(\mathbf{r},\mathbf{v},t;\mathbf{r}_0,\mathbf{v}_0)h(\mathbf{r},\mathbf{v})\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{v}.$$

Предполагается, что $f(\mathbf{r},\mathbf{v},-t)\equiv 0$, и вследствие этого $F(\mathbf{r},\mathbf{v},-t)\equiv 0$, при $t>0$.

Символом $f_t^{(m)}$ далее обозначается m-кратная производная от функции f по t.

Лемма 1. [2] *Пусть точка* $(\mathbf{r}_0,\mathbf{v}_0)$ *распределена для* $t_0\equiv 0$ *с плотностью* $f_0(\mathbf{r},\mathbf{v})$,

$$|f^{(m)}(\mathbf{r},\mathbf{v},t)\Big/f_0(\mathbf{r},\mathbf{v})|<C<+\infty,\quad m=0,1,\ldots,n,\quad \rho(K_p)<1$$

и $F(x)<C<+\infty$. *Тогда выполняется соотношение* $J^{(m)}(t)=\mathrm{E}\xi_t^{(m)}$, *где*

$$\xi_t^{(m)}=\sum_{n=0}^{N} Q_n h(\mathbf{r}_n,\mathbf{v}_n)f^{(m)}(\mathbf{r}_0,\mathbf{v}_0,t-t_n)\Big/f_0(\mathbf{r}_0,\mathbf{v}_0),\quad Q_0\equiv 1,$$

причем $\mathrm{D}\xi_t^{(m)}<+\infty, m=0,1,\ldots n$.

1 Оценка параметра экспоненциальной асимптотики по времени

Известно, что при выполнении довольно общих условий в случае источника, локализованного в точке $(\mathbf{r}_0,\mathbf{v}_0,0)$, имеет место асимптотическое при $t\to\infty$ соотношение [3,4]:

$$F(\mathbf{r},\mathbf{v},t)\sim C(\mathbf{r},\mathbf{v})\mathrm{e}^{\lambda t},\quad C(\mathbf{r},\mathbf{v})<C_0<+\infty, \tag{2.1}$$

где λ — ведущее характеристическое число соответствующего однородного стационарного уравнения переноса с заменой $\sigma_c\mapsto\sigma_c+\lambda\Big/v$. Эти условия, в частности, имеют место для односкоростного процесса переноса частиц в ограниченной среде с достаточно быстро убывающей по времени плотностью источника.

Теорема 1. *Если выполняются соотношения*

$$\int f^{(n)}(\mathbf{r},\mathbf{v},t)\mathrm{e}^{-\lambda t}dt<+\infty,\quad n=0,1, \tag{2.2}$$

соотношение (2.1) *и условия леммы* 1, *то*

$$\frac{J'(t)}{J(t)}\to\lambda\ \text{ при }\ t\to+\infty. \tag{2.3}$$

Доказательство. Прямое интегрирование с учетом соотношений (2.1), (2.2) дает равенство

$$J(t)=C\mathrm{e}^{\lambda t}[1+\varepsilon(t)],\quad J'(t)=C_1\mathrm{e}^{\lambda t}[1+\varepsilon_1(t)],$$

причем $\varepsilon(t)$, $\varepsilon_1(t)\to 0$ для $t\to+\infty$. Интегрируя функцию $J'(t)$ в пределах $(\tau,+\infty)$ при $\tau\to\infty$ для $\lambda<0$ получаем $J'(t)=C\lambda\mathrm{e}^{\lambda t}[1+\varepsilon_1(t)]$, т. е $C_1=\lambda C$. В случае $\lambda>0$ тот же результат получается путем введения дополнительного поглощения с коэффициентом $\sigma_c^{(0)}>\frac{\lambda}{|\mathbf{v}|}$. Это завершает доказательство теоремы 1.

2 Оценка величин $\mathrm{E}\lambda$ и $\mathrm{D}\lambda$ для случайной среды

Определяемую леммой 1 и теоремой 1 оценку метода Монте–Карло можно рандомизировать с целью изучения флуктуаций значения λ для процесса переноса частиц в случайной среде [2]. Будем полагать $\sigma(\mathbf{r})$ — случайное поле, причем отношения σ_s/σ, σ_f/σ, а также индикатрисы рассеяния и деления фиксированы. При этом $J(t) \equiv J(t,\sigma)$, $J'(t) \equiv J'(t,\sigma)$ и $\lambda(\sigma) \approx J'(t,\sigma)/J(t,\sigma)$ соответственно теореме 1. В дальнейшем аргумент t будет опускаться, т. е. $J(t,\sigma) \mapsto J(\sigma)$.

Практически важными являются величины $\mathrm{E}\lambda(\sigma)$ и $\mathrm{D}\lambda(\sigma)$. Логически простейший (прямой) способ их оценки методом Монте–Карло состоит в реализации достаточно точных оценок величины $J'(t,\sigma)/J(t,\sigma)$ для выборки σ достаточно большого объема [5]. Однако такой способ может быть слишком трудоемким для реалистических моделей среды и процесса переноса. Поэтому в настоящей работе для построения рандомизированного алгоритма используется соотношение

$$\mathrm{E}\frac{J'(\sigma)}{J(\sigma)} \approx \lambda_n = \mathrm{E}\left[J'(\sigma)\sum_{s=0}^{n}\frac{(-1)^s}{J_0^{s+1}}(J(\sigma)-J_0)^s\right],$$

где $J_0 \approx \mathrm{E}J(\sigma)$.

Простейшая (элементарная) несмещенная рандомизированная оценка величины λ_n на основе леммы 1 строится путем реализации $n+1$ условно независимых траекторий частиц для выбранного значения σ : $\lambda_n = \mathrm{E}\tilde{\lambda}_n$, где

$$\tilde{\lambda}_n = \xi'(\Omega_{n+1},\sigma)\sum_{s=0}^{n}\frac{(-1)^n}{J_0^{s+1}}\prod_{k=1}^{s}(\xi(\Omega_k,\sigma)-J_0)\,. \tag{3.1}$$

Для осреднения произведений из (3.1) целесообразно использовать несовпадающие сочетания по s элементов из группы $\Omega_1,\dots\Omega_{i-1},\Omega_{i+1},\dots\Omega_{n+1}$ после подстановки $\Omega_i \to \Omega_{n+1}$ в (3.1), в связи с тем, что в качестве выделенной $(n+1)$ - ой траектории можно рассматривать любую из совокупности $\{\Omega_i\}$, $i = 1,\dots,n+1$. Соответствующие различным сочетаниям произведения и суммы наиболее просто вычислять последовательным пересчетом соответственно возрастанию s. Получаемую таким образом статистическую оценку обозначим через $\tilde{\lambda}_n^{(i)}$. Пусть $\tilde{\lambda}_n^* = \frac{1}{n+1}\sum\limits_{i=1}^{n+1}\tilde{\lambda}_n^{(i)}$.

Теорема 2. *Если выполняются условия леммы 1, причем* $\rho\,(K_p(\sigma)) < 1-\varepsilon$, $(\varepsilon > 0)$ $\forall\sigma$, *то* $\mathrm{E}\tilde{\lambda}_n^* = \lambda_n$ *и* $\mathrm{D}\tilde{\lambda}_n^* \le \mathrm{D}\tilde{\lambda}_n < +\infty$.

Доказательство. Доказательство строится путем прямого осреднения и оценок на основе сформулированных выше утверждений и условной независимости траекторий $\Omega_1,\dots,\Omega_{n+1}$.

Оценку $\tilde{\lambda}_n$ естественно называть комбинированной. Отметим, что величина $\mathrm{D}\lambda_n^*$ может быть существенно меньше величины $\mathrm{D}\tilde{\lambda}_n$ при $\mathrm{D}\sigma \ll 1$ вследствие слабой коррелированности реализаций произведений в $\tilde{\lambda}_n^*$ для фиксированных значений s.

Аналогичная несмещенная элементарная оценка на основе леммы 1 для $\mu_n \approx \mathrm{E}\left[J'(\sigma)/J(\sigma)\right]^2$ строится по формуле

$$\tilde{\mu}_n = \xi'(\Omega_{n+1},\sigma)\cdot\xi'(\Omega_{n+2},\sigma)\sum_{s=0}^{n}\frac{(-1)^s(s+1)}{J_0^{s+2}}\prod_{k=1}^{s}(\xi(\Omega_k,\sigma)-J_0)$$

с использованием $n+2$ условно независимых траекторий. Более эффективная комбинированная оценка $\tilde{\mu}_n^* = \sum\limits_{i=1}^{n+1}\mu_n^{(i)}/(n+1)$ строится по аналогии с $\tilde{\lambda}_n^*$ для фиксированного значения $\xi'(\Omega_{n+2},\sigma)$.

3 Численные результаты

4.1. Для численного тестирования был рассмотрен односкоростной процесс переноса частиц с изотропным рассеянием (в том числе и после акта деления) в шаре радиуса $R = 7.72043$ с параметрами: σ — случайная величина, равномерно распределенная в интервале (0.7, 1.3);

$$\frac{\sigma_s}{\sigma} \equiv 0.97; \quad \frac{\sigma_f}{\sigma} \equiv 0.03; \quad \nu \equiv 2.5; \quad \sigma_c \equiv 0; \quad v = |\mathbf{v}| = 1.$$

Как указано в [2], такой шар при $\sigma \equiv 1$ с достаточной степенью точности критичен. Соответствующая прецизионная оценка величины λ методом Монте–Карло подтверждается расчетами в улучшенном диффузионном

приближении [6], которые дали значение $\lambda_d(\sigma = 1) = 0.000000\ldots$. С использованием этого приближения путем моделирования значений σ были получены следующие тестовые статистические оценки:

$$\mathrm{E}\lambda_d = -0.00103, \quad \mathrm{D}\lambda_d = 0.00022.$$

Среднеквадратическая погрешность имеет здесь порядок последнего десятичного разряда. Для построения эффективного алгоритма метода Монте–Карло на основе теоремы 2 в сформулированную модель было введено поглощение с постоянным неслучайным коэффициентом σ_c/v, который приводит к замене $\lambda \mapsto \lambda - \sigma_c/v\ \forall\sigma$. Отметим, что такой прием является универсальным и может существенно повысить эффективность весового метода, исключая необходимость ветвления моделируемых траекторий, как это видно из дальнейшего.

Используя уравнение переноса (см., например [1]), можно сделать замену $\sigma_f \mapsto \sigma_f + \sigma_c$, $\nu \mapsto \nu\sigma_f/(\sigma_f + \sigma_c)$. В численном эксперименте моделировался процесс переноса с константами: $\sigma_1^* = \sigma_1$, $\sigma_f^* = \sigma_f + \sigma_c$, $\nu^* = 1$. Вспомогательные веса при этом определяются формулой $Q_n = \big(\nu\sigma_f/(\sigma_f + \sigma_c)\big)^{n_1}$, где n_1 — число делений, предшествующих данному столкновению [1]. Было использовано значение $\sigma_c = 0.059$, для которого $\nu\sigma_f/(\sigma_f + \sigma_c) < 1$ и тем самым [1] $\rho(K_p) < 1\ \ \forall\sigma$. Плотность распределения первых столкновений взята в виде

$$f(\mathbf{r}, t) = 2t\exp(-2t)f_d(\mathbf{r}), \quad t > 0, \quad r = |\mathbf{r}| < R, \tag{4.1}$$

где $f_d(\mathbf{r}) = C\sin(ær)/r$ — диффузионное приближение к пространственной характеристической функции для $\sigma = 1, æ = 0.3739866$ [6]. Полагали также $h(\mathbf{r}, \mathbf{v}) \equiv \sin(æ(\rho)r)/r$, где (см. [6])

$$æ(\rho) = \frac{\pi(\sigma_s + \nu\sigma_f)\rho}{R(\sigma_s + \nu\sigma_f) + \alpha}.$$

Расчеты показали, что использование таких функциональных параметров алгоритма существенно улучшает сходимость в (2.3) сравнительно с $f_d(\mathbf{r}) \equiv \left(\frac{4}{3}\pi R^3\right)^{-1}$ и даже сравнительно с вариантом, в котором $f(\mathbf{r}, t)$ определяется формулой (4.1), а $h(\mathbf{r}, \mathbf{v}) \equiv 1$.

Нетрудно проверить, что для сформулированных выше характеристик вычислительной модели выполняются условия теорем 1, 2 и она тем самым позволяет построить оценку значений $\mathrm{E}\lambda$ и $\mathrm{D}\lambda$. Для построения статистических оценок этих величин было реализовано $16 \cdot 10^9$ случайных значений оптической плотности σ; моделирование траекторий осуществлялось параллельно на восьми процессорах Intel Core i7-3930K. При этом полагали $n = 5$, то есть для оценки $\mathrm{E}\lambda$ использовали шесть, а для оценки $\mathrm{E}\lambda^2$ — семь условно-независимых траекторий частиц. Результативные оценки определялись на основе анализа установления (с учетом вероятностной погрешности) двух старших значащих цифр при возрастании параметров t и s. Таким образом получены следующие результативные оценки: $\mathrm{E}\lambda = 0.00104 \pm 0.00001$ и $\mathrm{D}\lambda = 0.00020 \pm 0.00007$ для $t = 10$, $s = 4$. В качестве погрешности здесь приведено среднеквадратическое отклонение. Эти оценки вполне соответствуют диффузионным, достаточная точность которых проверена лишь для $\sigma = 1$.

Отметим, что полученные результаты переносятся на случай многоскоростного процесса переноса с анизотропным рассеянием соответственно [2].

4.2. Для проведения более сложных тестовых расчетов рассматривалась сферически симметричная среда с кусочно-постоянной случайной плотностью $\rho = \rho(r)$ в шаре радиуса $R = 7.72043$ с макроскопическими сечениями $\rho\sigma^{(0)}$, $\rho\sigma_s^{(0)}$, $\rho\sigma_f^{(0)}$, где

$$\sigma^{(0)} = 1,\ \sigma_s^{(0)} = 0.97,\ \sigma_f^{(0)} = 0.03,\ \nu = 2.5,$$

Для построения реализации среды шар делится на m одинаковых по объему сферических слоев, в каждом слое случайная величина $\rho \equiv \rho_i$ выбирается независимо и равномерно на отрезке $[1 - \varepsilon, 1 + \varepsilon]$.

Для такой системы в таблице 4.1 представлены результаты расчетов величины $\mathrm{E}k$ из [7] с помощью специальной гомогенизации (то-есть путем осреднения реализаций $\rho(r)$ с весом "диффузионной плотности" $g(r)$) и улучшенного диффузионного приближения; отметим, что в [7] рассматривалась величина $k = \rho(K)$ — коэффициент размножения по поколениям столкновений. В таблице 4.1 приведены также оценки величин $\mathrm{E}\lambda$, $\mathrm{D}\lambda$ с использованием этого приближения. Даны также основные здесь оценки, полученные методом Монте–Карло согласно изложенному в разделе 3. Представляет определенный интерес тот факт, что при $m = 6$ имеем $\mathrm{E}\lambda > 0$, хотя $\mathrm{E}k < 1$; результаты из [8] показывают, что при этом коэффициент размножения по поколениям делений $k_{eff} > 1$.

Таблица 4.1: Оценки $\mathrm{E}\lambda$ и $\mathrm{D}\lambda$

	Гомогенизация			Метод Монте–Карло		
m	$\mathrm{E}k$	$\mathrm{E}\lambda$	$\sqrt{\mathrm{D}\lambda}$	$\mathrm{E}k$	$\mathrm{E}\lambda$	$\sqrt{\mathrm{D}\lambda}$
1	0.9969	−0.00103	0.015	0,9968	−0.00104	0.014
2	0.9981	−0.00063	0.012	0.9983	−0.00017	0.014
6	0.993	−0.00022	0.0069	0.9994	0.00020	0.007

Некоторое занижение оценок величины $\mathrm{E}\lambda$, полученных методом стохастической гомогенизации, можно объяснить тем, что в этом методе фактически реализуется линеаризация по значениям плотности среды.

Список литературы

[1] Марчук Г.И., Михайлов Г. А., Назаралиев М.А. и др. Метод Монте-Карло в атмосферной оптике. – Новосибирск: Наука, 1976, 283 с.

[2] Лотова Г. З., Михайлов Г. А. Новые методы Монте-Карло для решения нестационарных задач теории переноса излучения // Журн. вычисл. матем. и матем. физики. 2002. Т. 42, № 4. С. 569–579.

[3] Дэвисон Б. Теория переноса нейтронов. – М.: Атомиздат, 1960, 514 с.

[4] Марчук Г.И. Методы расчета ядерных реакторов. М.: Госатомиздат, 1961.

[5] Larmier C., Zoia A., Malvagi F., Dumonteil E., Mazzolo A. Neutron multiplication in random media: Reactivity and kinetics parameters // Annals of Nuclear Energy. 2018. V. 111. P.391–406.

[6] Романов Ю. А. Точные решения односкоростного кинетического уравнения и их использование для расчета диффузионных задач // Исследование критических параметров реакторных систем. М.: Атомиздат, 1960. С. 3–26.

[7] Лотова Г.З., Михайлов Г.А. Методы Монте-Карло для оценки вероятностных распределений параметров критичности процесса переноса частиц в случайно возмущенной среде // Журн. вычисл. матем. и матем. физики. 2018. Т. 58. № 11, С. 1900–1910.

[8] Лотова Г.З., Михайлов Г.А. Моменты параметров критичности процесса переноса частиц в случайной среде //Журн. вычисл. матем. и матем. физики. 2008. Т. 48. № 12, С. 2225–2236.

Михайлов Геннадий Алексеевич — чл. корр. РАН, Советник РАН Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: gam@sscc.ru;
Лотова Галия Зуфаровна — к.ф.-м.н., старший научный сотрудник Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: lot@osmf.sscc.ru;
Дата поступления — 1 июня 2019 г.

МОДЕЛИРОВАНИЕ СТРУКТУРЫ И СТАБИЛЬНОСТИ БИМЕТАЛЛИЧЕСКИХ НАНОВОЛОКОН МЕТОДОМ МОНТЕ-КАРЛО

В. С. Мясниченко[1], Н. Ю. Сдобняков[1], Л. Кирилов[2], Р. Михов[2]

[1] *Тверской государственный университет, 170100, Тверь, Россия*

[2] *Институт информационных и коммуникационных технологий Болгарской Академии Наук, София, Болгария*

УДК 519.87: 538.911
DOI: 10.24411/9999-016A-2019-10056

Представлен метод оптимизации металлических наноструктур. Основой является решеточный метод Монте-Карло с различными решетками в сочетании с применением молекулярной динамики. Метод позволяет решать задачи с периодическими граничными условиями, может использоваться для моделирования одномерных и двумерных атомных структур. Если периодические граничные условия не заданы, то предполагаются конечные размеры модельной решетки. Кроме того, автоматическая релаксация кристаллической решетки может быть выполнена с целью минимизации потенциальной энергии системы. Разработана компьютерная реализация метода, использующая распространенный формат XYZ для описания атомарных структур и передачи входных параметров. Обнаружено, что наиболее стабильная конфигурация наносплава в виде нанопроволоки золото-серебро имеет Au-обогащенный приповерхностный слой.

Ключевые слова: глобальная оптимизация, симуляция отжига, атомная сегрегация, неустойчивость Рэлея-Плато.

Введение

Описание механизмов формирования внутренней структуры наночастиц и нановолокон позволяет прогнозировать их физико-химические свойства. Стабильная упорядоченная конфигурация нановолокна обладает наименьшей потенциальной энергией. Минимизация поверхностной энергии, вызванная термически активированной самодиффузией, приводит к поперечному разрыву нановолокон. Это наблюдалось для серебра [1], золота [2], а также для других переходных металлов.

Обзор последних исследований биметаллических наносплавов и нанокристаллов представлен в [3, 4]. Обсуждаются структурные характеристики нанокристаллов и различные подходы к их синтезу с контролируемыми свойствами. Также перечислены сферы применения в индустрии. В [5] исследованы механические и термодинамические свойства золотых нанопроводов для применения в наноразмерной электронике.

В работе [6] показано превращение тонких пленок Ag в наночастицы после одноимпульсного лазерно-индуцированного плавления. Единственным фактором для среднего размера частиц является начальная толщина пленки. Рэлеевская неустойчивость ультратонких золотых нанопроводов изучена в [7]. Авторы используют *in situ* просвечивающую электронную микроскопию (ПЭМ) для наблюдения за их поведением в реальных условиях эксплуатации. В [8] также изучена рэлеевская неустойчивость ультратонких (диаметром менее 10 нм) золотых нанопроводов. Авторы предлагают простой механический процесс для их самовосстановления после повреждения.

В [9] представлены исследования ПЭМ методом наноразмерных частиц $NiAu$ типа ядро-оболочка на нагреваемых сетках ПЭМ. При этом авторы отметили, что более репрезентативная оценка оптимального акта диффузии одного атома Ni в свободном икосаэдрическом кластере золота должна основываться на

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (проект № 18-03-00132), а также поддержке Министерства образования и науки Респ. Болгария (акроним проекта ICTinSES).

ISBN 978-5-901548-42-4

статистических подходах, таких как выборка по методу Монте-Карло. Поэтому компьютерное моделирование биметаллических наносистем является более сложным с точки зрения интерпретации результатов, по сравнению с однокомпонентными системами атомов.

1 Решеточный метод Монте-Карло

Поиск устойчивых конфигураций (и даже сравнение топологий двух конфигураций) наноструктур металлов и сплавов является NP-полной задачей с вычислительной точки зрения. Задача поиска эквивалентна минимизации поверхности потенциальной энергии с чрезвычайно большим количеством локальных минимумов. Сложность глобальной оптимизации атомных кластеров еще больше возрастает для наносплавов, где перестановка неодинаковых атомов также приводит к еще большему количеству изомеров [10]. Нецелесообразно применение численных методов, основанных на получении и переборе всех возможных минимумов, потому что они требуют огромных вычислительных ресурсов, таких как время и оперативная память [11]. Глобальный минимум аппроксимируется с помощью эффективных по времени стратегий оптимизации, так называемых метаэвристик. Существует множество методов для предсказания структуры наночастиц [12].

Для оптимизации структуры биметаллических одномерных нановолокон мы применяем алгоритмический подход имитации отжига с линейным убыванием температуры. Предлагаемая адаптация метода имеет следующие отличительных особенности. Сначала используется решетчатый метод Монте-Карло, универсальный относительно вида и размера решетки. Затем заполненная решетка подвергается автоматическому растяжению–сжатию для получения более оптимального решения.

Периодические граничные условия (1)–(2) позволяют использовать метод единым образом для моделирования атомных кластеров, одномерных (нанопроволока, трубки) и двумерных (нанопленочных) структур. Если периодические граничные условия отсутствуют, мы предполагаем конечные размеры модельной решетки.

$$r_{ij} = \sqrt{|\Delta x_{ij}|^2 + |\Delta y_{ij}|^2 + |\Delta z_{ij}|^2} \tag{1}$$

$$|\Delta x_{ij}| = \min\{|x_i - x_j|, L_x - |x_i - x_j|\} \tag{2}$$

Полученные структуры наночастиц релаксируются при низкой температуре в рамках молекулярной динамики, и один из изомеров выбирается в качестве аппроксимации глобального минимума.

Основной алгоритм метода состоит из следующих шагов:

1. Считать входные данные: начальные положения атомов, размеры окна для периодических граничных условий, другие параметры управления и т. д. Атомы, которые не имеют заданного начального положения, разместить случайным образом.

2. Для всех узлов предварительно вычислить списки соседей, окрестности и другую информацию, которая известна заранее.

3. Проверить критерий выхода. Цикл останавливается при превышении требуемого числа итераций или при достижении системой равновесия. Если да, перейти к шагу 11.

4. Отрегулировать температуру по следующей формуле:

$$T = \max\{1, T_0 + s\Delta T\} \tag{3}$$

   где T_0 и ΔT — константы, а s-номер итерации. Эта проверка выполняется раз в несколько тысяч итераций.

5. Выбрать атом с произвольным номером.

6. Выбрать соседний пустой узел. Если пустых соседей нет, вернуться к шагу 3.

7. Вычислить разность потенциальной энергии для атома, движущегося в выбранный пустой узел, с учетом периодических граничных условий.

8. Если энергия не будет увеличиваться, выполнить прыжок и вернуться к шагу 3.

9. В противном случае вычислить вероятность прыжка $P = \exp(-E/kT)$ и сгенерировать случайное число p $(0 \leq p < 1)$.

10. Если число p меньше вероятности, выполнить прыжок. В любом случае, вернуться к шагу 3.

11. Выполнить итеративное растяжение и сжатие решетки независимо вдоль каждой оси с шагом 0.01 (значение входного параметра по умолчанию) для дальнейшей минимизации потенциальной энергии.

Успешное применение имитации отжига в решающей степени зависит от правильного регулирования температуры, и одним из его наиболее важных параметров является начальная температура. Она должна быть достаточно высока, чтобы система могла исследовать различные глобальные конфигурации, не попадая в локальный минимум слишком рано. Но более низкая начальная температура позволяет получить решение с меньшим количеством итераций по методу Монте-Карло.

Обычно требуется параллельно запускать алгоритм имитации отжига несколько раз с различными последовательностями случайных чисел в надежде на лучшее решение. Это представляет компромисс: если общее время выполнения является постоянным, более низкое значение позволит выполнить больше испытаний. Другим важным параметром является степень заполнения узлов решеточной модели. Более детально метод и подбор его параметров описан в наших работах [13, 14].

2 Реализация численного метода

В данной работе мы применяем расширенный на одномерные биметаллические наноструктуры вариант метода имитации отжига [14]. В процессе медленного и контролируемого охлаждения атомы стремятся попасть в состояние с меньшей суммарной энергией. Однако, с определенной вероятностью они могут перейти и в состояние с большей энергией. Эта вероятность прыжка P уменьшается вместе с температурой.

Взаимодействие между атомами вычисляется с использованием многочастичного полуэмпирического потенциала Cleri & Rosato [15, 16]. Данный потенциал корректно описывает динамику решетки ГЦК и ГПУ металлов, а также их поведение при ненулевой температуре. Сравнение полученной энергии производится после охлаждения методом молекулярной динамики до 0,01 К.

В ходе реализации решалось несколько важных вопросов, вытекающих из особенностей предлагаемого метода и характера решаемой задачи. Мы минимизируем необходимые вычисления во время основного цикла, так как он должен выполняться для миллионов итераций. С этой целью широко использовались предварительные вычисления и кэширование.

Предварительные вычисления (предкомпиляция) позволяют выполнять каждый шаг основного цикла за постоянное время. Для расчета разности энергий “наивным способом” потребовалось бы перебирать не только $r-$окрестности выбранного атома и выбранного пустого узла, но и $r-$окрестности их окрестностей, так как учитываются не только парные межатомные взаимодействия. Предкомпиляция С позволяет вычислить правильную энергию, перебирая только первые окрестности.

Структуры данных организованы следующим образом. Узлы хранятся в двух массивах M и A, где M дает индекс узла в A; и в обратную сторону: A дает индекс узла в M. Массив M сортируется во входном порядке, в то время как A сортируется, чтобы начать с атомов и закончить пустыми узлами. Это позволяет нам запрашивать информацию об узлах и атомах, выбирать атомы случайным образом, добавлять, удалять и перемещать атомы, всё это за постоянное время.

Входные данные задаются в виде списка узлов, заданных их декартовыми координатами. Каждый узел может быть пустым или содержать атом определённого сорта. Расстояние между соседними узлами во входных данных могут незначительно отличаться (в пределах 15%). Используется распространённый текстовый формат *.XYZ для описания атомарных структур.

Другие аспекты задачи моделирования, такие как число и тип атомов металла, периодических или непериодических граничных условий, температуры системы и т. д. указываются в качестве дополнительных параметров. Они могут быть заданы либо в командной строке, либо во второй строке (строка “комментария”) входного файла *XYZ*. Выходные файлы имеют идентичный формат, при этом значения энергии системы и всех остальных параметров отображаются в строке “комментария”. Это позволяет выполнять несколько последовательных запусков программы на одних и тех же данных. Можно также вывести промежуточные результаты вычислений.

Программа написана на языке программирования C и работоспособна на Windows и Linux платформах. Поскольку предлагаемый метод не является детерминированным, для обеспечения стабильности поведения на каждом наборе данных выполняется несколько запусков алгоритма.

3 Результаты численных экспериментов

Для детального изучения структурообразования был проведен ряд экспериментов. Использовалась решетка ГЦК с 28830 узлами, линейными размерами 11,9 x 6,1 x 6,1 нм. Также использовалась квазикристаллическая решетка с осью симметрии пятого порядка. На ось X налагались периодические граничные условия.

Составы были одинаковыми: чистое золото и биметалл $Au - Ag$ в соотношении 1:1. Каждый состав был представлен 6-ю вариантами: 1000, 2000, 4000, 6000, 8000 и 16000 атомов. Было выполнено моделирование Монте-Карло с 360 миллионами итераций, с охлаждением от 2200K до 1K. Для каждого варианта проводилась серия из десяти экспериментов.

На рис. 1 приведены изображения срезов биметаллических нановолокон (нанопроволок), имевших различное число атомов в расчётном блоке, и соответственно различную степень заполнения узлов атомами η_{or} при использовании решетки одного размера. Количество одиночных нульмерных фрагментов (кластеров) для систем из одного металла (золота) и биметалла отдельно показано на рисунке 2. Как можно наблюдать, в случае моносостава образуется больше атомных кластеров, и средний размер каждого кластера меньше, чем в случае биметаллического состава. Когда степень заполнения узлов η_{or} превышает 25%, в обоих случаях образуются только непрерывные нанопроволоки.

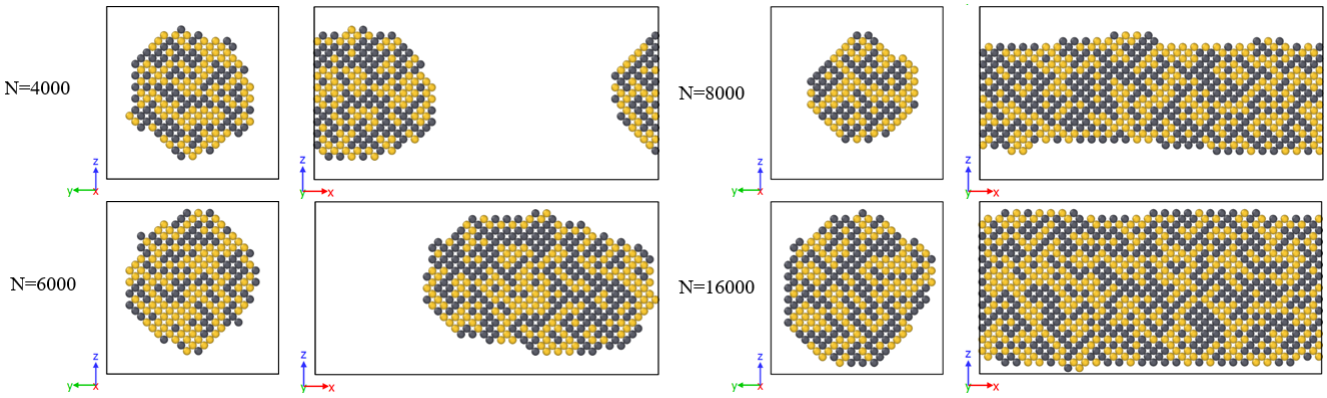


Рис. 1: Расположение атомов золота (желтого) и серебра (темно-серого цвета) в нанопроволоке различной толщины. Для каждого размера показаны сечения поверхностью (100) и поверхностью (010)

Исследовалось распределение атомов в получаемых биметаллических нановолокнах. Концентрационные колебания состава вблизи поверхности наблюдались в наночастицах и нановолокнах $AuAg_3$ и $AuAg$ [17] [18], где самый наружный слой был богат Ag, но приповерхностный слой богат Au. Подобные особенности атомного упорядочения наблюдаются (рис. 3) в результатах наших МК экспериментов. Например, хотя атомы Ag выделяются на поверхность, под поверхностью находятся в основном атомы Au (60-80%), а внутренний слой остается структурой сплава.

Координационные числа Z для графика на рис. 3 были рассчитаны суммарно по пяти координационным сферам с использованием радиуса обрезания 6,8A. Поверхностными считались атомамы с $Z \leq 50$ соседей. Приповерхностный слой состоит из атомов с координационным числом $Z = 51 - 72$ и показан на рис. 4 красным цветом. Как видно на рис. 3 и рис. 4, наблюдаемый эффект поверхностной сегрегации не зависит от размера и формы (единое нановолокно либо отдельные фрагменты) биметаллической наноструктуры.

Заключение

Предложен адаптированный решеточный метод Монте-Карло для моделирования одномерных и двумерных биметаллических наноструктур. Автоматическое растяжение/сжатие решетки делается для того, чтобы

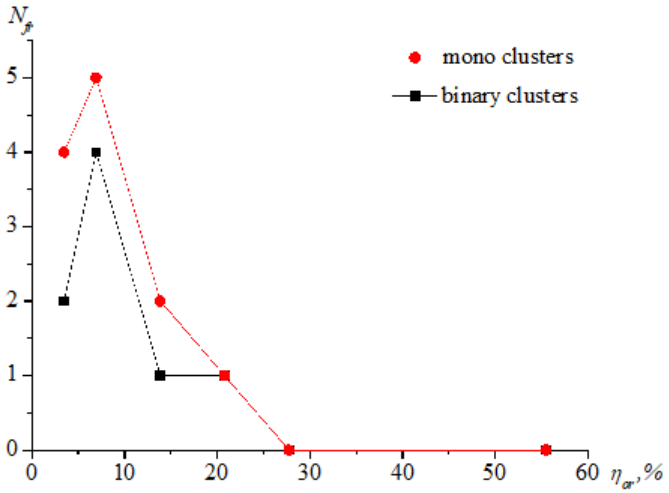


Рис. 2: Зависимость числа кластерных фрагментов N_{fr} от степени заполнения одномерной периодической решетки для моно- и биметаллических нановолокон

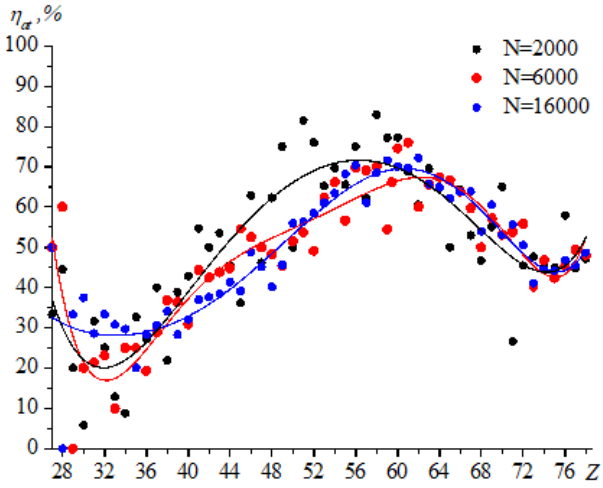


Рис. 3: Зависимость доли атомов золота от координационного числа Z для блоков нановолокна с разным количеством атомов N. Сплошные кривые — полиномиальные приближения для соответствующих данных

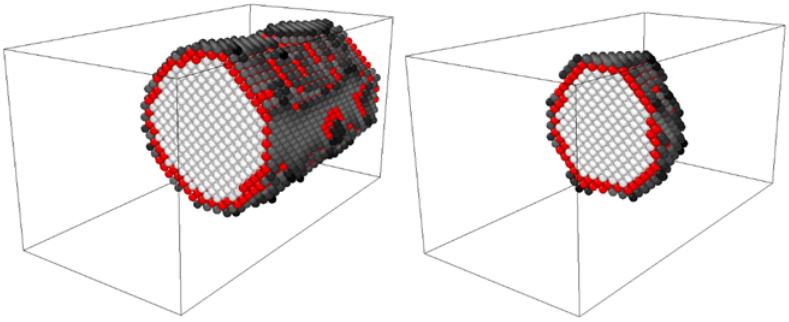


Рис. 4: Вид в разрезе нановолокна $N = 16000$ и кластера $N = 6000$ с цветовым кодированием: белый цвет соответствует атомам ядра, красный цвет — подповерхностным, а серый — поверхностным атомам

найти лучшее оптимальное решение; полученные нанопроволоки релаксируются при низкой температуре в молекулярной динамике.

При использовании описанного метода для изучения поведения сегрегации и структурных особенностей $AuAg$ нанопроволок, поверхностная сегрегация Ag обнаружена во всех рассмотренных наносистемах. Наиболее термодинамически стабильными являются биметаллические структуры с Ag-обогащенной поверхностью и Au-обогащенный приповерхностный слой. При степени заполнения решетки менее 25% наблюдалось разделение на отдельные наночастицы вследствие неустойчивости Рэлея-Плато.

Список литературы

[1] Langley D.P. et al. Metallic nanowire networks: effects of thermal annealing on electrical resistance. // Nanoscale. 2014. Vol. 6, № 22. P. 13535–13543.

[2] Karim S. et al. Morphological evolution of Au nanowires controlled by Rayleigh instability // Nanotechnology. 2006. Vol. 17, № 24. P. 5954–5959.

[3] Ferrando R., Fortunelli A., Johnston R.L. Searching for the optimum structures of alloy nanoclusters // Phys. Chem. Chem. Phys. The Royal Society of Chemistry, 2008. Vol. 10, № 5. P. 640–649.

[4] Gilroy K.D. et al. Bimetallic nanocrystals: Syntheses, properties, and applications // Chem. Rev. 2016. Vol. 116, № 18. P. 10414–10472.

[5] Naik J.P. et al. Liquid-like behaviour of gold nanowire bridges // Appl. Phys. Lett. 2017. Vol. 111, № 7.

[6] Oh Y., Lee M. Single-pulse transformation of Ag thin film into nanoparticles via laser-induced dewetting // Appl. Surf. Sci. Elsevier B.V., 2017. Vol. 399. P. 555–564.

[7] Xu S., Li P., Lu Y. In situ atomic-scale analysis of Rayleigh instability in ultrathin gold nanowires // Nano Res. 2018. Vol. 11, № 2. P. 625–632.

[8] Wang B. et al. Mechanically Assisted Self-Healing of Ultrathin Gold Nanowires // Small. 2018. Vol. 14, № 20. P. 1–7.

[9] Schnedlitz M. et al. Stability of Core-Shell Nanoparticles for Catalysis at Elevated Temperatures: Structural Inversion in the Ni-Au System Observed at Atomic Resolution // Chem. Mater. 2018. Vol. 30, № 3. P. 1113–1120.

[10] Lloyd L.D. et al. Theoretical investigation of isomer stability in platinum-palladium nanoalloy clusters // J. Mater. Chem. 2004. Vol. 14, № 11. P. 1691–1704.

[11] Doye J.P.K. Physical Perspectives on the Global Optimization of Atomic Clusters // Global Optimization. Springer US, 2006. Vol. 85, № November. P. 103–139.

[12] Gelfand S.B., Mitter S.K. Metropolis-Type Annealing Algorithms for Global Optimization in $\mathbb{R}^d$ // SIAM J. Control Optim. 1993. Vol. 31, № 1. P. 111–131.

[13] Myasnichenko V. et al. Monte Carlo Approach for Modeling and Optimization of One-Dimensional Bimetallic Nanostructures // The 9th International Conference on Numerical Methods and Applications NM&A'18. Lecture Notes in Computer Science. Borovetz, 2018. P. 133–141.

[14] Myasnichenko V. et al. Simulated Annealing Method for Metal Nanoparticle Structures Optimization // Advanced Computing in Industrial Mathematics. BGSIAM 2017. Studies in Computational Intelligence, vol 793. 2019. P. 277–289.

[15] Cleri F., Rosato V. Tight-binding potentials for transition metals and alloys // Phys. Rev. B. American Physical Society, 1993. Vol. 48, № 1. P. 22–33.

[16] Li J. et al. Interatomic potentials of the binary transition metal systems and some applications in materials physics // Phys. Rep. Elsevier, 2008. Vol. 455, № 1–3. P. 1–134.

[17] Deng L. et al. Au-Ag bimetallic nanoparticles: Surface segregation and atomic-scale structure // J. Phys. Chem. C. 2011.

[18] Wei L. et al. Surface segregation of Au-Ag bimetallic nanowires // Comput. Mater. Sci. 2013. Vol. 69. P. 374–380.

Мясниченко Владимир Сергеевич — науч. сотр. каф. общей физики Тверского государственного университета;
e-mail: virtson@gmail.com;
Сдобняков Николай Юрьевич — к.ф.-м.н., доцент каф. общей физики Тверского государственного университета;
e-mail: nsdobnyakov@mail.ru;
Кирилов Леонид М. — PhD, доцент Института информационных и коммуникационных технологий Болгарской Академии Наук;
e-mail: lkirilov@iinf.bas.bg;
Михов Россен — мл. науч. сотр Института информационных и коммуникационных технологий Болгарской Академии Наук;
e-mail: rmikhiv@bas.bg
Дата поступления — 01 июня 2019 г.

МЕТОДЫ ПРЕДОБУСЛАВЛИВАНИЯ, ОСНОВАННЫЕ НА ГРАФОВЫХ ПРЕОБРАЗОВАНИЯХ

Г. А. Омарова, Д. В. Перевозкин

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск

УДК 519.1+519.612
DOI: 10.24411/9999-016A-2019-10057

В работе продолжено исследование, начатое в [1] и дополненное [2]. Рассматриваемые методы предобуславливания основаны на алгоритмах построения остовных деревьев, где в качестве исходного графа выступает граф, построенный на матрице СЛАУ. Методы исследуются применительно к решению различных СЛАУ с квадратной вещественной матрицей. Приводятся результаты численных экспериментов.

Ключевые слова: предобуславливатели, графы, корневые деревья, остовные деревья.

Введение

Данная работа посвящена реализации и построению предобуславливателей с применением теории графов. В конце 80-х — начале 90-х годов вышли самые первые работы, в которых теоретико-графовые понятия (главным образом, пути) используются при анализе предобуславателей [3–7]. Эти идеи были расширены Вайдией [1], он описал свою работу в беседе в 1991 году, но не опубликовал статью. Вайдия использовал метод перехода от СЛАУ к теории графов для построения остовного дерева и создания семейства новых предобусловливателей на основе поддеревьев.

Позднее в начале 2000-х годов алгоритм Вайдия был реализован в [2] и экспериментально показано, что эти предобуславливатели обладают некоторыми замечательными свойствами для узкого класса матриц. Авторы показали, что в классе симметричных матриц с диагональным преобладанием вычислительная стоимость и скорость сходимости зависят от структуры матрицы, но не от ее числовых значений. Авторы провели ряд сравнений по эффективности работы предобуславливателей Вайдии с другими предобуславливателями, в частности, с неполным разложением Холецкого. Основные выводы, которые были получены, заключаются в том, что предобуславливатели Вайдии:

- обспечивают почти постоянную скорость сходимости по целому ряду задач, поведение очень отличается от поведения сходимости предобуславливателей на основе неполного разложения Холецкого;
- чувствительны только к портрету матрицы коэффициентов, а не к значениям элементов (внутри класса симметричных матриц с диагональным преобладанием, далее SDD-матриц), опять-таки по-другому, чем поведение неполного Холецкого;
- обеспечивают аналогичную или лучшую производительность по сравнению с неполным Холецким по широкому кругу двумерных проблем;
- обеспечивают более низкую производительность, чем неполный Холецкий, по некоторым трехмерным задачам, но значительно лучше работают на других трехмерных задачах.

В работе [8] некоторые свойства были доказаны для узкого класса матриц.

Целью нашей работы является исследование поведения работы предобуславливателя Вайдии на различных вещественных матрицах, в том числе несимметричных. В первом параграфе данной работы представлен алгоритм построения предобуславливателя, предложенного в [2]. Во втором параграфе рассматривается

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 18-01-00295 А).

ISBN 978-5-901548-42-4

модифицированный алгоритм, предложенный авторами данной статьи. В третьем параграфе приводятся результаты численных экспериментов.

1 Алгоритм построения предобуславливателя

Вайдя предложил построить предобуславливатель M для SDD-матрицы A, отбросив диагональные ненулевые значения от A и факторизуя M (полностью). В работе [2] авторы акцентируют свое внимание на применении предобуславливателей Вайдии к подклассу SDD-матриц с неположительными внедиагональными элементами и классу матриц SPDDD — диагонально-доминантным матрицам Стилтьеса. Построение предобуславливателя основано на 3-х этапах.

Входными данными является SPDDD-матрица $n \times n$, с $2m$ не нулевыми внедиагональными элементами и параметром t.

Пусть $A = \{a_{i,j}\}_{i,j=1}^{n}$ — SPDDD-матрица размерности $n \times n$. G_A — граф, построенный по данной матрице. $G_A = (V_A, E_A)$, где $V_A = \{1, 2, \ldots, n\}$ — множество вершин графа, $E_A = \{(i,j), i \neq j,\ a_{i,j} \neq 0\}$ — множество ребер, вес ребра равен значению $a_{i,j}$. Так как, матрица A симметричная, то построенный граф — неориентированный. Параметр t определяет число подграфов в методе для построения предобуславливателя. При $t = n$ предобуславливатель $M = A$, при $t = 1$ M имеет плохой результат сходимости. Цель данного метода построить предобуславливатель M по графу G_M, который является подграфом графа G_A. Граф G_M имеет тоже множество вершин $V_M = V_A$, а множество ребер является подмножеством E_A, $E_M \subset E_A$.

На первом шаге для графа G_A строится остовное дерево максимального веса. Требуется найти такое дерево графа, которое бы соединяло все его вершины, и при этом обладало наибольшим возможным весом (сумма весов рёбер дерева). Корневое дерево — это дерево с выделенной вершиной. Дерево - это набор рёбер, соединяющих все вершины, причём из любой вершины можно добраться до любой другой ровно одним простым путём. Для реализации этого шага используется алгоритм Прима [9], на выходе получают корневое дерево, заданное в массиве π. Целочисленный массив π размерности $2n$ для каждой вершины i содержит её родителя.

На втором шаге с помощью рекурсивной функции дерево разбивается на k поддеревьев с множествами вершин $V_1, V_2, \ldots, V_k$. Каждая вершина дерева имеет различную степень, поэтому разбиение происходит произвольно. Число вершин в каждом поддереве V_i варьируется между значениями n/t и $(dn/t) + 1$, где d — максимальная степень в дереве. В идеале хотелось бы разбить дерево на эквивалентные по количеству вершин поддеревья, но не для всех случаев это возможно. В работе [2] приводится доказательство теоремы подтверждающее утверждение не эквивалентности размерности полученных поддеревьев, в противном случае нарушается связность поддерева, а это противоречит идеи алгоритма.

На третьем шаге происходит построение матрицы M. Полученное корневое дерево T запишем, как граф G_M. Формируем G_M, добавляя к T самые тяжелые ребра между V_i и V_j для всех i и j. Ничего не добавляется, если между V_i и V_j нет ребер. Предобуславливатель M — это матрица, графом которой является G_M. Обозначим через M_t предобуславливатель Вайдия, построенный с параметром t, тогда $M_n = A$. Предобусловливатель M_1 состоит исключительно из остовного максимального веса дерева без добавленных ребер.

Также в [2] указано, что при построении матрицы M ее диагональные элементы выбираются таким образом, чтобы совпадали строчные суммы A и M, т.е. $Ae = Me$, где $e = (1, \ldots, 1)^T$.

2 Описание алгоритма

Построение предобуславливателя для наших матриц основано на 3-х этапах, но с некоторыми модификациями. Входными данными является произвольная матрица размерности n, с m не нулевыми внедиагональными элементами и параметром t.

Пусть $A = \{a_{i,j}\}_{i,j=1}^{n}$ — произвольная матрица размерности $n \times n$. G_A — граф, построенный по данной матрице. $G_A = (V_A, E_A)$, где $V_A = \{1, 2, \ldots, n\}$ — множество вершин графа, $E_A = \{(i,j), i \neq j,\ a_{i,j} \neq 0\}$ — множество ребер, вес ребра равен значению $a_{i,j}$. Так как, матрица A произвольная, то построенный граф ориентированный. Параметр t определяет число подграфов в методе для построения предобуславливателя. Цель данного метода построить предобуславливатель M по графу G_M, который является подграфом графа G_A.

- Восстанавливаем граф G_A до неориентированного графа следующим образом: вес ребра $w(i,j) = \max\{|a_{i,j}|, |a_{j,i}|\}$.

- Создаем массив ребер, где ребро задается в виде структуры: начало ребра, конец ребра, вес. Сортируем в порядке убывания.

- Реализуем алгоритм для нахождения подграфа G_M с максимальным весом и без циклов, на основе алгоритма Краскала. Алгоритм Краскала [10] предназначен для решения задачи о построении минимального остовного дерева во взвешенном неориентированном графе. Последовательная сложность алгоритма $O(m \log n)$. В отличие от алгоритмов Прима, алгоритм Краскала не требует информации о рёбрах конкретной вершины, вместо этого на его вход подаётся общий список рёбер графа в произвольном порядке. Кроме этого, каждое (ненаправленное) ребро достаточно представить лишь одной из его направленных дуг, что на практике означает в два раза меньший объём вычислений. На выходе получаем граф $G_M = (V_A, E_M)$, где $E_M \subset E_A$, который представляет собой лес, каждая компонента представляет дерево в виде списка ребер.

- Разбиваем полученный граф G_M на компоненты связности и в каждой компоненте выделяем корневую вершину. Реализация алгоритма основана на алгоритме поиска в ширину в графе, выбор основан на локализации вершин по уровню удаленности от корня. Поиск в ширину (англ. breadth-first search, BFS) — метод обхода графа и поиска пути в графе [11]. Если t меньше количества компонент связности, то мы получим вырожденный случай, как и при $t = 1$. При t больше числа компонент связности, делим компоненты с большим числом вершин на несколько частей, чтобы в одной компоненте было не больше чем n/t вершин. Выше приводились пояснения, что в различных поддеревьях не всегда одинаковое число вершин. Деление компоненты связности реализуется с помощью алгоритма поиска в глубину не рекурсивным обходом, так как в случае рекурсии при больших матрицах получаем переполнение стека. Сам же алгоритм позволяет сохранить связность. Поиск в глубину (англ. Depth-first search, DFS) — метод обхода графа [11]. Стратегия поиска в глубину, состоит в том, чтобы идти «вглубь» графа, насколько это возможно. На выходе получаем массив цветов, где вершины, попавшие в одну компоненту связности красим в один цвет.

- Построение матрицы M. Из графа, полученного в пункте 3, строим матрицу M следующим образом: берем ребро из полученного списка и добавляем его в матрицу M, аналогично, как и в матрице A. Далее формируем G_M, добавляя к T самые тяжелые ребра между V_i и V_j для всех i и j, используя матрицу цветов, и ничего не добавляется, если i и j имеют один цвет. Диагональному элементу $M_{i,i}$ соответствует сумма строк.

Предобуславливатель M — это матрица, графом которой является G_M. Обозначим через M_t предобусловливатель Вайдия, построенный с параметром t.

3 Численные результаты

В качестве СЛАУ для экспериментального исследования выбирались примеры из общедоступной коллекции Флоридского университета, имеющие правую часть. Последнее требование связано с тем, что скорость сходимости итерационных методов зависит не только от матрицы, но и от правой части.

В таблице 1 представлено количество итераций для различных методов. Через V10, V100, V1000 обозначены предобуславливатели Вайдия с параметром $t = 10,\ 100,\ 1000$. В правом столбце представлено количество итераций для метода ILU0. Также указан тип матриц, в данном случае SPD означает симметричную положительно определенную матрицу, а пропуск — произвольную квадратную вещественную матрицу.

Предобуславливатели Вайдия и ILU0 использовались в качестве предобуславливателей в методе FGMRes с периодом рестарта 100, ограничением в 1000 на общее число итераций и критерием выхода $\|Ax_i - b\|_2 / \|b\|_2 < 10^{-7}$, где x_i — текущее итерационное приближение, а b — правая часть.

Заключение

Эксперименты показывают, что даже для симметричных положительно определенных матриц метод Вайдия может не дать сходимости. В то же время видно, что область применимости этого метода не ограничивается лишь вышеуказанными матрицами. Также добавляет неясности вопрос выбора параметра t, практически не освещенный в литературе.

Таблица 1: Количество итераций

SLAE	type	V10	V100	V1000	ILU0
af_0_k101	SPD	1000	1000	1000	1000
nasa4704	SPD	1000	1000	1000	1000
offshore	SPD	1000	1000	1000	145
parabolic_fem	SPD	1000	820	262	1000
Pres_Poisson	SPD	1000	1000	1000	538
smt	SPD	1000	1000	1000	1000
thermal1	SPD	576	200	71	1000
hcircuit		14	13	12	–
jan99jac120		17	17	17	–

Список литературы

[1] P. M. Vaidya, Solving Linear Equations with Symmetric Diagonally Dominant Matrices by Constructing Good Preconditioners, manuscript, 1991.

[2] Chen D, Toledo S. Vaidya's preconditioners: Implementation and experimental study. ETNA, 16, 2003.

[3] R. Beauwens, Upper eigenvalue bounds for pencils of matrices, Linear Algebra Appl., 62 (1984), pp. 87–104.

[4] R. Beauwens, Approximate factorizations with modified S/P consistently ordered M-factors, Numer. Linear Algebra Appl., 1 (1994), pp. 3–17.

[5] Y. Notay, Solving positive (semi)definite linear systems by preconditioned iterative methods, in Preconditioned Conjugate Gradient Methods, Lecture Notes in Math. 1457, Springer-Verlag, New York, 1990, pp. 105–125.

[6] Y. Notay, Conditioning analysis of modified block incomplete factorizations, Linear Algebra Appl., 154–156 (1991), pp. 711–722.

[7] Y. Notay, Conditioning of Stieltjes matrices by S/P consistently ordered approximate factorizations, Appl. Numer. Math., 10 (1992), pp. 381–396.

[8] M. Bern, J. Gilbert, B. Hendrickson, N Nguyen, S. Toledo, Support-graph preconditioners, SIAM J. MATRIX ANAL. APPL. Vol. 27, No. 4, pp. 930–951, 2006, Society for Industrial and Applied Mathematics

[9] R. C. Prim, Shortest connection networks and some generalizations, Bell System Technical Journal, 36 (1957), pp. 1389–1401.

[10] Kruskal, Joseph B. "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem." Proceedings of the American Mathematical Society 7, no. 1 (January 1956): 48–50. doi:10.1090/S0002-9939-1956-0078686-7.

[11] Томас Х. Кормен, Чарльз И. Лейзерсон, Рональд Л. Ривест, Клиффорд Штайн. Алгоритмы: построение и анализ// 3-е издание — М.: «Вильямс», 2013. — 1328 с.

Омарова Гульзира Алимовна — к.ф.-м.н., н.с. Института вычислительной математики и математической геофизики СО РАН;
e-mail: gulzira@rav.sscc.ru;
Перевозкин Данил Валерьевич — м.н.с. Института вычислительной математики и математической геофизики СО РАН;
e-mail: foxillys@gmail.com.
Дата поступления — 30 апреля 2019 г.

КЛЕТОЧНО-АВТОМАТНОЕ МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННЫХ ПРОЦЕССОВ НА ТРИАНГУЛЯЦИОННЫХ СЕТКАХ

А. В. Павлова, С. Е. Рубцов, П. Р. Родионов

Кубанский государственный университет, 350040, Краснодар

УДК 510.67:554
DOI: 10.24411/9999-016A-2019-10058

Работа посвящена реализации пространственных КА-моделей диффузии для различных природных процессов на триангуляционных сетках. Полученные результаты могут быть использованы для построения более сложных композиционных КА, одним из правил перехода в которых является диффузия. Реализована КА-модель распространения фронта (поверхности, до которой дошел процесс распространения частиц). Для этого применен так называемый композиционный клеточный автомат, обычно используемый при моделировании явлений, включающих несколько различных процессов. Для случаев, когда при моделировании процесса требуется учитывать рельеф, применение триангуляционных сеток, позволяющих строить системы непересекающихся треугольников с вершинами в опорных точках поверхности, имеет преимущества, так как алгоритмы разбиения на треугольники имеют меньшую вычислительную сложность, нежели при использовании других полигонов.

Ключевые слова: клеточно-автоматное моделирование, трехмерная модель триангуляционная сетка, диффузия, распространения фронта.

Введение

В настоящее время развитием клеточно-автоматных моделей (КА-моделей) в плане возможности имитации различных физико-химических процессов на основе применения КА занимаются многие исследователи. Среди таких моделей можно выделить так называемые диффузионные, агрегационные и реакционные [1–6, и др.].

Область применения таких моделей постоянно расширяется. В частности, клеточные автоматы используются при создании моделей распространения фронта пожара и распределения концентрации продуктов горения с течением времени. Среди всего многообразия антропогенных и природных факторов, негативно воздействующих на состояние экосистемы, пожары можно выделить как один из самых разрушительных, наносящих огромный экологический и материальный ущерб. И использование КА-моделей служит расширению и развитию подходов к моделированию этого чрезвычайно сложного процесса, разработке эффективных в эксплуатации моделей.

При создании клеточно-автоматных моделей прямоугольные сетки получили наибольшее распространение. Но зачастую при моделировании того или иного процесса требуется учет формы поверхности, например, рельефа местности. В таких случаях применение триангуляционных сеток, позволяющих строить системы непересекающихся треугольников с вершинами в опорных точках поверхности, является более предпочтительным. Альтернатива в виде триангуляционных сеток имеет преимущества, так как алгоритмы разбиения на треугольники имеют меньшую вычислительную сложность, нежели при использовании других полигонов [5]. Рассмотрение КА-моделей на различных триангуляционных сетках позволяет имитировать сложные процессы различной природы, моделируемые совместной работой множества простых вычислителей, на криволинейных поверхностях в трехмерном пространстве. При этом треугольники в триангуляции не обязательно должны быть правильными и одинаковыми по площади. Эту особенность стоит принимать во

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований и администрации Краснодарского края (код проекта 19-41-230005).

ISBN 978-5-901548-42-4

внимание при реализации клеточных автоматов, требующих уточнения имитации процесса на определенных участках поверхности.

1 Построение клеточных-автоматов, имитирующих различные процессы, на триангуляционных сетках

В настоящей работе реализованы КА-алгоритмы диффузии и распространения фронта с использованием триангуляционных сеток. Рассматриваемая область пространства представляется клеточным массивом Ω с множеством имен $M = \{m_i : \ i = 1, \ldots, N\}$. При этом задан алфавит состояний $A = \{0, 1\}$ — булев алфавит. Каждой клетке ставится в соответствие пара (a, m), где $a \in A$, $m \in M$, именуемая в дальнейшем состоянием клетки и обозначаемая также $a(m)$. В клеточном пространстве конечных размеров $\Omega = \{a(m), m : \ m \in M\}$ определены подмножества, называемые шаблоном соседства для клетки (a, m), включающие близлежащие клетки. Каждый треугольник сетки может иметь не более трех соседей, при этом число соседей может быть меньше трех (для граничных клеток). При использовании триангуляционных сеток клетки считаются соседями (a, m_0) в случае, если соответствующие им треугольники имеют с нею общую сторону [5].

На множестве M клеточных имен определены так называемые именующие функции $f(m) : \ M \to M$. Их значения описывают местоположения клеток, взаимодействующих с заданной клеткой (a, m_0), т.е. для любой (a, m_0) именующие функции определяют ее ближайших соседей: $m_j = f_j(m_0)$, $j = \overline{1,3}$. Множество именующих функций определяет шаблон соседства: $T = \{m_0, f_1(m_0), f_2(m_0), f_3(m_0)\}$. Нумерация клеток в определенном таким образом шаблоне может быть любой. При этом принято считать, что $f_0(m_0) = m_0$.

При моделировании наивной диффузии режим функционирования КА асинхронный, т.е. за глобальную итерацию КА $\Omega(t) \to \Omega(t + 1)$ правило перехода будет последовательно применено в случайном порядке для каждого элемента клеточного массива. Правило перехода формулируется следующим образом: клетка меняется своим значением с одной из соседствующих ячеек, при этом вероятность обмена обратно пропорциональна площади соседа.

Для проведения вычислительных экспериментов создано приложение с использованием средств программирования 3D графики OpenGL, реализующее КА модель наивной диффузии и композиционные КА на различных триангуляционных сетках. Предлагаемый алгоритм программной реализации КА-диффузии позволяет легко иллюстрировать процесс на любой замкнутой или незамкнутой поверхности (рис. 1). Рис. 1 соответствует 1000 итераций работы автомата, реализующего наивную диффузию.

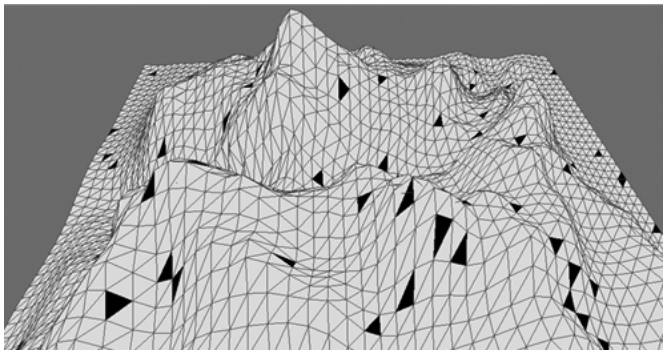


Рис. 1: Реализация наивной диффузии на произвольной незамкнутой поверхности

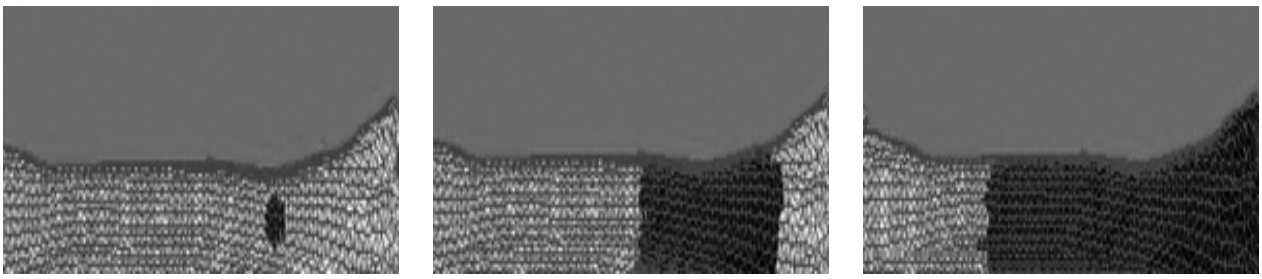
Созданное приложение предусматривает возможность перехода от булевых значений к непрерывным функциям распределения концентрации частиц, осуществляемого посредством осреднения значений состояний клеток автомата по задаваемой пользователем окрестности, например, осреднение по ближайшим соседствующим ячейкам. В таком случае подсчет клеток с состоянием «1» производится среди заданного числа соседей по формуле $\varphi = \frac{N_1}{N_2+1}$, где $N_2 \in \{0, 1, 2, 3\}$. Вычислительная сложность этого алгоритма составляет $O(n)$ [5], однако результат такого осреднения не очень нагляден, так как концентрация принимает лишь одно из семи значений $\phi \in \left\{0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1\right\}$. Осредненная концентрация будет отображаться более наглядно, если распространить алгоритм с первого на j-й уровень соседства, $j \geqslant 2$.

Полученные результаты могут быть применены для построения более сложных композиционных КА, в которых одним из правил перехода будет диффузия. Так, реализована КА-модель распространения фрон-

та (поверхности, до которой дошел процесс распространения частиц). Для этого применен так называемый композиционный клеточный автомат, обычно используемый при моделировании явлений, включающих несколько различных процессов [7]. При построении такого КА к клеткам массива на каждом такте применяется несколько правил, в данном случае следующих: после итерации диффузии для всего введенного клеточного пространства $\Omega = \{a(m), m : \ m \in M\}$ производится осреднение, т.е. определяются значения φ; затем эти значения заменяются в каждой клетке на $\psi = 0,5\varphi(1-\varphi)$, тем самым моделируется поток; далее производится операция, обратная осреднению, называемая дискретизацией, т.е. каждой клетке, содержащей значение $\psi \in [0,1]$, ставится в соответствие булево значение по формуле

$$a' = \begin{cases} 1, & \tau < \psi; \\ 0, & \tau \geqslant \psi, \end{cases}$$

где τ — случайное число из $[0,1]$. Описанный клеточный автомат при учете реальных особенностей процесса может, в частности, имитировать распространение пожара. На рис. 2 представлены результаты работы КА-модели распространения фронта на триангуляции (а — 0 итераций, б — 75 итераций, в — 200 итераций).



а) б) в)

Рис. 2: Результаты моделирования процесса диффузионного распространения фронта

Из рис. 2 видно, что фронт распространился равномерно во все стороны, охватывая со временем всю поверхность, что качественно соответствует физической природе процесса.

Заключение

В работе рассмотрены клеточные автоматы пространственных процессов, реализованные с использованием триангуляционных сеток. Созданное приложение позволило также осуществлять переход от булевых значений к непрерывным функциям путем осреднения по ближайшим соседствующим клеткам задаваемого уровня.

Для построения уникальных клеточных автоматов в рамках конкретных моделей для заданных условий и расчетных областей в дальнейшем будут реализованы механизмы генерации триангуляционных сеток с использованием карт высот. Предполагается также реализация построенных моделей, имитирующих эволюцию клеточных автоматов больших размеров, на многопроцессорных вычислителях, что позволит обеспечить их высокую производительность. Представленные результаты применимы для построения более сложных композиционных КА, включающих интерпретацию нескольких явлений.

Список литературы

[1] Boccara N. Reaction-Diffusion complex systems. Berlin: Springer, 2004.

[2] Bandman O.L. A method for construction of cellular automata simulating pattern formation processes // Theoretical background of applied discrete mathematics. 2010. Iss. 4. P. 91–99.

[3] Бандман О.Л. Отображение физических процессов на их клеточно-автоматные модели // Вестник томского государственного университета. 2008. № 2(3). С. 5–17.

[4] Бандман О.Л. Клеточно-автоматное моделирование пространственной динамики. Новосибирск: СО РАН, 2000.

[5] Евсеев А.А., Нечаева О.И. Клеточно-автоматное моделирование диффузионных процессов на триангуляционных сетках // Прикладная дискретная математика. 2009. № 4. С. 72–83.

[6] Павлова А.В., Рубцов С.Е. К клеточно-автоматным моделям конвекционно-диффузионных процессов // Марчуковские научные чтения – 2017: труды международной конференции по вычислительной и прикладной математике (ВПМ 2017). Новосибирск: ИВМиМГ СО РАН, 2017. С. 677–679.

[7] Рубцов С.Е., Павлова А.В., Родионов П.Р. К клеточно-автоматным моделям на триангуляционных сетках // Экологический вестник научных центров Черноморского экономического сотрудничества. 2018. № 2. С. 5–11.

Павлова Алла Владимировна — д.ф.-м.н., профессор кафедры математического моделирования Кубанского государственного университета;
e-mail: pavlova@math.kubsu.ru;
Рубцов Сергей Евгеньевич — к.ф.-м.н., доцент кафедры математического моделирования Кубанского государственного университета;
e-mail: ub_serg@mail.ru;
Родионов Павел Рольданович — студент Кубанского государственного университета;
e-mail: kmm@fpm.kubsu.ru.
Дата поступления — 30 апреля 2019 г.

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ ТРЕХСЛОЙНЫХ РАЗНОСТНЫХ СХЕМ ДЛЯ УРАВНЕНИЯ ШРЕДИНГЕРА С СИММЕТРИЧНЫМИ МНОГОТОЧЕЧНЫМИ ОПЕРАТОРАМИ

В. И. Паасонен

Институт вычислительных технологий СО РАН, 630090, Новосибирск
Новосибирский государственный университет, 630090, Новосибирск

УДК 519.6
DOI: 10.24411/9999-016A-2019-10059

Рассматриваются произвольные трехслойные симметричные компактные схемы с многоточечными операторами, коэффициенты которых определяются из условий аппроксимации с максимально возможным порядком по основной переменной. Для схем этого класса исследуется выполнение необходимого критерия устойчивости Неймана. В результате анализа сформулированы общие условия устойчивости, и на их основе для ряда схем, представляющих практический интерес, вычислены ограничения на соотношение шагов, при которых критерий Неймана выполняется. Показано также, что все схемы рассматриваемого класса подобно известной схеме Ричардсона в случае уравнения теплопроводности абсолютно неустойчивы.
Ключевые слова: компактная схема, условие устойчивости, уравнение Шредингера.

Введение

Для решения задач нелинейной волновой оптики [1] традиционные разностные схемы для достижения достаточной точности требуют существенной детализации сетки, что весьма расточительно по вычислительным затратам, поэтому чаще используются более точные спектральные методы [2], хотя разностные методы являются более гибкими и более простыми. В частности они легко допускают параллельную реализацию. Использование трехточечных компактных схем четвертого порядка точности [3–5] несколько улучшает ситуацию, но пока не обеспечивает уверенного преимущества, требуется еще больший порядок точности. Однако в рамках трехточечных шаблонов повышение порядка выше четвертого невозможно. В связи с этим представляется целесообразным выйти за пределы традиционного трехточечного шаблона в поисках устойчивых схем высших порядков точности на расширенных шаблонах. При этом на трехслойных шаблонах попутно легко достигается повышенный четвертый порядок аппроксимации также и по эволюционной переменной.

Использование многоточечных шаблонов сопряжено с необходимостью решать системы уравнений с матрицами ленточной структуры и ставить дополнительные граничные условия. Однако, во-первых, задачи с матрицами ленточной структуры по сложности и экономичности близки к обычной трехточечной прогонке и, во-вторых, среди задач волновой оптики имеются важные классы задач с периодическими решениями, с известной асимптотикой или с финитным носителем, для которых постановка дополнительных граничных условий не составляет проблемы. В данной работе с целью поиска эффективных высокоточных методов для решения таких задач исследуется устойчивость произвольных трехслойных схем максимального порядка аппроксимации на расширенных симметричных шаблонах.

Работа выполнена при частичной финансовой поддержке Российского научного фонда (код проекта 17-72-30006).

ISBN 978-5-901548-42-4

1 Постановка задачи

Итак, решается периодическая или краевая задача для нелинейного уравнения Шредингера:

$$i\,U_t + dU_{xx} + s|U|^2\,U = 0, \quad 0 < t < T, \quad x_1 < x < x_2,$$

где i — мнимая единица, t — эволюционная переменная, x — время (или конфигурационная переменная), d и s — вещественные коэффициенты. При $t = 0$ задается начальное данное $U(x, 0) = U_0(x)$.

Умножив уравнение на $-i$ приведем его к виду, совпадающую по форме с уравнением теплопроводности

$$U_t = cU_{xx} - f. \tag{1}$$

Отличие состоит в том, что в уравнении теплопроводности коэффициент c вещественный положительный, а в уравнении Шредингера $c = id$ — величина чисто мнимая, и правая часть имеет специальный вид. Кроме того, в рассматриваемых уравнениях наименования независимых переменных различны — в уравнении теплопроводности t и x — временная и пространственная переменная. С точки зрения аппроксимации различия между уравнениями носят лишь формальный характер, поэтому можно рассматривать оба варианта. Дисперсионные соотношения схем в обоих случаях также идентичны, хотя вытекающие из них условия невозрастания гармоник зависят от того, вещественное c или мнимое.

Введем равномерную сетку с шагом τ по t и шагом h по x. Рассмотрим трехслойную схему в канонической форме на произвольном шаблоне

$$A\frac{u^{n+1} - u^{n-1}}{2\tau} = c\frac{B}{h^2}u^n - g^n, \quad A = \sum_{i=-p}^{p} a_i T^i, \quad B = \sum_{j=-q}^{q} b_j T^j, \tag{2}$$

где произвольные целые числа $p \geq 0$, $q \geq 1$ определяют «габариты» разностных операторов схемы, а T — оператор сдвига ($Tu(x) = u(x+h)$). Коэффициенты схемы $a_i, (i = -p, -p+1, ..., p)$ и $b_j, (j = -q, -q+1, ..., q)$ будем определять из системы условий аппроксимации с максимально возможным порядком относительно h и с некоторым заданным порядком относительно τ.

Представим погрешность аппроксимации схемы (2) на достаточно гладких решениях дифференциального уравнения (1) в виде разложения по степеням шага τ:

$$\Psi = c\left\{A\left(u_{xx} + \frac{c^2\tau^2}{6}u_{xxxxxx} + O(\tau^4)\right) - \frac{B}{h^2}u\right\} +$$

$$+\left\{g^n - A\left(f + \frac{\tau^2}{6}(f_{tt} + cf_{xxt} + c^2 f_{xxxx})\right)\right\} + O(\tau^4).$$

Используя затем разложения разностных операторов A и B:

$$A = \sum_{m=0}^{\infty}\frac{h^m}{m!}A_m\frac{\partial^m}{\partial x^m}, \quad B = \sum_{m=0}^{\infty}\frac{h^m}{m!}B_m\frac{\partial^m}{\partial x^m}, \quad A_m = \sum_{i=-p}^{p} a_i i^m, \quad B_m = \sum_{j=-q}^{q} b_j j^m, \tag{3}$$

получим разложение погрешности также и по степеням h:

$$\Psi = c\left\{\sum_{m=0}^{\infty}A_m\frac{h^m}{m!}\frac{\partial^m}{\partial x^m}\left(u_{xx} + \frac{c^2\tau^2}{6}u_{xxxxxx}\right) - \sum_{m=0}^{\infty}B_m\frac{h^{m-2}}{m!}\frac{\partial^m u}{\partial x^m}\right\} + \tag{4}$$

$$+\left\{g^n - \sum_{m=0}^{\infty}A_m\frac{h^m}{m!}\frac{\partial^m}{\partial x^m}\left(f + \frac{\tau^2}{6}(f_{tt} + cf_{xxt} + c^2 f_{xxxx})\right)\right\} + O(\tau^4). \tag{5}$$

Для построения схемы максимального порядка аппроксимации по h вначале следует определить коэффициенты схемы (2) так, чтобы максимально возможное число первых членов разложения в первом слагаемом (4) обратилось в нуль, а после нахождения коэффициентов схемы необходимо правую часть схемы g^n задать как функционал от $f(x, t)$ таким образом, чтобы второе слагаемое (5) также приближенно обратилось в нуль, желательно с тем же порядком. В случае различных порядков малости этих двух слагаемых порядок погрешности в целом определится порядком менее точного приближения.

Ввиду симметрии шаблона схема максимального порядка обладает также симметрией в отношении коэффициентов, т. е. $a_{-i} = a_i$ $(i = 1, ...p)$ и $b_{-j} = b_j$ $(j = 1, ...q)$. Следовательно, выражения A_m, B_m из (3) при нечетных индексах m обращаются в нуль, а при четных имеют вид

$$A_0 = a_0 + 2\sum_{i=1}^{p} a_i, \quad B_0 = b_0 + 2\sum_{j=0}^{q} b_j, \quad A_{2m} = 2\sum_{i=1}^{p} a_i i^{2m}, \quad B_{2m} = 2\sum_{j=1}^{q} b_j j^{2m}, \quad m \geq 1. \tag{6}$$

Так как разностный оператор B/h^2 должен аппроксимировать оператор двойного дифференцирования по x, то следует потребовать выполнения равенства $B_0 = 0$, а из соображений нормировки коэффициент при f в разложении правой части схемы g^n должен быть равен единице, т. е. $A_0 = 1$.

Приводя подобные (с производными одинакового порядка) в первой скобке выражения (4), с учетом всего сказанного выше относительно специфики симметричного случая, получим систему условий (аппроксимации однородного уравнения)

$$B_0 = 0, \quad A_0 = 1, \tag{7}$$

$$\frac{A_{2m}}{(2m)!} + \frac{R^2}{3!}\frac{A_{2m-4}}{(2m-4)!} = \frac{B_{2m+2}}{(2m+2)!}, \quad m = 0, 1, ..., L. \tag{8}$$

Из контекста вывода условий аппроксимации непосредственно следует способ повышения погрешности относительно τ. Очевидно, для этого уравнение (8) следует заменить более сложным:

$$\frac{A_{2m}}{(2m)!} + \frac{R^2}{3!}\frac{A_{2m-4}}{(2m-4)!} + \frac{R^4}{5!}\frac{A_{2m-8}}{(2m-8)!} + \cdots + \frac{R^{2k}}{(2k+1)!}\frac{A_{2m-4k}}{(2m-4k)!} = \frac{B_{2m+2}}{(2m+2)!}, \quad m = 0, 1, ..., L, \tag{9}$$

при этом порядок погрешности по τ составит величину $2(k+1)$.

Уравнения в системах (8) и (9) при некоторых значениях m содержат факториалы отрицательных чисел и выражения A_k с отрицательными индексами k. Предполагается, что первые считаются равными единице, а вторые — нулю, тогда запись систем оказывается содержательно верной для любых $m \geq 0$.

Так как выражения (3) являются линейными функциями коэффициентов схемы a_i, b_j, то (7), (8) или (7), (9) есть система $L + 3$ линейных уравнений для определения $p + q + 2$ искомых коэффициентов схемы. Следовательно, если строки системы (3) линейно независимы, то максимально возможное значение $L = p + q - 1$. При этом, если система совместна, коэффициенты схемы определяются однозначно, и максимально возможный порядок аппроксимации по h равен $S = 2(p + q)$.

Предположим, что фиксированное k выбрано и система условий нормировки и аппроксимации (7), (9)) решена, т. е. коэффициенты схемы однозначно определены. Однако для аппроксимации неоднородного уравнения $(f \neq 0)$ требуется приближенное равенство нулю также и второй фигурной скобки в выражении погрешности, приведенного в (5) лишь для частного случая погрешности четвертого порядка относительно τ. В общем случае правая часть схемы g^n с погрешностью порядка $O(\tau^{2(k+1)})$ должна иметь такое же разложение, как функционал

$$g^n = \sum_{m=0}^{L} A_{2m}\frac{h^{2m}}{(2m)!}\frac{\partial^{2m} F}{\partial x^{2m}} + O(\tau^{2k+1} + h^{2(L+1)}), \tag{10}$$

где

$$F = F_0 + \frac{\tau^2}{3!}F_2 + \frac{\tau^4}{5!}F_4 + \cdots + \frac{\tau^{2k}}{(2k+1)!}F_{2k}, \quad F_{2m} = \sum_{r=0}^{m} c^r \frac{\partial^{r+m} f}{\partial x^{2r}\partial t^{m-r}}, \quad m = 0, 1, \ldots, k.$$

К сожалению, аппроксимировать правую часть с высокой точностью по τ не всегда удается, поскольку для этого необходимо знать выражения старших производных от правой части по t. При задании или при вычислении f как сеточной функции аппроксимация даже третьей производной по трем слоям невозможна, поэтому в общем случае максимально достижимый порядок по t лишь четвертый, когда

$$F = f + \frac{\tau^2}{6}(f_{tt} + cf_{txx} + c^2 f_{xxxx}).$$

Попытка повысить порядок до шестого вынуждает дополнить функционал слагаемым, содержащим третью производную по эволюционной переменной, для аппроксимации которой требуется минимум четыре слоя.

Элементарный анализ показывает, что в случае четвертого порядка аппроксимации по τ при любых заданных параметрах p, q числа узлов в шаблоне достаточно для аппроксимации правой части (1) с тем же порядком точности по h, т. е. с порядком $2(p + q)$.

2 Численное исследования устойчивости схем

Определение выражений коэффициентов схемы из системы (7), (9) в симметричном случае при небольших габаритах шаблона не представляет труда. С привлечением машинных символьных вычислений можно получить выражения коэффициентов практически для любого случая, однако при больших габаритах шаблона громоздкие выражения коэффициентов схемы вряд ли полезны. С другой стороны, схемы на слишком больших шаблонах не представляют и практической ценности. Интересны именно схемы с умеренными значениями габаритов (до пяти–семи, максимум до девяти узлов на слое). Для исследования их устойчивости воспользуемся прямыми вычислениями коэффициентов схемы и множителей возрастания гармоник. При фиксированных p и q коэффициенты схемы максимально возможного порядка аппроксимации зависят от единственного параметра — соотношения шагов $R = c\tau/h^2$, а собственные значения операторов A и B, соответствующие гармонике $\exp(i\omega x)$:

$$\lambda_A = 1 - 4\sum_{k=1}^{p} a_k \sin^2\frac{k\psi}{2}, \quad \lambda_B = -4\sum_{k=1}^{q} b_k \sin^2\frac{k\psi}{2}, \quad \psi = \omega h,$$

зависят только от R и от частоты ψ. Коэффициент возрастания гармоники $\rho = \rho(R, \psi)$ удовлетворяет дисперсионному уравнению

$$\rho^2 - 2\,C\,\rho - 1 = 0, \quad C = \frac{R\,\lambda_B}{\lambda_A},$$

корни которого имеют вид

$$\rho_\pm = C \pm \sqrt{C^2 + 1}.$$

Заметим, что в случае уравнения теплопроводности параметр $R = c\,\tau/h^2$ вещественный, и месте с ним число C также вещественно, поэтому один из корней дисперсионного уравнения непременно превосходит по модулю единицу. Следовательно, все рассматриваемые симметричные трехслойные схемы в случае уравнения теплопроводности абсолютно неустойчивы. Вероятно, именно по этой причине такие схемы не рассматривались прежде (как якобы заведомо бесперспективные) также и для уравнения Шредингера.

Однако в случае уравнения Шредингера ситуация совершенно иная. Коэффициенты схемы a_i, b_j вещественны (!), так как мнимый параметр $R = i\,d\,\tau/h^2$ входит в систему условий аппроксимации (7), (9) только в четных степенях. Ввиду симметрии схемы собственные числа операторов λ_A и λ_B также вещественны. Отсюда следует, что параметр C в выражении корней дисперсионного уравнения вместе с мнимым R также чисто мнимый. Следовательно

$$\rho_\pm = i\tilde{C} \pm \sqrt{1 - \tilde{C}^2}, \quad \tilde{C} = \frac{r\,\lambda_B}{\lambda_A}, \quad r = Im\,R.$$

При $|\hat{C}| > 1$ оба корня мнимые, причем один из них больше единицы по модулю, т. е. схема абсолютно неустойчива. Если же $|\hat{C}| \le 1$, то корни комплексно-сопряженные, и $|\rho_\pm| = 1$, т. е. схема устойчива, но не диссипативна. Таким образом, в случае уравнения Шредингера условием устойчивости схемы является неравенство

$$G(|R|, \psi) = |R|\frac{|\lambda_B|}{|\lambda_A|} \le 1 \quad \forall\,\psi. \in (0, 2\,\pi).$$

Нарушение этого условия влечет неустойчивость.

Очевидно, при фиксированных параметрах шаблона схемы p и q функция $z = G(|R|, \psi)$ зависит только от двух переменных, и устойчивость имеет место при тех значениях $|R|$, при которых $z \le 1$ для всего спектра частот. Функция $z = G(|R|, \psi)$ вычислялась на равномерной сетке по $|R|$ и ψ в области $(0 \le |R| \le R_{max}) \times (0 \le \psi \le 2\,\pi)$, причем промежуток $[0, R_{max}]$ изменения $|R|$ подбирался так, чтобы найти наиболее широкую область устойчивости. Там, где значения z превышали единицу, они обрезались до величины $1 + \varepsilon$. При изображении поверхности значения G, превышающие единицу, игнорировались, поэтому на рисунках белые (непрорисованные) зоны соответствовали областям, где критерий устойчивости Неймана не выполнялся.

Расчеты проводились для параметров шаблона схемы в пределах $0 \le p \le 4$, $1 \le q \le 4$. Этого более чем достаточно с практической точки зрения, так как в этот класс попадают, в частности, все явные и неявные максимально точные схемы, аппроксимирующие с четными порядками от второго до десятого относительно h, и, кроме того, целый ряд схем более точных вплоть до шестнадцатого. Критерий Неймана выполняется при ограничении вида $|R| \le M(p, q)$; соответствующие значения $M(p, q)$ (пороги устойчивости) приведены в таблице 1.

Таблица 1: Порог устойчивости максимально точных схем $O(\tau^4 + h^{2(p+q)})$

$q\backslash p$	0	1	2	3	4
1	0.251*	0.168*	0.170*	0.167*	0.167*
2	0.188*	0.170	0.133	0.166	0.166
3	0.176*	0.167	0.166	0.092	0.166
4	0.171*	0.166	0.166	0.166	0.072

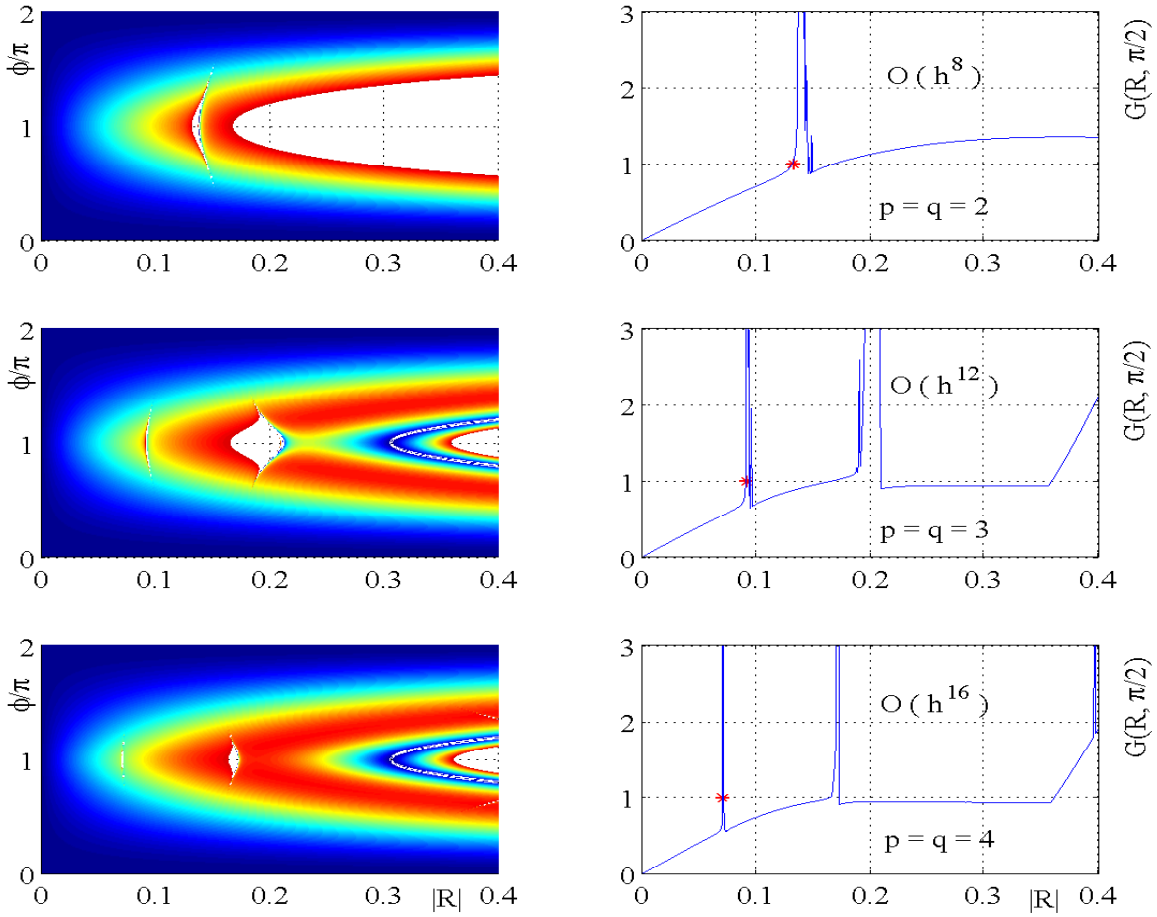


Рис. 1: Формы поверхностей $G(|R|, \psi)$ при $p = q$

В расчетах наблюдалось несколько видов поверхностей $G(|R|, \psi)$, их типичные представители изображены на рисунках 1–2 в осях $|R|$ и ψ/π. Все рисунки устроены одинаково — в левой части даны изображения поверхностей $G(|R|, \psi)$ в плане, а в правой части представлены эпюры поверхностей при $\psi = \pi$, представляющие собой графики $z = G(|R|, \pi)$ на оси симметрии поверхностей. Звездочкой на правых рисунках помечены пороги устойчивости, там же приведены параметры шаблона схемы и порядок точности. Для всех схем с трехточечным оператором B (т. е. с $q = 1$) и для явных схем ($p = 0$) характерен приблизительно параболический профиль линий уровня (см. первую и третью поверхности на рисунке 2) без вырождений. Схемы, для которых вырождение поверхности в расчетах не наблюдалось, помечены в таблице 1 символом $*$ (звездочкой). В других вариантах схем основа поверхности имеет аналогичный гладкий вид, но на ее фоне имеются вырождения с бесконечными локальными значениями, соответствующими обращению в нуль собственных чисел λ_A. Эти аномалии иногда сами определяют порог устойчивости, как на всех поверхностях рисунка 1, или могут находиться за порогом устойчивости, как на второй поверхности рисунка 2. В некоторых вариантах область неустойчивости трансформировалась в полосу, ограниченную одинаково ориентирован-

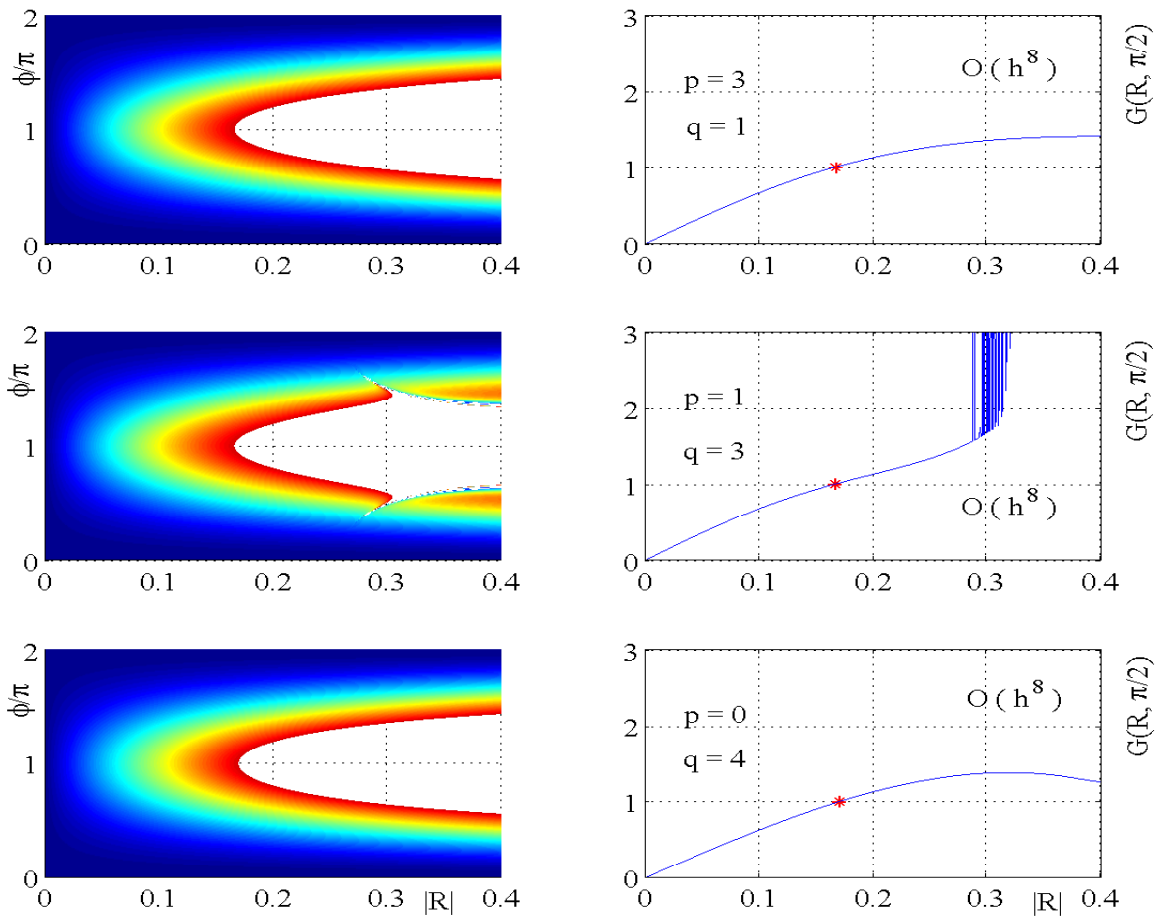


Рис. 2: Формы поверхностей $G(|R|, \psi)$ для трех схем восьмого порядка точности

ными кривыми параболического вида, или при некоторых параметрах шаблона приобретала ограниченную овальную форму или форму, ограниченную двумя причудливыми кривыми, как на рисунке 1.

Заключение

1. Все рассмотренные симметричные схемы в случае уравнения теплопроводности абсолютно неустойчивы. В случае уравнения Шредингера эти же схемы условно устойчивы, причем множители возрастания гармоник для них по модулю равны единице, т. е. схемы не являются диссипативными. Среди исследованных схем для уравнения Шредингера не нашлось ни одной схемы абсолютно устойчивой или абсолютно неустойчивой.

2. Следует заметить, что чем выше порядок точности по h, тем объективно менее жестким представляется ограничение на соотношение шагов. Так, например, сравним схемы $O(\tau^4 + h^8)$ и $O(\tau^4 + h^{12})$. В обоих случаях желательно соблюдать баланс погрешностей. В пером случае имеем естественное соотношение шагов $\tau = O(h^2)$, при котором $R = O(1)$, поэтому условие устойчивости вида $|R| \le M$ следует контролировать специально выбором шага по эволюционной переменной. Во втором случае требование баланса погрешностей диктует выбор шага $\tau = O(h^3)$, что соответствует $R = O(h)$, поэтому любое условие устойчивости вида $|R| \le M$ при при достаточно малом h выполнится автоматически.

3. По характеру условий устойчивости неявные схемы в данном случае не дают преимуществ перед явными, а по простоте уступают им. По этой причине, вероятно, неявным схемам следует предпочесть явные, соответствующие $p = 0$.

4. Особое место занимают схемы с параметром $p = 1$. Они реализуются привычной трехточечной прогонкой и в приграничных узлах выступают за контур области лишь на n-ом слое, где решение уже известно,

что упрощает реализацию. Схемы с $p > 1$ имеет многоточечный (более чем трехточечный) оператор A, требующий обращения, и поэтому требует доопределения еще неизвестного решения на верхнем слое, что усложняет краевые условия и реализацию в целом.

Список литературы

[1] Akhmediev N. N., Afanasiev V. V. Singularities and special soliton solutions of the cubic-quintic complex Ginsburg-Landau equation // Physical Review E. 1996. V. 53, iss.1, P. 1190–1201.

[2] Lu S., Lu Q., Twizell E. H. Fourier spectral approximation to long-time behaviour of the derivative three-dimensional Ginzburg-Landau equation // J. Comput. Appl. Math. 2007. V. 198, P. 167–186.

[3] Shu-Sen Xie, Guang-Xing Li, Sucheol Yi. Compact finite difference schemes with high accuracy for one-dimensional nonlinear Schrödinger equation b,2.// Comput. Methods Appl. Mech. Engrg. 2009. V. 198, P. 1052–1061.

[4] Паасонен В. И., Федорук М. П. Компактная диссипативная схема для нелинейного уравнения Шредингера// Вычислительные технологии. 2011. Т. 16, № 6, С. 68–73.

[5] Паасонен В. И., Федорук М. П. Компактная безитерационная схема для нелинейного уравнения Шредингера// Вычислительные технологии.2012. Т. 17, № 3, С. 83–90.

Паасонен Виктор Иванович — к.ф.-м.н., ст. науч.сотр. Института вычислительных технологий СО РАН;
e-mail: paas@ict.nsc.ru.
Дата поступления — 30 апреля 2019 г.

КОМПАКТНЫЕ СХЕМЫ НА РАСШИРЕННЫХ АСИММЕТРИЧНЫХ ШАБЛОНАХ ДЛЯ УРАВНЕНИЯ ШРЕДИНГЕРА

В. И. Паасонен[1,2], М. П. Федорук[2,1]

[1] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.6
DOI: 10.24411/9999-016A-2019-10060

В работе исследуются произвольные двухслойные и трехслойные компактные схемы с многоточечными операторами общего вида, аппроксимирующие уравнение Шредингера с максимально возможным порядком по основной переменной. Установлено, что множество схем со слабой асимметрией (с асимметрией в один узел) фактически являются симметричными, имея нулевой коэффициент в узле, нарушающем симметрию. Доказано, что в симметричном случае схем с операторами, определенными на четном числе узлов, происходит вырождение системы уравнений для вычисления коэффициентов схемы, построенной из соображений формального равенства числа узлов схемы и числа условий аппроксимации, в связи с этим максимально возможный порядок оказывается на две единицы ниже формально ожидаемого. Исследована устойчивость схем путем численной проверки выполнения необходимого критерия Неймана. В результате расчетов установлено, что асимметричные двухслойные схемы либо симметризуются, либо являются абсолютно неустойчивыми, а в классе трехслойных схем существуют условно устойчивые симметричные или симметризуемые схемы, абсолютно неустойчивые существенно асимметричные или вырожденные.

Ключевые слова: компактная схема, условие устойчивости, уравнение Шредингера.

Введение

Для решения различных проблем нелинейной волновой оптики [1] имеет смысл использовать схемы на расширенных шаблонах, повышая тем самым порядок точности расчетов. Такого рода схемы на симметричных шаблонах, являющиеся обобщением компактных схем четвертого порядка точности [2, 3], изучены в работе [4]. В частности установлено, что различные схемы одного и того же порядка аппроксимации по основной переменной, определенные на одном и том же шаблоне, могут быть устойчивыми или неустойчивыми в зависимости от числа удержанных членов разложения погрешности по эволюционной переменной.

Но для схем на произвольных несимметричных шаблонах вопрос оставался открытым, хотя такие схемы предположительно могли быть востребованы для решения некоторых типов задач волоконной оптики. В связи с этим авторам представлялось важным исследовать схемы максимального порядка аппроксимации в случае произвольных расширенных шаблонов, не обладающих симметрией, и этой проблеме посвящена данная работа.

1 Схемы на произвольных шаблонах

Итак, решается периодическая или краевая задача для нелинейного уравнения Шредингера:

$$\pm i\,U_t + dU_{xx} + s|U|^2\,U = 0, \quad 0 < t < T, \quad x_1 < x < x_2,$$

Работа выполнена при частичной финансовой поддержке Российского научного фонда (код проекта 17-72-30006).

ISBN 978-5-901548-42-4

где i — мнимая единица, t — эволюционная переменная, x — время (или конфигурационная переменная), d и s — вещественные коэффициенты. При $t = 0$ задается начальное данное $U(x, 0) = U_0(x)$.

Умножив уравнение на $\mp i$ приведем его к виду, совпадающую по форме с уравнением теплопроводности

$$U_t = cU_{xx} - f. \tag{1}$$

Отличие состоит в том, что в уравнении теплопроводности коэффициент c вещественный положительный, а в уравнении Шредингера — мнимый, и правая часть имеет специальный вид.

Введем сетку с шагом τ по t и шагом h по x и рассмотрим схему в канонической форме

$$A\,\hat{u}_t = c\frac{B}{h^2}u^n - g^n, \quad A = \sum_{i=-I_0}^{I_1} a_i T^i, \quad B = \sum_{j=-J_0}^{J_1} b_j T^j, \quad Tu(x) = u(x+h), \tag{2}$$

где

$$\hat{u}_t = \frac{u^{n+1} - u^n}{\tau}, \quad \text{либо} \quad \hat{u}_t = \frac{u^{n+1} - u^{n-1}}{2\,\tau},$$

T — оператор сдвига, а произвольные целые числа $I_k, J_k, (k = 0, 1)$ определяют «габариты» операторов. Коэффициенты схемы $a_i, (i = -I_0, -I_0 + 1, ..., I_1)$ и $b_j, (j = -J_0, -J_0 + 1, ..., J_1)$ будем определять из системы условий аппроксимации с максимально возможным порядком относительно h и с заданным порядком относительно τ.

Представим погрешность аппроксимации двухслойной схемы (2) на достаточно гладких решениях дифференциального уравнения (1) в виде разложения по степеням шага τ.

$$\Psi = A\left(u_t + \frac{\tau}{2!}u_{tt} + \frac{\tau^2}{3!}u_{ttt} + O(\tau^3)\right) - c\frac{B}{h^2}u + g^n =$$

$$= c\left\{A\left(u_{xx} + \frac{c\tau}{2}u_{xxxx} + \frac{c^2\tau^2}{6}u_{xxxxxx}\right) - \frac{B}{h^2}u)\right\} +$$

$$\left\{g^n - A\left(f + \frac{\tau}{2}(f_t + cf_{xx}) + \frac{\tau^2}{6}(f_{tt} + cf_{xxt} + c^2 f_{xxxx})\right)\right\} + O(\tau^3).$$

Используя затем разложения по степеням h разностных операторов A и B:

$$A = \sum_{m=0}^{\infty} \frac{h^m}{m!} A_m \frac{\partial^m}{\partial x^m}, \quad A_m = \sum_{i=-I_0}^{I_1} a_i i^m, \quad B = \sum_{m=0}^{\infty} \frac{h^m}{m!} B_m \frac{\partial^m}{\partial x^m}, \quad B_m = \sum_{j=-J_0}^{J_1} b_j j^m, \quad m = 0, 1, ..., \tag{3}$$

получим разложение погрешности также и по степеням h:

$$\Psi = c\left\{\sum_{m=0}^{\infty} A_m \frac{h^m}{m!}\frac{\partial^m}{\partial x^m}\left(u_{xx} + \frac{c\tau}{2}u_{xxxx} + \frac{c^2\tau^2}{6}u_{xxxxxx}\right) - \sum_{m=0}^{\infty} B_m \frac{h^{m-2}}{m!}\frac{\partial^m u}{\partial x^m}\right\} + \tag{4}$$

$$+ \left\{g^n - \sum_{m=0}^{\infty} A_m \frac{h^m}{m!}\frac{\partial^m}{\partial x^m}\left(f + \frac{\tau}{2}(f_t + cf_{xx}) + \frac{\tau^2}{6}(f_{tt} + cf_{xxt} + c^2 f_{xxxx})\right)\right\} + O(\tau^3). \tag{5}$$

Для построения схемы максимального порядка аппроксимации по h вначале следует определить коэффициенты схемы (2) так, чтобы по возможности большее число первых членов разложения в первой скобке (4) обратилось в нуль, а после нахождения коэффициентов схемы необходимо правую часть схемы g^n задать как функционал от $f(x, t)$ таким образом, чтобы вторая скобка (5) также обратилось в нуль с тем же порядком. В случае различных порядков малости этих двух слагаемых порядок аппроксимации схемы в целом определяется менее порядком менее точного приближения.

Очевидно, разностный оператор B/h^2 должен аппроксимировать оператор двойного дифференцирования по x, поэтому в (3) коэффициенты $B_0 = B_1 = 0$, а из соображений нормировки коэффициент при f в разложении правой части схемы g_n должен быть равен единице, поэтому

$$B_0 = \sum_{j=-J_0}^{J_1} b_j = 0, \quad B_1 = \sum_{j=-J_0}^{J_1} b_j j = 0, \quad A_0 = \sum_{i=-I_0}^{I_1} a_i = 1. \tag{6}$$

Приводя подобные (с производными одинакового порядка) в первой скобке выражения (4), получим систему условий

$$\frac{A_m}{m!}+\frac{R}{2}\frac{A_{m-2}}{(m-2)!}+\frac{R^2}{6}\frac{A_{m-4}}{(m-4)!}=\frac{B_{m+2}}{(m+2)!},\quad m=0,1,...,M,\quad (R=\frac{c\tau}{h^2}). \tag{7}$$

Условимся факториалы отрицательных чисел считать равными единице, а выражения A_k с отрицательными индексами равными нулю, тогда запись системы (7) оказывается корректной и содержательно верной для любых $m \geq 0$.

Таким образом, при выполнении условий (6)–(7) погрешность аппроксимации для однородного уравнения ($f \equiv 0$) составляет величину $O(\tau^3 + h^{M+1})$.

Из контекста вывода системы (7) следует очевидный путь повышения порядка аппроксимации относительно τ. Для аппроксимации с погрешностью $O(\tau^{K+1} + h^{M+1})$ вместо (7) следует формулировать более общие условия

$$\frac{A_m}{m!}+\frac{R}{2!}\frac{A_{m-2}}{(m-2)!}+\frac{R^2}{3!}\frac{A_{m-4}}{(m-4)!}+...+\frac{R^K}{(K+1)!}\frac{A_{m-2K}}{(m-2K)!}=\frac{B_{m+2}}{(m+2)!},\quad m=0,1,...,M. \tag{8}$$

Ясно, что в случае трехслойной схемы $\hat{u}_t$ разлагается по четным степеням τ, при этом условия аппроксимации с погрешностью $O(\tau^{2(K+1)} + h^{M+1})$, аналогичные (8), будут выглядеть так:

$$\frac{A_m}{m!}+\frac{R^2}{3!}\frac{A_{m-4}}{(m-4)!}+\frac{R^4}{5!}\frac{A_{m-8}}{(m-8)!}+...+\frac{R^{2K}}{(2\,K+1)!}\frac{A_{m-4K}}{(m-4K)!}=\frac{B_{m+2}}{(m+2)!},\quad m=0,1,...,M. \tag{9}$$

В однородном случае ($f = 0$) вторая скобка в выражении погрешности отсутствует, и тогда (6), (8) (а для трехслойной схемы (6), (9)) есть система условий аппроксимации однородного уравнения с максимально возможным порядком. Так как выражения (3) являются линейными функциями коэффициентов схемы a_i, b_j, то (6), (8) есть система M+4 линейных уравнений для определения I_1+I_0+J_1+J_0+2 искомых коэффициентов схемы. Следовательно, если строки системы (3) линейно независимы, то максимально возможное значение $M = I_1 + I_0 + J_1 + J_0 - 2$. Ниже будет описан целый класс схем, для которых данный формально ожидаемый порядок M не достигается по причине линейной зависимости строк. В симметричных случаях порядок может быть также и выше ожидаемого из-за отсутствия в разложении погрешности нечетных степеней h.

Будем предполагать, что число K, определяющее максимальную степень удержанных членов разложения погрешности по τ, зафиксировано, а показатель M степени удержанных членов разложения по x выбран максимально возможным, и соответствующая система условий нормировки и аппроксимации решена, т. е. коэффициенты схемы однозначно определены. Однако для аппроксимации неоднородного уравнения требуется приближенное равенство нулю также и второй фигурной скобки в выражении погрешности, приведенного выше лишь для частного случая схемы третьего порядка аппроксимации по τ. В общем случае для двухслойной схемы правая часть схемы g^n с погрешностью порядка $O(\tau^{k+1})$ должна иметь такое же разложение, как функционал

$$g^n=\sum_{m=0}^{M}A_m\frac{h^m}{m!}\frac{\partial^m F}{\partial x^m}+O(\tau^{K+1}+h^{M+1}),\quad F=F_0+\frac{\tau}{2!}F_1+\frac{\tau^2}{3!}F_2+\cdots+\frac{\tau^(K)}{(K+1)!}F_K. \tag{10}$$

где

$$F_k=\sum_{r=0}^{k}c^r\frac{\partial^{r+k}f}{\partial x^{2r}\partial t^{k-r}},\quad k=0,1,\ldots,K.$$

К сожалению, аппроксимировать правую часть с любой точностью по τ не всегда удается. Например, если f задана или вычислена как сеточная функция, то аппроксимация ее второй производной по t невозможна по двум слоям, поэтому для двухслойных схем максимально достижимый порядок по τ лишь второй. Для трехслойных схем разложение в (10) содержит лишь четные степени τ, поэтому здесь максимально возможный порядок будет четвертым.

Однако это обстоятельство отнюдь не требует решать систему условий аппроксимации (6), (8) непременно при малых K. Как показано в [4] в частном случае симметричных схем, неустойчивая при некотором K схема может на том же шаблоне стать абсолютно устойчивой и даже диссипативной при другом K, поэтому полезно комбинировать аппроксимацию $O(\tau^2)$ для правой части (так как здесь повышение порядка невозможно) с аппроксимацией однородного уравнения с другим порядком K, обеспечивающим устойчивость. Порядок точности по τ в целом будет в любом случае вторым, но такой подход расширяет выбор схем.

2 Численное исследования схем

Определение выражений коэффициентов схемы из системы условий аппроксимации при небольших габаритах шаблона не представляет труда, особенно в операционной среде, допускающей символьные вычисления, однако при больших габаритах шаблона громоздкие выражения коэффициентов практически бесполезны. С другой стороны, схемы на очень больших шаблонах не представляют интереса. Интересны схемы с умеренными параметрами шаблонов. Для исследования их свойств, в частности устойчивости, можно воспользоваться прямыми вычислениями коэффициентов схем и множителей возрастания гармоник. При фиксированных I_l, J_l $(l = 0, 1)$ коэффициенты двухслойной максимально возможного порядка аппроксимации зависят от единственного параметра — соотношения шагов $R = c\tau/h^2$, а собственные значения операторов A и B, соответствующие гармонике $\exp(i\omega x)$:

$$\lambda_A = \sum_{k=-I_0}^{I_1} a_k \exp(k\psi), \quad \lambda_B = \sum_{k=-J_0}^{J_1} b_k \exp(k\psi), \quad (\psi = \omega h),$$

зависят только от R и от частоты ψ. Коэффициент возрастания гармоники для двухслойных схем имеет вид $\rho = 1+C$, а для трехслойных — $\rho = C \pm \sqrt{C^2+1}$, где $C = R\,\lambda_B/\lambda_A$, и, таким образом, областью устойчивости будет отрезок $|R| \le M$, где выполняется неравенство

$$G(|R|, \psi) = |\rho| \le 1 \quad \forall\, \psi. \in (0, 2\,\pi).$$

Очевидно, при фиксированных параметрах шаблона схемы функция $z = G(|R|, \psi)$ зависит только от двух переменных. Она вычислялась на равномерной сетке в области $\{0 \le |R| \le R_{max}) \times (0 \le \psi \le 2\,\pi\}$. Промежуток $[0, R_{max}]$ изменения $|R|$ подбирался так, чтобы найти наиболее широкую область устойчивости. Там, где значения z превышали единицу, они искусственно ограничивались до величины $1 + \varepsilon$. При визуализации значения $z > 1$ игнорировались, поэтому при изображении поверхности в плане (виде сверху) белые (не прорисованные) зоны в точности соответствовали областям нарушения устойчивости. В результате расчетов определялся порог устойчивости M.

2.1 Вырождение и симметризация схем

При проведении расчетов выяснилось, что при четном числе узлов в шаблонах операторов A и B, расположенных симметрично относительно точки $x = X_s$ симметрии шаблона, система условий аппроксимации несовместна, что проявлялось в переполнении разрядов и крайне плохой оценке числа обусловленности. Например, такими являются варианты схем с $q = p+1 = 2$ или $q = p+1 = 4$. Ввиду геометрической симметрии таких схем и требования максимального порядка аппроксимации коэффициенты операторов также симметричны. На этой основе нетрудно доказать, что определитель системы условий аппроксимации действительно равен нулю.

В самом деле, пусть оператор A записан по четному числу узлов равномерной сетки и симметричен. Ввиду четности числа узлов координата $x = X_s$ точки симметрии оператора расположена посередине между узлами, и в точке симметрии оператор можно записать в виде

$$A\,w(X_s) = \sum_{l=1}^{L} a_l \{w(X_s + lh - h/2) + w(X_s - lh + h/2)\}.$$

Утверждается, что для такого A найдется симметричный оператор, записанный по нечетному множеству узлов

$$C\,w(x) = \sum_{l=-L+1}^{L-1} c_l w(x + lh), \quad c_{-l} = c_l, l = 1, ..., L-1$$

такой, что $A = C\,(T_{h/2} + T_{-h/2})/2$.

В самом деле, данное операторное равенство означает выполнение системы уравнений на коэффициенты искомого оператора C:

$$c_{L-1} = 2a_L, \quad c_{l-1} + c_l = a_l, \quad l = 1, ..., L-2.$$

Матрица данной системы двухдиагональная с единицами на главной и примыкающей к ней диагонали. Определитель системы, очевидно, равен единице. Следовательно, из любого симметричного оператора вида A (с

четным числом узлов) можно выделить ненулевой множитель, представляющий собой оператор осреднения как полусумму операторов сдвига на полшага вправо и влево.

Выделив этот множитель из операторов схемы A и B, а также из ее правой части (поскольку она тоже выражена симметричными выражениями), и сократив на него всю схему, получим эквивалентную симметричную схему с нечетным (на единицу меньшим) числом узлов на каждом слое шаблона. Так, приведенные выше примеры схем с $q = p+1 = 2$ или $q = p+1 = 4$ будут соответственно эквивалентны схемам с $q = p+1 = 1$ или $q = p + 1 = 3$.

На практике это означает, что в рассматриваемых симметричных случаях с четным числом узлов система уравнений для определения коэффициентов схемы, формально построенная из соображений совпадения числа неизвестных (числа узлов) и числа уравнений, оказывается несовместной. Удаление из системы последнего уравнения, соответствующего старшей степени разложения погрешности, приводит к совместной системе, однозначно определяющей коэффициенты схемы, которая имеет порядок на две единицы ниже формально ожидаемого.

В расчетах замечено также, что в тех случаях, когда шаблон схемы отличается лишь одним узлом от симметричного шаблона, происходит автоматическая симметризация схемы. Она проявляется в том, что коэффициент схемы в узле, нарушающем симметрию, в результате решения системы уравнений аппроксимации автоматически оказывается равным нулю. Это легко доказать в общем виде, исходя из того соображения, что в симметричных схемах разложение погрешности содержит только четные степени.

Установлено, что при наличии двух возможных путей симметризации почти симметричных схем реализуется тот вариант симметризации, который ведет к корректной схеме, а не к вырождающейся. Так например, схема с параметрами $p = q = 4$ имеет 5-точечный оператор A и 6-точечный B. Теоретически возможны два пути симметризации: Обращение в нуль либо коэффициента a_{-4}, либо b_1. Но первый случай ведет к вырождающейся схеме с четным числом узлов на обоих слоях(четыре и шесть), и он не реализуется при решении системы. Второй путь ведет к устойчивой схеме восьмого порядка точности с пятью узлами на обоих слоях, и именно он реализуется.

2.2 Результаты расчетов

Серия расчетов проведена для шаблонов специального вида (см. рисунок 1), для которого правые габариты $I_1 = 0$, $J_1 = 1$ зафиксированы, а левые $I_0 = p$, $J_0 = q$ изменялись. Нумерация узлов условна, в данном случае нулевой номер соответствует точке A. Предполагаемый порядок точности по h для при таких шаблонах составляет величину $S = p + q$, если только система для определения коэффициентов совместная. При симметрии шаблона порядок может понижаться на две единицы ввиду вырождения или повышаться на единицу, когда S нечетно, но условий для вырождения нет.

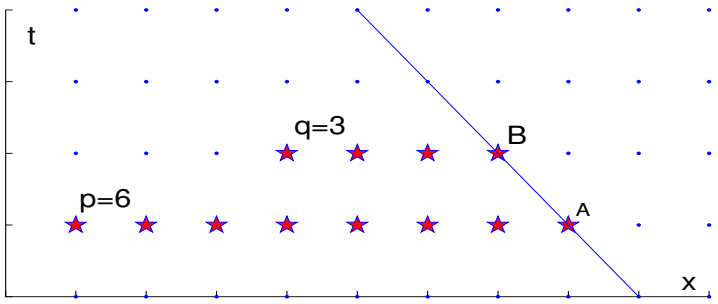


Рис. 1: Пример шаблонов операторов A и B.

Расчеты множителя возрастания гармоник проводились для параметров $0 \le p \le 5$, $1 \le q \le 5$. В этот класс попадает ряд схем от второго до десятого порядка аппроксимации относительно h. Расчеты показывают, что при существенной асимметрии (более, чем в один узел) все двухслойные схемы абсолютно неустойчивы. Численно подтверждены все случаи вырождения и симметризации. Условно устойчивых схем среди исследованных двухслойных не обнаружено. Устойчивость, когда она имеет место, в каждом случае абсолютна, и среди абсолютно устойчивых двухслойных схем найдена единственная диссипативная схема. А именно, множество устойчивых схем из числа исследованных ограничивается следующим перечнем: Это

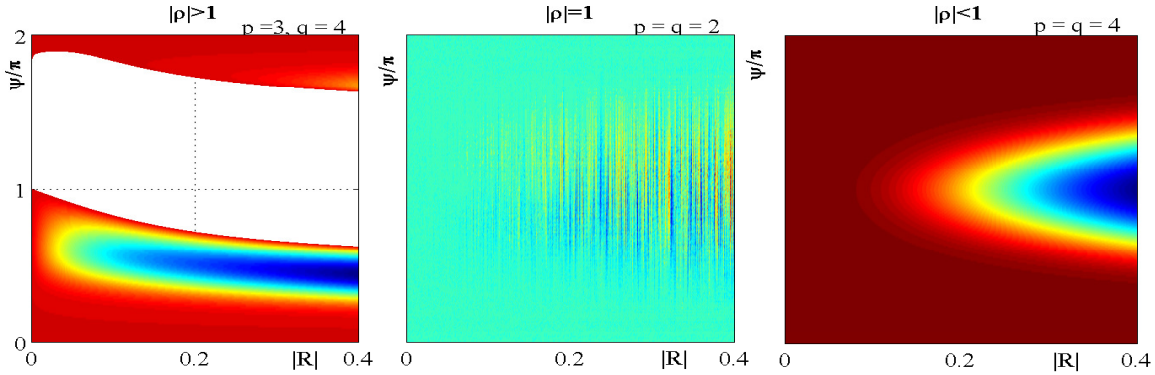


Рис. 2: Зависимость модуля возрастания ошибки схемы при различных параметрах шаблона.

абсолютно устойчивые не диссипативные ($|\rho| \equiv 1$) схемы с $p = q = 2$ при всех $K = 1, 2, 3, 4$ и схемы с $p = q = 4$ при $K = 3, 4$, а также единственная диссипативная ($|\rho| < 1$) схема с $p = q = 4$ при $K = 2$. Все остальные схемы абсолютно неустойчивые.

На рисунке 2 в частном случае $K = 2$ приведены три примера поверхности $z = |\rho|(|R|, \psi)$ для двухслойных схем (частотная координата на рисунках отнесена к π): для абсолютно неустойчивой схемы (слева), абсолютно устойчивой без диссипации (в центре) и абсолютно устойчивой диссипативной схемы (справа). Некоторая рябь на центральной картинке есть следствие вычислительной погрешности z-координаты поверхности, теоретически имеющей во всей области постоянное значение $z = 1$.

Пороги устойчивости трехслойных схем приведены в таблице 1. Здесь совершенно нет абсолютно устойчивых схем, но тоже имеет место и симметризация, и вырождение. Существенно несимметричные схемы (в таблице — отстоящие от диагонали более чем на единицу) абсолютно неустойчивые. Вырожденным схемам соответствует метка $Det = 0$, а звездочкой отмечены схемы, при поиске коэффициентов которых число обусловленности было менее 10^{-10}.

Таблица 1: Порог устойчивости максимально точных трехслойных схем $O(\tau^4 + h^{2(p+q)})$

$q \backslash p$	0	1	2	3	4	5
1	0.250	0.250	0	0	0	0
2	0.250	$Det = 0$	0.167	0	0	0
3	0	0.188	0.170	0.169	0*	0*
4	0	0	0.164	$Det = 0$*	0.013*	0*
5	0	0	0*	0.015*	0.008*	$Det = 0$*

Таким образом, из проведенного исследования можно сделать следующие выводы: Двухслойные и трехслойные существенно асимметричные схемы на расширенных шаблонах, записанные в первой канонической форме, бесперспективны ввиду их абсолютной неустойчивости. Схемы, близкие по построению к симметричным, фактически являются симметричными ввиду автоматической симметризации. Все симметричные схемы с четным числом узлов на каждом слое, а также почти симметричные, которые к таковым схемам автоматически симметризуются, являются вырожденными.

Следует заметить, что при возрастании суммы $p + q$ ограничение на параметр $R = c\tau/h^2$ фактически смягчается. Например, рассмотрим три схемы с $p = 1, 2, 3$ при фиксированном $q = 1$. Порог устойчивости для них (ограничения от $|R| \le 0.188$ до $|R| \le 0.169$ в четвертой строке таблицы) примерно одинаков. Однако естественное требование $\tau^2 = O(h^{p+q})$ сбалансированности погрешностей аппроксимации по τ и h приводит для этих схем к эквивалентностям $R = const$, $R = O(\sqrt{h})$ и $R = O(h)$ соответственно. Если в первом случае необходимо контролировать неравенство $|R| \le 0.188$ путем выбора шага по эволюционной переменной, то во втором и третьем случае ограничения по устойчивости при достаточно малых h выполнятся автоматически.

Список литературы

[1] Akhmediev N. N., Afanasiev V. V. Singularities and special soliton solutions of the cubic-quintic complex Ginsburg-Landau equation // Physical Review E. 1996. V. 53, iss.1, P. 1190–1201.

[2] Shu-Sen Xie, Guang-Xing Li, Sucheol Yi. Compact finite difference schemes with high accuracy for one-dimensional nonlinear Schrödinger equation b,2.// Comput. Methods Appl. Mech. Engrg. 2009. V. 198, P. 1052–1061.

[3] Паасонен В. И., Федорук М. П. Компактная безитерационная схема для нелинейного уравнения Шредингера// Вычислительные технологии.2012. Т. 17, № 3, С. 83–90.

[4] Паасонен В. И. Классификация разностных схем максимально возможной точности на расширенных симметричных шаблонах для уравнения Шредингера и уравнения теплопроводности // СибЖВМ. 2019. В печати.

Паасонен Виктор Иванович — к.ф.-м.н., ст. науч.сотр. Института вычислительных технологий СО РАН;
e-mail: paas@ict.nsc.ru;
Федорук Михаил Петрович — член-корр. РАН, ректор Новосибирского государственного университета;
e-mail: mife@ict.nsc.ru
Дата поступления — 30 апреля 2019 г.

ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ СХОДИМОСТИ ДВУХСТАДИЙНОЙ ДИСКРЕТНО-АНАЛИТИЧЕСКОЙ ЧИСЛЕННОЙ СХЕМЫ ДЛЯ УРАВНЕНИЙ ТИПА ПРОДУКЦИИ-ДЕСТРУКЦИИ

А. В. Пененко[1,2], А. Б. Салимова[2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.62
DOI: 10.24411/9999-016A-2019-10061

Приводятся результаты численного исследования сходимости одностадийной и двухстадийной дискретно-аналитических численных схем для системы обыкновенных дифференциальных уравнений типа продукции-деструкции. Исследование скорости сходимости проводится с помощью сравнения численного решения с аналитическим решением уравнения Смолуховского. В проведенном численном эксперименте двухстадийная схема показывает второй порядок сходимости по времени.

Ключевые слова: системы продукции-деструкции, дискретно-аналитические схемы, порядок сходимости, уравнение Смолуховского

Введение

Модели продукции-деструкции широко применяются для моделирования различных природных, техногенных и социальных процессов. В частности, они описывают процессы трансформации химических веществ, динамику аэрозольных популяций (уравнение Смолуховского), процессы в живых системах (например, в рамках теории генных сетей) и т.д.

В приложениях часто требуется восстанавливать неизвестные параметры моделей, решая обратные задачи. При использовании алгоритма из [1, 2], обратная задача, сформулированная в виде системы обыкновенных дифференциальных уравнений, переформулируется в виде квази-линейного операторного уравнения с оператором чувствительности. Оператор чувствительности строится из ансамбля решений сопряженных уравнений. Следовательно, для успешного применения алгоритмов требуется оптимизировать вычисление сопряженных уравнений. Если для обеспечения требуемой точности численного решения приходится использовать сетки с большим числом точек, то хранение решений сопряженных уравнений становится сложной технической задачей. В этой связи актуальным является вопрос об уменьшении количества шагов сетки при сохранении точности решения. Решение сопряженной задачи по вычислительным затратам эквивалентно решению прямой задачи. Для уменьшения количества требуемых шагов сетки мы рассмотрели двухстадийные дискретно-аналитические схемы, полученные аналогично двухстадийным схемам из [3]. Достоинством схем является сохранение неотрицательности решения при неотрицательных входных данных, что важно для сохранения физического смысла получаемого решения. Для численного исследования порядка сходимости схем в этой работе мы использовали аналитическое решение для уравнения Смолуховского [4, 5].

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований по проекту 19-07-01135 (в части проведения численного исследования).

ISBN 978-5-901548-42-4

1 Постановка задачи

Рассмотрим систему вида продукции-деструкции

$$\frac{d\varphi_l(t)}{dt} + P_l(t,\varphi(t))\varphi_l(t) = \Pi_l(t,\varphi(t)) + r_l(t),\ t \in [0,T],\ l = 1,...,N_c, \tag{1}$$

$$\varphi_l(0) = \varphi_l^0,\ l = 1,...,N_c, \tag{2}$$

где $\varphi(t) \in \mathbb{R}^{N_c}$ — вектор функция с элементами $\varphi_l(t)$, $r_l(t)$ — функция источника для l-го элемента в зависимости от времени, P_l, Π_l — l-ые компоненты операторов деструкции и продукции P, Π соответственно, φ_l^0 — начальные значения для l-го элемента. При этом если $\varphi_l \geq 0$, то $P_l(t,\varphi) \geq 0$ и $\Pi_l(t,\varphi) \geq 0$. По заданным P, Π, φ^0, r требуется найти φ.

2 Дискретно-аналитические численные схемы

Введем на временном интервале $[0,T]$ расчетную сетку $\omega_t = \left\{0 = t^1 < t^2 < ... < t^{N_t} = T\right\}$ с шагом по времени $\Delta t^{j+0.5} = t^{j+1} - t^j$. Следуя [2, 3], рассмотрим для системы (1) одностадийную явную дискретно-аналитическую численную схему, известную как QSSA (Quasi Steady State Approximation) [6]:

$$\phi_l^{j+1} = \phi_l^j L^j(P_l(t^j,\phi^j)) + G^j(P_l(t^j,\phi^j))(\Pi_l(t^j,\phi^j) + r_l^j),\quad j = 1,...,N_t - 1, \tag{3}$$

$$\phi_l^1 = \phi_l^0,\ l = 1,...,N_c. \tag{4}$$

Здесь $\phi \in \mathbb{R}^{N_c \times N_t}$ является дискретным эквивалентом φ, а операторы $L^j(P)$ и $G^j(P)$ введены для работы с жесткими системами. Использование этих операторов позволяет разделить элементы вектора ϕ на группы «быстрых», «средних» и «медленных» в зависимости от их «времени жизни» (величины скорости деструкции). Элементы из этих групп обрабатываются по-разному.

$$L^j(P) = \begin{cases} 0, & P\Delta t > \varepsilon_{\max} \\ e^{-P\Delta t}, & P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ 1 - P\Delta t, & 0 < P\Delta t < \varepsilon_{\min} < 1 \end{cases}, \tag{5}$$

$$G^j(P) = \begin{cases} 1/P, & P\Delta t > \varepsilon_{\max} \\ \frac{1-e^{-P\Delta t}}{P\Delta t}\Delta t, & P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ \Delta t, & 0 < P\Delta t < \varepsilon_{\min} < 1 \end{cases}, \tag{6}$$

и $\Delta t = \Delta t^{j+0.5}$. Данная схема при неотрицательных входных данных и вектор-функциях продукции-деструкции дает неотрицательное решение. Она имеет первый порядок аппроксимации по времени.

Далее рассмотрим двухстадийную дискретно-аналитическую схему, аналогично [3]. Результат работы первой стадии совпадает с работой вышеописанной одностадийной схемы (обозначим ее решение через $\tilde{\phi}^{j+1}$).

$$\phi_l^{j+1} = \begin{cases} \frac{1}{2\hat{P}_l}\left(\Pi_l(t^j,\phi^j) + r_l^j + \Pi_l(t^{j+1},\tilde{\phi}^{j+1}) + r_l^{j+1}\right), & \hat{P}\Delta t > \varepsilon_{\max} \\ \phi_l^j e^{-\hat{P}_l\Delta t} + \frac{1-e^{-\hat{P}_l\frac{\Delta t}{2}}}{\hat{P}_l\frac{\Delta t}{2}}\left(\left(\Pi_l(t^j,\phi^j) + r^j\right)e^{-\hat{P}_l\frac{\Delta t}{2}} + \left(\Pi_l(t^{j+1},\tilde{\phi}^{j+1}) + r_l^{j+1}\right)\right)\frac{\Delta t}{2}, \\ \hat{P}\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ \phi_l^j + \frac{\Delta t}{2}\left(\left(\Pi_l(t^j,\phi^j) + r_l^j - P_l(t^j,\phi^j)\phi_l^j\right) + \left(\Pi_l(t^{j+1},\tilde{\phi}^{j+1}) + r_l^{j+1} - P_l(t^{j+1},\tilde{\phi}^{j+1})\tilde{\phi}_l^{j+1}\right)\right), \\ \hat{P}\Delta t < \varepsilon_{\min} \end{cases} \tag{7}$$

$$\phi_l^1 = \phi_l^0,\quad l = 1,...,N_c, \tag{8}$$

где $\hat{P}_l = \frac{P_l(t^j,\phi^j)+P_l(t^j,\tilde{\phi}^{j+1})}{2}$. Используя операторную запись, перепишем численную схему в следующем виде:

$$\phi_l^{j+1} = \phi_l^j \hat{L}_1^j(\hat{P}_l) + \tilde{\phi}_l^{j+1}\hat{L}_2^j(\hat{P}_l) + \hat{G}_1^j(\hat{P}_l)(\Pi_l(t^j,\phi^j) + r_l^j) + \hat{G}_2^j(\hat{P}_l)(\Pi_l(t^{j+1},\tilde{\phi}^{j+1}) + r_l^{j+1}), \tag{9}$$

$$\phi_l^1 = \phi_l^0, \quad l = 1, ..., N_c, \tag{10}$$

$$\hat{L}_1^j(P) = \begin{cases} 0, & P\Delta t > \varepsilon_{\max} \\ e^{-P\Delta t}, & P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ 1 - P\frac{\Delta t}{2}, & 0 < P\Delta t < \varepsilon_{\min} < 1 \end{cases}, \quad \hat{L}_2^j(P) = \begin{cases} 0, & P\Delta t > \varepsilon_{\max} \\ 0, & P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ -P\frac{\Delta t}{2}, & 0 < P\Delta t < \varepsilon_{\min} < 1 \end{cases}, \tag{11}$$

$$\hat{G}_1^j(P) = \begin{cases} \frac{1}{2P}, & P\Delta t > \varepsilon_{\max} \\ \frac{1-e^{-P\frac{\Delta t}{2}}}{P\frac{\Delta t}{2}} e^{-P\frac{\Delta t}{2}}\frac{\Delta t}{2}, & P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ \frac{\Delta t}{2}, & 0 < P\Delta t < \varepsilon_{\min} < 1 \end{cases}, \tag{12}$$

$$\hat{G}_2^j(P) = \begin{cases} \frac{1}{2P}, & P\Delta t > \varepsilon_{\max} \\ \frac{1-e^{-P\frac{\Delta t}{2}}}{P\frac{\Delta t}{2}}\frac{\Delta t}{2}, & P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ \frac{\Delta t}{2}, & 0 < P\Delta t < \varepsilon_{\min} < 1 \end{cases}. \tag{13}$$

3 Численное исследование сходимости

В качестве основного примера рассмотрим модель динамики аэрозольных популяций в атмосфере, описываемую уравнением Смолуховского. На временном интервале $[0, T]$ и на конечном интервале размеров частиц $[0, X]$ с сеткой $\omega_x = \{0 < x_1 < ... < x_{N_c} = X\}$ определим задачу Коши для дискретизированного по размерам частиц с использованием квадратур трапеций непрерывного уравнения Смолуховского:

$$\frac{d\varphi_l(t)}{dt} = \frac{1}{2}\sum_{i=1}^{l-1} k_{l-i,i}\varphi_i(t)\varphi_{l-i}(t)dx_i - \varphi_l(t)\sum_{i=1}^{N_c} k_{li}\varphi_i(t)dx_i + r_l(t), \tag{14}$$

$$\varphi_l(0) = \varphi_l^0, \ l = 1, ..., N_c, \tag{15}$$

где N_c — количество рассматриваемых размеров частиц $\{x_l\}_{l=1}^{N_c}$, $\varphi_l(t)$ — концентрация частиц, соответствующих l -му размеру x_l в момент времени t, $\varphi(t) \in \mathbb{R}^{N_c}$ — вектор функция с элементами $\varphi_l(t)$, $r_l(t)$ — функция источника частиц размера x_l в зависимости от времени, k — дискретное ядро коагуляции, и

$$dx_l = \begin{cases} (x_{l+1} - x_l)/2 & l = 1 \\ (x_{l+1} - x_{l-1})/2 & 1 < l < N_c \\ (x_l - x_{l-1})/2 & l = N_c \end{cases}.$$

Из физического смысла явления коагуляции следует симметричность и положительность ядра $k : k_{li} = k_{il} \geq 0$. При введении обозначений

$$P : \begin{cases} [0,T] \times \mathbb{R}^{N_c} \to \mathbb{R}^{N_c} \\ \{t, \varphi\} \mapsto \left\{\sum_{i=1}^{N_c} k_{li}\varphi_i dx_i\right\}_{l=1}^{N_c} \end{cases}, \quad \Pi : \begin{cases} [0,T] \times \mathbb{R}^{N_c} \to \mathbb{R}^{N_c} \\ \{t, \varphi\} \mapsto \left\{\frac{1}{2}\sum_{i=1}^{l-1} k_{l-i,i}\varphi_{l-i}\varphi_i dx_i\right\}_{l=1}^{N_c} \end{cases}, \tag{16}$$

система (14)–(15) приводится к виду (1)–(2).

Для интегро-дифференциального уравнения Смолуховского известны лишь несколько аналитических решений, в число которых входит аналитическое решение для постоянного ядра. Поэтому случай единичного ядра ($k_{li} = 1$) был выбран для проверки алгоритма в самых простых, но при этом уже содержательных, условиях. Матрица, описывающая ядро, является входным параметром соответствующей компьютерной программы. В будущем предполагается рассмотреть и другие ядра. При численном решении неограниченный интервал размеров частиц заменяется на достаточно большой ограниченный интервал.

Для проверки алгоритма решения прямой задачи используем аналитическое решение интегро-дифференциального уравнения Смолуховского [7], спроектировав его в пространство сеточных функций:

$$\varphi_l(t^j) = \frac{1}{\left(1 + \frac{t^j}{2}\right)^2} \exp\left(-\frac{x_l}{\left(1 + \frac{t^j}{2}\right)}\right). \tag{17}$$

Рассмотрим сходимость приближенного решения ϕ к аналитическому при $N_c = 1000$, $T = 100$, $X = 100$ и: 1) $N_t = 10000$, 2) $N_t = 2500$, 3) $N_t = 1000$, 4) $N_t = 500$. Относительные ошибки, вычисленная по Фробениусовой норме, представлены на рис. 1 для схемы первого порядка синей линией и для схемы второго порядка красной линией. В представленном эксперименте алгоритмы решения показывают соответственно первый и второй порядки сходимости.

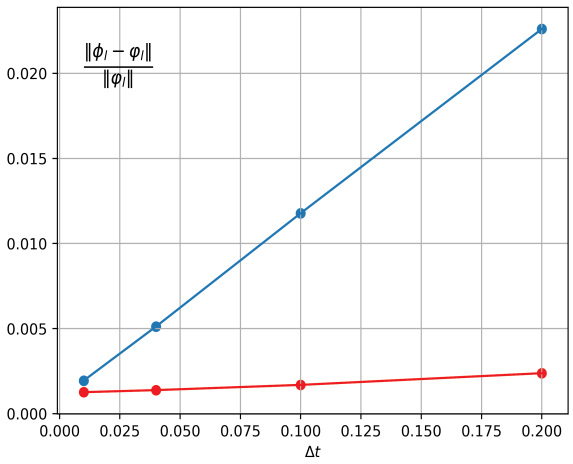


Рис. 1: Сравнение относительных ошибок для схем (3)–(4) (синим) и (9)–(10) (красным) при различных N_t ($\Delta t = T/N_t$).

Заключение

В работе приведен пример реализации дискретно-аналитических численных схем первого и второго порядка для решения прямой задачи. В проведенном численном эксперименте для уравнения Смолуховского схема второго порядка позволила получить при заданном числе точек по времени большую точность, чем одностадийная схема, что делает ее перспективной для применения в алгоритмах решения обратных задач на основе ансамблей решений сопряженных уравнений. В дальнейшей работе планируется построить на базе рассмотренной двухстадийной схемы алгоритм решения обратной задачи.

Список литературы

[1] Penenko A. V. A Newton-Kantorovich method in inverse source problems for production-destruction models with time series-type measurement data. Novosibirsk: Numerical Analysis and Applications, jan 2019. Vol. 12, № 1, P. 51–69.

[2] Penenko A. V. Consistent numerical schemes for solving nonlinear inverse source problems with gradient-type algorithms and Newton-Kantorovich methods. Novosibirsk: Numerical Analysis and Applications, jan 2018. Vol. 11, № 1, P. 73–88.

[3] Пененко В. В., Цветова Е. А. Вариационные методы построения монотонных аппроксимаций для моделей химии атмосферы. Новосибирск: Сиб. журн. вычисл. матем., 2013. Т. 16, № 3, С. 243–256.

[4] Smoluchowski M. V. Uber brownsche molekularbewegung unter einwirkung außerer krafte und deren zusammenhang mit der verallgemeinerten diffusionsgleichung. // Annalen der Physik, 1916. Vol. 353, no. 24, P. 1103–1112.

[5] Mueller Hans Zur allgemeinen theorie ser raschen koagulation. // Kolloidchemische Beihefte, 1928. Vol. 27, no. 6-12, P. 223–250.

[6] Hesstvedt Eigil, Hov Oystein, Isaksen Ivar S.A. Quasi-steady-state approximations in air pollution modeling: Comparison of two numerical schemes for oxidant prediction. // International Journal of Chemical Kinetics, 1978. Vol. 10, no. 9, P. 971–994.

[7] Галкин В. А. Уравнение Смолуховского. Москва: ФИЗМАТЛИТ, 2001.

Пененко Алексей Владимирович — к.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: a.penenko@yandex.ru;
Салимова Акжан Бауыржановна — студент, Новосибирский государственный университет;
e-mail: aksalimova@yandex.ru;
Дата поступления — 31 мая 2019 г.

ПОСТРОЕНИЕ МОДЕЛИ ИЗМЕНЕНИЯ ВО ВРЕМЕНИ СОДЕРЖАНИЯ ТЕМАТИЧЕСКИХ КЛАСТЕРОВ В КОЛЛЕКЦИЯХ НАУЧНЫХ ТЕКСТОВ

И. С. Пименов[1], Н. В. Саломатина[2]

[1] *Новосибирский государственный университет, 630090, Новосибирск*

[2] *Институт математики им. С.Л. Соболева, 630090, Новосибирск*

УДК 81'322.2
DOI: 10.24411/9999-016A-2019-10062

В работе представлены результаты решения задачи по оцениванию изменений, происходящих в отдельной предметной области с течением времени. Реализован подход, опирающийся на анализ полных текстов. Извлечение терминов, связей между ними и тематическая кластеризация проведены с помощью свободно распространяемой программы VosViewer, позволяющей выявлять термины в форме именных словосочетаний, выполнять их кластеризацию с использованием алгоритма, основанного на применении функции модулярности. Для отслеживания динамики развития и формирования тематических кластеров предлагается использовать ориентированные графы, построенные с учетом значимых изменений в кластере. Апробация подхода проведена на материале докладов представительных конференций по катализу "EuropaCat" за десятилетний период с 2005 г. по 2015 г. Количественно охарактеризованы тенденции в изменении тематических кластеров.

Ключевые слова: концептуальный анализ текстов, ориентированные графы, эволюция предметной области, катализ.

Введение

Исследование изменений, происходящих в отдельной предметной области (ПО) с течением времени, предполагает, что в определенные моменты нам известна ее структура. Распространенный способ представления структуры ПО — научные карты. На данный момент известны два основных подхода к построению карт. Первый опирается на представление структуры ПО исходя из совместного цитирования, например, в одной статье двух авторов или двумя авторами одной статьи [1,2], другой — из совместной встречаемости терминов ПО в текстах [3]. Как было показано в работе [4], результаты картирования, получаемые на одном и том же материале в соответствии с первым и со вторым методом построения карт, очень близки.

Для автоматического построения карт используют коллекции текстов, которые, в основном, формируются из статей, полученных по запросам к базам данных сетей цитирования и\или из рейтинговых публикаций в тематических журналах. Структура ПО представляется ассоциированными с тематической терминологией подобластей ПО кластерами, имеющими не только внутрикластерные, но и межкластерные связи.

При применении совместного цитирования группы тематически связанных текстов выявляются по авторским ссылкам. В случае использования совместной встречаемости терминов тематически связанными текстами считаются те, что содержат в значительном объеме ту же тематическую лексику. На основе различных мер ассоциаций строится сеть связанных ссылками документов или связанных терминов. С целью выявления структуры ПО проводится кластеризация терминов на сети (или до построения сети другими известными методами кластеризации). Если сеть построена на основе цитирования, то в этом случае узлами сети являются документы, из которых извлекаются термины, характеризующие построенные кластеры. В качестве терминов используют как ключевые слова (авторские или выбранные из словарей индексирования), так и слова, словосочетания, выявленные на основе анализа полных текстов, часто с применением частеречных шаблонов.

Работа выполнена при финансовой поддержке программы фундаментальных исследований Российской академии наук (код проекта 0314-2019-15).

ISBN 978-5-901548-42-4

Картирование каждой ПО представляет отдельный интерес, как в фиксированный момент времени, так и с точки зрения эволюции ПО, прогноза дальнейшего ее развития. Например, исследование структуры и динамики развития химии за тридцатилетний период представлены в [5], химии полимеров — в [3], теории нечетких множеств — в [4] и др. Рассматриваемая в данной работе ПО "катализ" отличается мультидисциплинарностью, включает разделы различных дисциплин: общую, неорганическую и органическую химию, физическую химию, общую химическую технологию и пр.

Исследование динамики изменений, происходящих в тематических кластерах с течением времени, может проводиться с помощью специальных алгоритмов классификации. Например, в работе [6] применяется метод скользящего окна, структура кластеров фиксируется в перекрывающихся временных интервалах, что позволяет максимально точно устанавливать зависимость состава кластеров от большого числа данных.

Особенность проводимого нами исследования заключается, в частности, в том, что оно реализовано на материале трудов конференций. Для них, в отличие от журнальных статей, как правило, рейтинг неизвестен. Но статус конференции и сам факт прохождения отбора говорит о значимости анализируемых текстов. Следует признать, что риск ошибок в оценивании тематических изменений ПО по материалам конференций выше, чем по рейтинговым статьям. Тем не менее, такое исследование представляет интерес, поскольку позволяет оперативно получать представление о смещении научных интересов в сторону той или иной темы на самом раннем этапе.

Цели работы: 1) разработать метод для отслеживания формирования и развития тематических кластеров в смежных временных периодах, пригодный для отбора важных изменений во всем анализируемом периоде; 2) провести апробацию на коллекции текстов.

В работе представлены предварительные результаты исследования динамики изменения тематических кластеров, проведенного на материале представительных конференций по катализу EuropaCat за период с 2005 г. по 2015 г.

1 Методы

Методы, которые применяются в этом исследовании для решения задачи картирования ПО, опираются на co-word анализ (анализ совместной встречаемости слов в текстах коллекции). Извлечение терминов, их кластеризация и визуализация в отдельные временные периоды проводятся с помощью свободно распространяемой программы VOSviewer [7].

Изменения тематических кластеров во времени прослеживаются как на уровне смежных временных отрезков с точки зрения формирования и развития кластеров, так и во всем анализируемом временном периоде.

1.1 Методы, реализованные в программе VOSviewer

Термин в co-word анализе, предоставляемом программой VOSviewer, определяется как последовательность существительных и прилагательных, оканчивающаяся на существительное. Сеть формируется путем связывания терминов i и j согласно мере ассоциации (s_{ij}):

$s_{ij} = 2ma_{ij}/k_ik_j$, где a_{ij} — вес связи между терминами i и j, равный числу текстов, в которых они встречаются совместно, а k_i, k_j — суммарный вес связей для i-го и j-го терминов соответственно, m — суммарный вес всех связей в сети [2].

Унифицированный (единый) подход к кластеризации и визуализации, реализуемый программой VOSviewer, опирается на нахождение минимума функции модулярности, в которой имеется параметр "резолюция", позволяющий строить сеть с необходимой степенью детализации. Термины, объединенные в кластеры, характеризуют отдельные темы ПО.

1.2 Стратегические диаграммы

Стратегические диаграммы (см. [3]) позволяют оценить разработанность темы и перспективность ее развития. Кластеры, построенные на коллекции текстов, принадлежащих одному временному периоду, можно разбить на четыре группы по показателям плотности (d) и центральности (c), которые вычисляются на базе индекса эквивалентности и характеризуют связность терминов внутри и вне кластера соответственно. Попадание кластера в каждый квадрант диаграммы можно интерпретировать следующим образом: 1-й

квадрант: важная тема, связанная с другими темами (значения $c > 0$ и $d > 0$); 2-й квадрант: хорошо развитая и узкоспециальная тема ($c \leq 0$ и $d > 0$); 3-й квадрант: слабо развитая тема (обычно это появляющиеся и исчезающие темы, значения $c \geq 0$ и $d \geq 0$) и 4-й квадрант: важные, но слаборазвитые темы, т.е. перспективные в смысле дальнейшего развития (значения $c > 0$ и $d \leq 0$). Начало координат устанавливается в пересечении медианных зачений $\{c_i\}$ и $\{d_i\}$ (i — число кластеров). Стратегические диаграммы часто применяют для "ручного" визуального анализа тематических кластеров, построенных для разных временных периодов.

1.3 Анализ изменения кластеров во времени

В данной работе для отслеживания временных изменений в кластере используются ориентированные графы, отражающие два аспекта функционирования кластера при переходе от одного временного периода к другому: формирование кластера в период времени t из кластеров, зафиксированных в период времени $t-1$, и развитие кластера в период $t+1$.

Пусть $C_t = \{c_i^t\}$ — множество кластеров в коллекции текстов, относящихся к периоду времени t, $i = 1, \ldots,$ K_t, K_t — число кластеров. Каждый кластер имеет метку $w_{c_i^t}$, ассоциируемую с тематическим содержанием. Метка — пара или тройка терминов, повторяющихся в разные временные периоды и имеющих наибольшее число внутрикластерных связей.

Пусть при переходе от временного среза t к $t+1$ каждый кластер c_i^t трансформируется в упорядоченное (по убыванию числа содержащихся терминов) множество $\{c_j^{t+1}\}$, $j = 1, \ldots, K_{t+1}$. При переходе от периода времени $t-1$ к t каждый кластер c_i^t формируется из упорядоченного (по убыванию числа содержащихся терминов) множества $\{c_j^{t-1}\}$, $j = 1, \ldots, K_{t-1}$. В анализе трансформации кластера c_i^t при переходе к периоду $t+1$ (или формировании кластера при переходе из $t-1$) учитываются четыре типа преобразований кластеров, удовлетворяющие следующим условиям (символами "cd" обозначены типы развития, "cf " — типы формирования, под Δ понимается разница объемов общих элементов кластеров):
$cf : |c_i^t \cap c_1^{t+1}|/|c_i^t| - |c_i^t \cap c_n^{t+1}|/|c_i^t| = \Delta_{1,n}$; $cd : |c_i^t \cap c_1^{t-1}|/|c_i^t| - |c_i^t \cap c_n^{t-1}|/|c_i^t| = \Delta_{1,n}$, $n = 2, \ldots$.
cd:
1 кластер c_i^t трансформировался большей частью в c_1^{t+1}, так что $\Delta_{1,2} > p$;
2 кластер c_i^t трансформировался большей частью в c_1^{t+1} и c_2^{t+1}, так что $\Delta_{1,2} \leq p$ и $\Delta_{1,3} > p$;
3 число кластеров, в которые трансформировался большей частью кластер c_i^t, превышает два ($j > 2$, $\Delta_{1,j} \leq p$);
4 кластер c_i^t отсутствует во множестве $\{c_j^{t+1}\}$ большей частью своих элементов ($\Delta_{0,1} > p$, $c_0^t = c_i^t \setminus \{c_j^{t+1}\}$).
cf:
1 кластер c_i^t сформировался большей частью из c_1^{t-1}, так что $\Delta_{1,2} > p$;
2 кластер c_i^t сформировался большей частью из c_1^{t-1} и c_2^{t-1}, так что $\Delta_{1,2} \leq p$ и $\Delta_{1,3} > p$;
3 число кластеров, из которых сформировался большей частью кластер c_i^t, превышает два ($j > 2, \Delta_{1,j} \leq p$;
4 кластер c_i^t сформировался большей частью из терминов, не принадлежащих множеству $\{c_j^{t-1}\}$ ($\Delta_{0,1} > p$, $c_0^t = c_i^t \setminus \{c_j^{t-1}\}$). Символом p обозначен порог, определяемый экспериментально.
Случаи $cd = 1$ и $cf = 1$ можно интерпретировать как сохранение темы, возможно, с обновлением, $cd = 2$ и $cf = 2$ — как выделение темы в самостоятельную и слияние двух соответственно. Случаи $cd = 3$ и $cd = 4$ — это два типа прекращения существования темы, а $cf = 3$ и $cf = 4$ — формирование новых тем из разных источников: из уже существовавших тем и из новых, ранее не зафиксированных. Выполнение условий обеспечивает отбор кластеров, играющих существенную роль в трансформации.

2 Апробация

Результаты кластеризации терминов для каждого временного периода были получены с помощью программы VOSviewer при следующих установленных параметрах:

1) На этапе выбора объектов визуализации использован co-word анализ коллекции текстов, представленных в требуемом формате. Термины отфильтрованы по Стоп-словарю, содержавшему около 500 элементов.

2) Вычисление веса связи между терминами производилось согласно совместной встречаемости в текстах, что обеспечивалось установкой для параметра counting method значения "binary". В рассмотрение включались термины, совместно встретившиеся не менее чем в 10 текстах.

3) Число кластеров регулировалось выбором значения параметра resolution, в нашем случае он был равен 1.10.

Таблица 1: Количественные характеристики коллекций текстов

	$t=1$ 2005 г.	$t=2$ 2007 г.	$t=3$ 2009 г.	$t=4$ 2011 г.	$t=5$ 2013 г.	$t=6$ 2015 г.	Весь период (T) 2005-2015гг.
m_t	951	1496	1259	1133	1571	1080	7490
M_t	528465	1041884	852446	676498	1463022	801662	5363977
n_t	819	1693	1398	1131	2530	1322	2908

Полученные результаты существенно зависят от установленных параметров, в том числе от частоты совместной встречаемости терминов. В данной работе рассмотрению подлежат довольно высокочастотные термины, которые характеризуют трансформацию тем на уровне широких понятий ПО.

2.1 Данные

Материалом исследования служат англоязычные тезисы конференций по катализу EuropaCat за период продолжительностью в 10 лет: с 2005 г. по 2015 г. Число конференций T (количество рассматриваемых временных периодов t) равно 6-ти. В таблице 1 указано число докладов (m_t), представленных на конференции, объем текстов докладов в словоупотреблениях (M_t), число различных терминов n_t.

К анализируемым характеристикам в библиометрии относится такой показатель как всплеск числа публикаций, появление новых авторов. Можно отметить, что в периоды $t=2$ и $t=5$ наблюдается существенный рост значений по этим характеристикам.

2.2 Характеризация выделенных терминов и построенных кластеров

Из шести коллекций текстов было выделено 2908 различных терминов. Они представляют собой именные словосочетания до 4 слов длиной: biodiesel, vegetable oil, oxygen storage capacity, x ray photoelectron spectroscopy и др.

Для каждого временного периода t было построено от 10 до 12 кластеров объемом до 509 терминов ($t=5$). Отдельные одноэлементные кластеры образованы общими для ПО терминами, такими как: catalyst ($t=1,2,3,4,5,6$), industrial application ($t=6$), catalytic performance ($t=1$), metal ($t=2$). Самые объемные кластеры неоднородны по составу, содержат много общих для ПО терминов, характеризуются слабыми внешними (c) и внутренними (d) связями.

Отсутствие термина среди кластеризующихся не означает, что он не встречался в текстах вовсе. Часто, он был отсечен по порогу частоты встречаемости как, например, BIODIESEL, имеющий высокую суммарную абсолютную ($f_a^T=681$) и текстовую ($f_t^T=229$) частоту во всем рассматриваемом периоде T, кроме первого: $f_1^t=2$; $f_2^t=25$; $f_3^t=62$; $f_4^t=46$; $f_5^t=39$; $f_6^t=55$. Т.е., значительный рост интереса к теме жидкого биотоплива возникает, начиная с момента $t=2$. В качестве меток ($w_{c_i^t}$) кластеров выбираются термины с максимальным значением d, обладающие свойством сохраняться во всем временном периоде T. Их перечень приведен в таблице 2.

В таблице указаны метки для кластеров, насчитывающих не менее десяти элементов. Аббревиатурами в метках обозначены используемые в ПО методы (ESR — метод магнитного резонанса), названия веществ (ZSM — цеолитный катализатор) и др.

2.3 Кластеры на стратегических диаграммах

Стратегические диаграммы построены для каждого временного периода. На рисунке 1(а) и 1(б) показано, как кластеры начального и конечного периодов расположились на стратегических диаграммах. В качестве единицы измерения c и d принимаются не значения, а ранги в упорядочении значений по убыванию. По оси абсцисс (c) откладываются ранги центральности кластеров, по оси ординат (d) — ранги плотности. Толщина и непрерывность линии окружности зависит от числа текстов, термины из которых представлены в кластере. Чем тоньше и прерывистей линия, тем из меньшего числа тектов извлечены термины.

Смещение интересов прослеживается и на рассматриваемых нами высокочастотных (довольно общих) терминах, что проявляется, в том числе, и на уровне меток кластеров, обладающих большой устойчивостью. Кластер “Photocatalysis”, не слишком изменившись в доле текстов от их общего числа в периоде, переместился из 2-го квадранта в 1-й, т.е. связь этой темы с другими темами возросла. Кластеры “Acidity, Zeolite” и

Таблица 2: Перечень меток кластеров ($w_{c_i^t}$) для временного периода T

i	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	t = 6
1	Kinetic Model, Oxygen	Solvent, Alcohol, Biodiesel	Alcohol, Biodiesel, Oil	Solvent, Alcohol, Biodiesel	Solvent, Biomass	Solvent Alcohol
2	Acidity, Zeolite	DFT, Oxygen	Gold, Cerium, Oxygen	DFT, Kinetic Model	DFT DFT	XPS, Gold
3	Methane, Hydrogen	Photocatalysis	Zeolite, Acidity, Pore, XRD	XRD, LDH	Methane, Hydrogen	Methane, Hydrogen, Syngas
4	Cerium, TPR, XRD	Methane, Hydrogen, Syngas	Methane, Hydrogen, TPR	Photocatalysis, Gold	XRD, LDH	Zeolite, Pore, Acidity
5	H2O2, Solvent	XRD, Cerium, TPR	Photocatalysis	Methane, Hydrogen, Syngas	XPS, Gold	XRD, Cerium
6	Pore	Zeolite, Acidity	SCR, NSR	Zeolite, ZSM, Acidity	Photocatalysis	Oil, Viscosity
7	Aluminium, Platinum	SCR, NSR	ODH, Vanadium	Cobalt, FTS	Zeolite, Pore, Acidity	HDS, Aluminium
8	Photoca- talysis	Gold	HDS, Aluminium	SCR, NSR		Photocatalysis
9	SCR, Sulphur, Diesel	HDS		Cerium, XPS		
10	ESR, ZSM	Vanadium, XPS		HDS, Biomass		
11				Viscosity		

“Pore” трансформировались с обновлением в один и переместились в 4-й квадрант, у нового кластера потенциал возможного развития повысился. Кластер с меткой “H2O2, Solvent”, трансформировавшись в “Solvent, Biomass, Alcohol” и прибавив в доле текстов, демонстрирует слабую связность как с терминами внутри, так и с терминами из других тематических кластеров. Новый кластер “Oil, Viscosity” появился в период $t = 6$, а кластер “ZSM, ESR” прекратил свое существование. Поведение кластера с меткой “Methane, Hydrogen”очень стабильно: обновившись (“Methane, Hydrogen, Syngas”), он остался в том же квадранте (1-м).

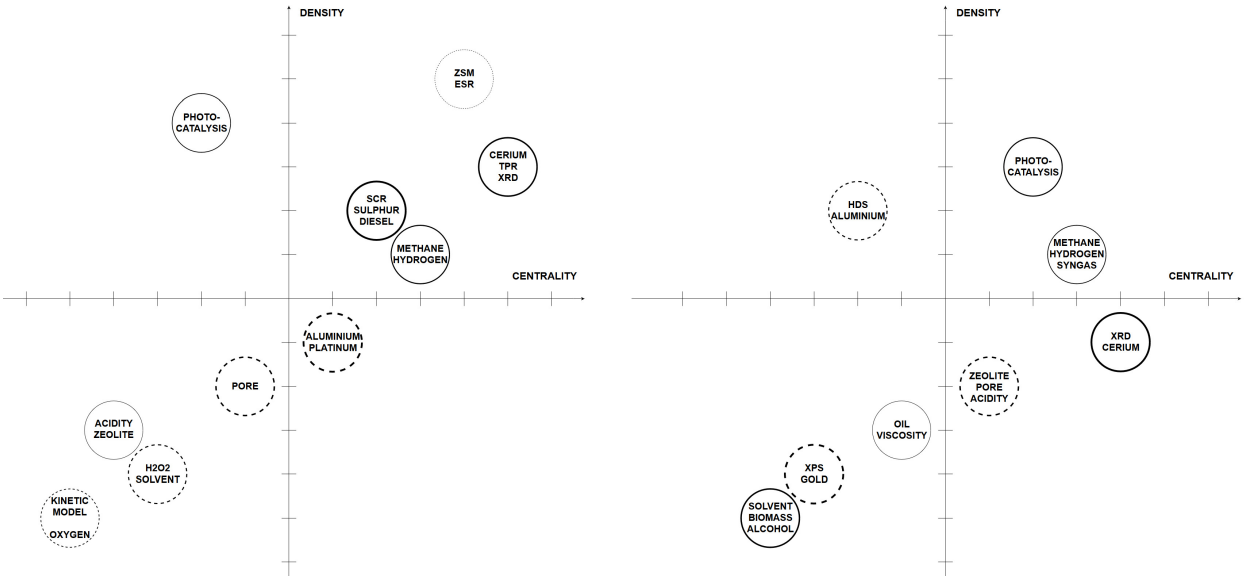
Более подробно временная трансформация тематических кластеров представлена в следующем разделе.

2.4 Графы формирования и развития кластеров

Эксперимент показал, что в процессе развития и формирования кластеров на уровне соседних временных периодов в течение всего анализируемого периода T чаще реализуются типы трансформации $cd = 1$ (в 61% случаев) и $cf = 1$ (43%). Например, при переходе от $t = 1$ к $t = 2$ развитие всех кластеров, кроме двух (7-го и 10-го) происходит по типу $cd = 1$.

Расщепление и слияние тем ($cd = 2$ и $cf = 2$) — происходит существенно реже — в 11% случаев. Исчезновение и возникновение новых тем происходит в 15% и 34% случаев соответственно. Периоды наибольшего обновления совпадают с временными отсчетами, в которые наблюдалось увеличение числа опубликованных докладов: $t = 2$ и $t = 5$. Кластеры 5-й и 6-й, $t = 5$ можно считать практически новыми, поскольку они содержат до 90% новых терминов.

На рисунках 2 и 3 показан граф формирования кластеров и граф их развития. Форма узлов и толщина стрелок зависят от того, какой тип развития или наследования реализуется. Прорисовывается не больше 3-х стрелок для $cf = 3$ и $cd = 3$. Один и тот же кластер на рисунке 2 и 3 может иметь разное количество вхо-



(а) t = 1 (б) t = 6

Рис. 1: Стратегические диаграммы двух периодов (а) и (б)

дящих стрелок, т.к. прорисовывание стрелок зависит от значения $\Delta_{1,n}$, которое вычисляется для каждого случаев cd и cf различным образом. Вершины, не имеющие входящих стрелок на рисунке 2, являются вновь сформированными в данном временном периоде. Вершины, не имеющие выходящих стрелок на рисунке 3, являются исчезающими. Отсутствие полного набора кластеров в фиксированный период времени на рисунках фиксирует отсутствие вновь образованных (или сильно обновленных новой терминологией) кластеров (рис. 2, $t = 6$) или переставших существовать (рис. 3, $t = 1$).

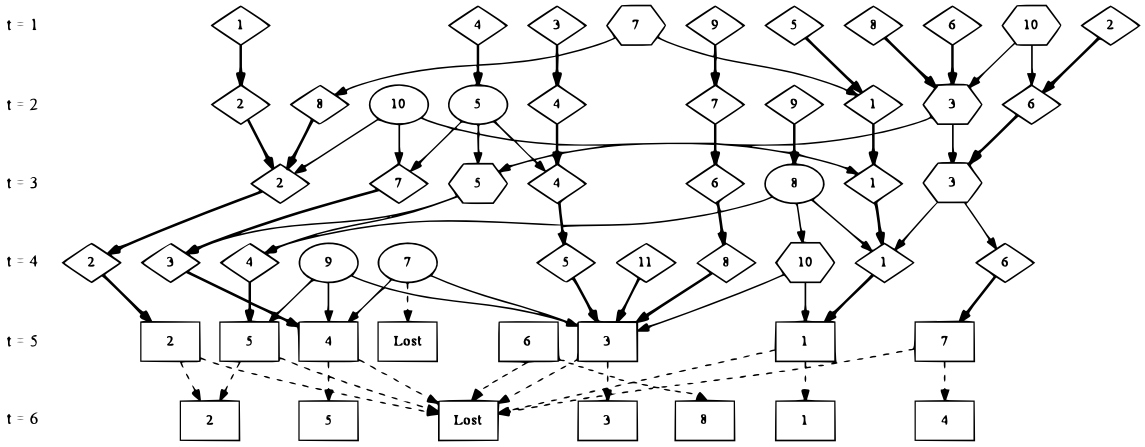


Рис. 2: Сеть развития кластеров

Если рассмотреть реализацию типов развития и формирования кластера в совокупности ($cd\&cf$) в отдельный временной период, то за весь рассматриваемый период (T) самым частым вариантом формирования и дальнейшего развития является сохранение темы ($cf = 1\&cd = 1$) в 25 % случаев. В 19% случаев реализуется схема $cf = 4\&cd = 1$ — образование новой темы с последующим сохранением. Следующий по числу случаев (14%) тип — $cf = 4\&cd = 4$, он характеризует случай кратковременного существования темы (характерна для $t = 2$ и $t = 5$). Оказались нереализованными 5 из 16 возможных типов, например, такие как $cf = 1\&cd = 4$ (кластеры, преимущественно сформированные из одного, не попали в исчезающие) и $cf = 3\&cd = 4$ (кластеры, сформированные из нескольких, не сохранились в следующем периоде).

Таким образом, кластеры, наследующие и сохраняющие преимущественно одну тему в отдельные перио-

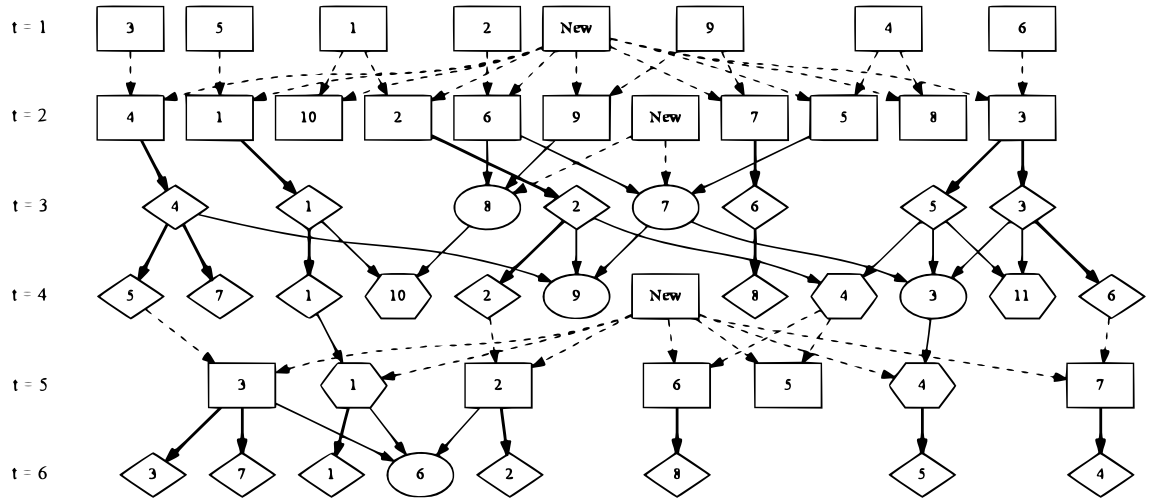


Рис. 3: Сеть формирования кластеров

ды времени, в среднем реже прекращают существование, тогда как кластеры, сформированные из нескольких проявляют неустойчивость: они могут в следующий период времени как сохраняться, так распадаться на несколько.

Заключение

На основе кластеров, построенных с помощью программы Vosviewer на 6-ти коллекциях текстов конференций по катализу, проведен анализ смещения интересов исследователей в рамках отдельной ПО. Выявлены закономерности в поведении кластеров на уровне их формирования и развития для каждого временного периода. Построены графы формирования и развития кластеров на протяжении всего рассматриваемого временного периода. Проанализированы стратегические диаграммы начального и завершающего периода времени.

Дальнейшим развитием исследования может стать рассмотрение поведения кластеров, построенных на более узких специфических терминах с невысокой текстовой частотой. Интерес представляет сравнение полученных результатов с результатами, которые может дать анализ рейтинговых публикаций в журналах ПО.

Список литературы

[1] Small, H. Tracking and predicting growth areas in science [Электронный ресурс]. http://www.scimaps.org/exhibit/docs/small.pdf (дата обращения: 10.04.2019).

[2] Van Eck, N.J., Waltman, L. Visualizing Bibliometric Networks [Электронный ресурс]. https://link.springer.com/chapter/10.1007

[3] Callon, M., Courtial, J.P., Laville F. Co-word analysis as a tool for describing the network of interaction between basic and technological research: the case of polymer chemistry // Scientometrics. 1991. N 22, P. 155–205.

[4] Cobo, M.J., Lypez-Herrera, A.G., Herrera-Viedma, E., Herrera, F. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field [Электронный ресурс]. https://www.sciencedirect.com/science/article/pii/ S1751157710000891 (дата обращения: 10.04.2019). 2019/03/23.

[5] Boyak, K.W., Borner, K., Klavans, R. Mapping the structure and evolution of chemistry research [Электронный ресурс]. https://www.researchgate.net/publication/298808287_Mapping_the_Structure_and _Evolution_of_Chemistry_Research (дата обращения: 12.04.2019).

[6] Kandilas, V., Uphum, S. P., Ungar, L. H. Analyzing knowledge communities using foreground and background clusters [Электронный ресурс]. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.146.3141&rep=rep1&type=pdf (дата обращения: 12.04.2019).

[7] VOSviewer Homepage [Электронный ресурс]. http://www.vosviewer.com/ (дата обращения: 23.03.2019).

[8] RAWGraphs Homepage [Электронный ресурс]. http://app.rawgraphs.io/ (дата обращения: 23.03.2019).

Пименов Иван Сергеевич — магистрант Новосибирского государственного университета;
e-mail: pimenov@yandex.ru;
Саломатина Наталья Васильевна — к.ф.-м.н., с.н.с. Института математики им. С. Л. Соболева;
e-mail: salomatina_nv@live.ru.
Дата поступления — 30 апреля 2019 г.

ЗАДАЧА ПРОДОЛЖЕНИЯ РЕШЕНИЯ ТРЕХМЕРНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ С ДАННЫМИ НА ЧАСТИ ГРАНИЦЫ

А. Ю. Приходько[1,2], М. А. Шишленин[1,2,3]

[1] *Новосибирский государственный университет, 630090, Новосибирск*

[2] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[3] *Институт математики им. С.Л. Соболева СО РАН, 630090, Новосибирск*

УДК 519.622
DOI: 10.24411/9999-016A-2019-10063

Исследуется задача продолжения решения трехмерного параболического уравнения с данными, заданными на временоподобной поверхности. Прямая задача решается конечно-разностным методом расщепления по направлениям. Задача продолжения сводится к обратной задаче, которая является некорректной. Задача продолжения решается численно методом наискорейшего спуска. Получена формула градиента функционала. Приведены результаты численных расчетов.

Ключевые слова: задача продолжения, уравнение теплопроводности, метод наискорейшего спуска

Введение

При исследовании процессов тепло и массообмена в экспериментах основными измеряемыми величинами являются температура и плотность теплового потока.

Методы измерения температуры на данный момент хорошо развиты. Имеются как локальные, так и полевые методы измерения температуры поверхности исследуемых объектов с высокой точностью, например, инфракрасная термография.

Измерение плотности теплового потока задача более сложная. Имеющиеся на данный момент методы обладают невысокой точностью, датчики теплового потока имеют большие размеры и не могут быть успешно использованы, особенно в мини и микросистемах. Прямых методов дистанционного измерения плотности теплового потока на сегодняшний момент нет. Во многих задачах плотность теплового потока вычисляют по температурным измерениям.

В данной работе исследуется изменения температуры нагреваемого элемента в эксперименте (Рис. 1). При исследовании теплообмена в жидких и двухфазных средах измерить температуру стенки нагреваемой фольги на поверхности теплообмена сложно, а подчас невозможно, поэтому измеряют температуру границы, доступной для наблюдения. Температура верхней поверхности (границы) измеряется инфракрасным излучением. На недоступной поверхности фольги происходит теплообмен.

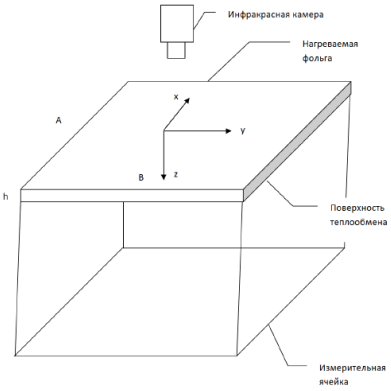
Задача состоит в определении потока тепла на недоступной части границы, а именно необходимо решить задачу продолжения решения с части границы.

В работах [12,17] исследована одномерная задача продолжения решения для параболического уравнения. Доказана теорема единственности решения, и численно сравниваются метод обращения разностной схемы, сингулярное разложение и градиентный метод.

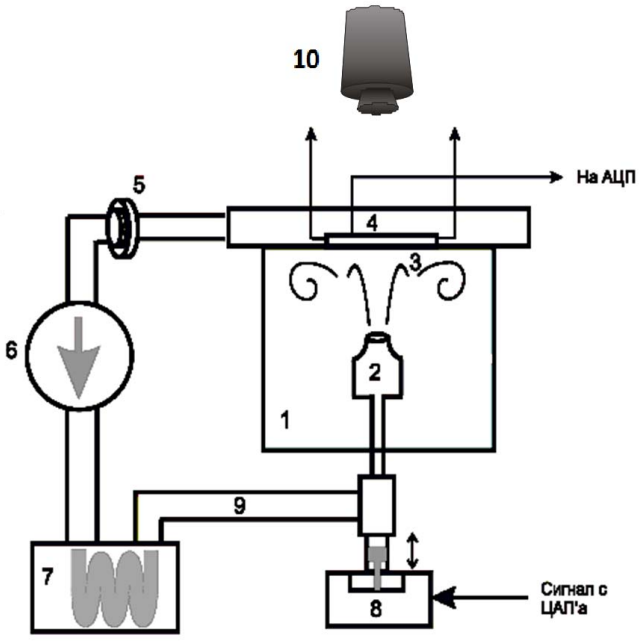
В работе [13] исследуется задача линейного теплообмена в однородном объекте, когда измерения температуры проводятся только в граничной области и требуется вычислить температуру в точках внутреннего контроля. Эта задача представляется в виде одномерной обратной задачи для параболического уравнения

Работа выполнена при поддержке РФФИ, проект 19-01-00694.

ISBN 978-5-901548-42-4



Нагреваемый элемент



1 — рабочий объем; 2 — струя (D=15 мм); 3 — блок; 4 — нагрев элемент; 5 — расходомер; 6 — насос; 7 — термостат; 8 — внешнее возмущение; 9 — система трубопроводов; 10 — инфракрасный сканер.

Рис. 1

в частных производных с ненулевыми условиями на одной граничной области стержня. В этой задаче требуется вычислить неизвестную граничную функцию на другом конце и найти значения температуры во внутренних точках. Авторы сводят обратную задачу к интегральному уравнению с помощью преобразований Лапласа. Полученное уравнение характеризует прямую зависимость неизвестной граничной функции от исходных данных.

В работе [15] предложен метод определения температуры при нелинейном теплопереносе внутри цилиндрического тела с неизвестной начальной температурой. Исходными данными являются результаты измерений температурных функций вблизи поверхности тела. Рассматриваемая задача возникает при термической обработке изделий в камерных печах. Математическая модель процесса включает в себя нелинейное уравнение теплопроводности, учитывающее зависимость теплофизических свойств материала от температуры, и граничные условия, характеризующие процесс теплообмена на поверхности тела. Для определения температуры внутри тела предложен метод дискретной регуляризации, позволяющий последовательно находить температуру в направлении от поверхности к оси цилиндра. Вычислительная схема метода предполагает использование стабилизирующих функционалов, а также конечно-разностных уравнений для определения температурных значений по пространственной переменной, что позволяет уменьшить влияние неизвестных начальных условий.

В работах [11, 14] предложен прямой метод решения задачи продолжения. Линейная обратная задача после дискретизации сводится к решению системы линейных алгебраических уравнений.

В данной работе построен численый метод решения трехмерной задачи продолжения с данными на части времениподобной границы.

1 Постановка задачи

Рассмотрим задачу продолжения решения с части границы $z = 0$ трехмерного уравнения теплопроводности в области $\Omega = \{(x, y, z) : x \in (0, A), y \in (0, B), z \in (0, h)\}$:

$$\frac{\partial u}{\partial t} = a(x, y, z)\Delta u + G(x, y, z, t) \qquad (x, y, z) \in \Omega, \qquad t \in (0, T), \tag{1}$$

$$u|_{t=0} = u_0(x, y, z), \tag{2}$$

$$u|_{z=0} = u_1(x, y, t), \tag{3}$$

$$u_z|_{z=0} = f(x, y, t), \tag{4}$$

$$u_x|_{x=0} = g_0(y, z, t), \qquad u_x|_{x=A} = g_1(y, z, t), \tag{5}$$

$$u_y|_{y=0} = h_0(x, z, t), \qquad u_y|_{y=B} = h_1(x, z, t). \tag{6}$$

Здесь $u(x, y, z, t)$ — температура, $a(x, y, z) = \frac{1}{\rho C_p}$, ρ — плотность, C_p — теплоемкость, u_0 — температура в начальный момент времени, $G(x, y, z, t)$ — функция источника.

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Задача продолжения заключается в определении функции $u(x, y, z, t)$ в области $\Omega \times (0, T)$.

Данная задача является некорректной. Приведем пример неустойчивости для одномерного случая [16]:

$$u_t^{(n)} = u_{xx}^{(n)}, \qquad x \in R, t > 0, \tag{7}$$

$$u|_{x=0} = \frac{2}{n}\sin(2n^2 t), \tag{8}$$

$$u_x^{(n)}|_{x=0} = 0. \tag{9}$$

нетрудно показать, что функция

$$u^{(n)}(x, t) = \frac{1}{n}[e^{nx}\sin(2n^2 t + nx) + e^{-nx}\sin(2n^2 t - nx)] \tag{10}$$

является решением задачи продолжения (7), (8), (9).

Когда данные (8) $\to 0$, при $n \to \infty$, решение $u^{(n)} \to \infty$ при $n \to \infty$.

Сведем задачу продолжения (1)–(6) к обратной задаче.

Рассмотрим прямую (корректную) задачу :

$$\frac{\partial u}{\partial t} = a(x, y, z)\Delta u + G(x, y, z, t) \qquad (x, y, z) \in \Omega, \qquad t \in (0, T), \tag{11}$$

$$u|_{t=0} = u_0(x, y, z), \tag{12}$$

$$u|_{z=0} = u_1(x, y, t), \tag{13}$$

$$u_z|_{z=h} = q(x, y, t), \tag{14}$$

$$u_x|_{x=0} = g_0(y, z, t), \qquad u_x|_{x=A} = g_1(y, z, t), \tag{15}$$

$$u_y|_{y=0} = h_0(x, z, t), \qquad u_y|_{y=B} = h_1(x, z, t). \tag{16}$$

Обратная задача состоит в опредлении q по дополнительной информации (4).

Задача продолжения (1)–(6) эквивалентна обратной задаче (11)–(16), (4) в следующем смысле. Если решена задача продолжения и найдена функция $u(x, y, z, t)$ во всей области $\Omega \times (0, T)$, то решение обратной задачи находится по формуле $q = u_z(x, y, h, t)$. И в обратную сторону, если решена обратная задача, т.е. найдена функция $q(x, y, t)$, то решение обратной задачи подставляем в уравнения (11)–(16), решением которых будет функция $u(x, y, z, t)$ в $\Omega \times (0, T)$, таким образом будет решена задача продолжения.

2 Численный метод решения

2.1 Решение прямой задачи

Для решения трехмерного неоднородного уравнения теплопроводности с переменным коэффициентом применяется метод расщепления Яненко [2].

$$\begin{cases} \dfrac{u^{n+1/3}-u^n}{\tau}+L_{1h}(\alpha u^{n+1/3}+\beta u^n)=f^n, \\ \dfrac{u^{n+2/3}-u^{n+1/3}}{\tau}+L_{2h}(\alpha u^{n+2/3}+\beta u^{n+1/3})=0, \\ \dfrac{u^{n+1}-u^{n+2/3}}{\tau}+L_{3h}(\alpha u^{n+1}+\beta u^{n+2/3})=0. \end{cases} \tag{17}$$

Здесь $L_{ih}=-\Lambda_{ii}$ — дискретный аналог второй производной.

$$\Lambda_{ii}u_j^n=\frac{u_{j-1}^n-2u_j^n+u_{j+1}^n}{h_i^2}+O(h_i^2), \tag{18}$$

$\alpha\in[0,1]$ — параметр схемы, $\beta=1-\alpha$.
Приведенная схема обладает свойствами:

- экономичность (реализация прогонкой в каждом отщепленном уравнении), при $\alpha>0$;
- абсолютная устойчивость (при $\alpha\geq 1/2$);
- аппроксимация с порядком $o(\tau+h_i^2)$.

2.2 Решение обратной задачи

Сформулируем обратную задачу (11) - (16), (4) в виде линейного операторного уравнения

$$Aq=f. \tag{19}$$

Здесь $q=q(x,y,t)$ — искомая функция, $f=f(x,y,t)$ — данные обратной задачи.

Обозначим $\Omega_{z=0}=\{(x,y,z,t):x\in(0,A),y\in(0,B),z=0\}$

Будем решать обратную задачу минимизацией следующего функционала

$$J(q)=\|Aq-f\|^2=\int\limits_0^T\int\limits_{\Omega_{z=0}}(u_z(x,y,0,t;q)-f(x,y,t))^2\mathrm{d}\Omega_z\mathrm{d}t\to\min_q, \tag{20}$$

Для решения задачи минимизации (20) применим метод наискорейшего спуска [1]. Если A — линейный ограниченный оператор, то справедлива оценка

$$J(q_n)\leq\frac{||A||^2||q_0-q_n||^2}{n}. \tag{21}$$

Рассмотрим метод наискорейшего спуска

$$q^{(n+1)}=q^{(n)}-\alpha_nJ'(q^{(n)}) \tag{22}$$

где $q^{(0)}$ — начальное приближение, $q^{(n)}$ — приближенной решение на n — итерации, $J'(q^{(n)})$ — градиент функционала, α_n — параметр спуска.

$$\alpha_n=\arg\min_{\alpha>0}J(q^{(n)}-\alpha J'(q_n))=\frac{\|J'(q^n\|^2}{2\|A(J'(q^n)\|^2}. \tag{23}$$

Градиент функционала вычисляется по формуле

$$J'(q)=(a\psi)|_{z=h}, \tag{24}$$

здесь ψ — решение сопряженной задачи

$$\begin{cases} -\dfrac{\partial \psi}{\partial t} = \Delta a\psi \qquad (x,y,z) \in \Omega, \qquad t \in (0,T), \\ (a\psi)|_{t=T} = 0, \\ (a\psi)|_{z=0} = 2(u_z(x,y,0,t;q) - f(x,y,t)), \\ (a\psi)_z|_{z=h} = 0, \\ (a\psi)_x|_{x=0} = (a\psi)_x|_{x=A} = 0, \\ (a\psi)_y|_{y=0} = (a\psi)_y|_{y=B} = 0. \end{cases} \tag{25}$$

Алгоритм :

1. Задаем начальное приближение $q^{(0)}$
2. Предположим что решение на n итерации найдено, покажем как найти решение на $n+1$ итерации.
3. Решаем прямую задачу (11)–(16).
4. Решаем сопряженную задачу (25).
5. Вычисляем градиент (24).
6. Вычисляем параметр спуска (23).
7. $q^{(n+1)} = q^{(n)} - \alpha_n J'(q^{(n)})$, Переход к шагу 2.

Критерий остановки $\|J(q^{(n)})\| \le \varepsilon$ или ограничение на количество итераций.

3 Численные результаты

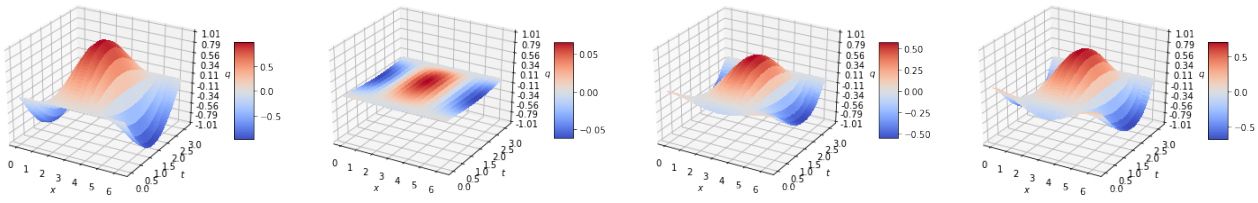
Рассмотрим результаты численных расчетов для симулированных данных: Зафиксируем $A = B = 2\pi$, $h = \pi/2$, $T = \pi/2$, $N_x = N_y = N_z = 40$, $N_t = 100$,

$$u(x,y,z,t) = \sin x \sin y \sin z \cos t. \tag{26}$$

И функции $G, u_0, u_1, g_0, g_1, h_0, h_1, q, f$ выбираются так, что бы удовлетворять уравнениям в прямой задаче.

Теперь $q = q^{(*)}$ — считаем неизвестной искомой функцией в обратной задаче, а f^δ — данными обратной задачи.

В качестве начального приближения зафиксируем $q_0(x,y,t) \equiv 0$, $y_{fix} = B/4$. N — количество итераций в градентном методе, $q^{(N)}$ — приближенное решение на N итерации.



$q^{(*)}(x, y_{fix}, t)$ $q^{(5)}(x, y_{fix}, t)$ $q^{(10)}(x, y_{fix}, t)$ $q^{(15)}(x, y_{fix}, t)$

Рис. 2: Сходимость градиентного метода q_n

На (Рис. 2) показаны результаты работы градиентного метода.

В таблице 1 используется гауссовский шум, $N = 50$ — количество итераций.

	$J(q)$	$\|q^{(*)} - q^{(N)}\|$
$\delta = 0\%$	0.0037	0.00058
$\delta = 2\%$	0.0590	0.00060
$\delta = 5\%$	0.3466	0.00627
$\delta = 10\%$	1.3961	0.00790

Таблица 1: Результаты для различного уравня шума δ

4 Заключение

Разработан численный метод решения задачи продолжения обладает высокой скоростью сходимости и приемлимой устойчивостью к шуму в данных. Для быстрого решения прямой задачи использован OpenMP и библиотека Intel MKL. Отметим, что можно существенно уменьшить число итераций, учитывая априорную информацию об искомом решении [6, 9].

Список литературы

[1] Кабанихин С. И. Обратные и некорректные задачи. – 2009.

[2] Яненко Н.Н. Метод дробных шагов решения многомерных задач математической физики. – Издательство “Наука”, Сибирское отделение, 1967.

[3] Алифанов О. М., Артюхин Е. А., Ненарокомов А. В. Обратные задачи в исследовании сложного теплообмена //М.: Янус-К. – 2009.

[4] Galaktionov V. A., Vazquez J. L. Continuation of blowup solutions of nonlinear heat equations in several space dimensions //Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences. – 1997. – Т. 50. – №. 1. – С. 1–67.

[5] Frankel J. I. Regularization of inverse heat conduction by combination of rate sensor analysis and analytic continuation //Journal of Engineering Mathematics. – 2007. – Т. 57. – №. 2. – С. 181–198.

[6] Kabanikhin S.I., Shishlenin M.A. Quasi-solution in inverse coefficient problems // Journal of Inverse and Ill-Posed Problems. 2008. V. 16, iss. 7. P. 705–713.

[7] Alifanov O. M. et al. Mathematical Model of Heat Transfer in High-Porous Materials //2010 14th International Heat Transfer Conference. – American Society of Mechanical Engineers, 2010. – С. 947–957.

[8] Алифанов О. М., Гребенников А. И. Быстрое решение обратных задач теплопроводности методом обобщенных лучей и его применение к моделированию наноструктурных материалов //Вестник Московского авиационного института. – 2010. – Т. 17. – №. 3. – С. 10–10.

[9] Кабанихин С.И., Шишленин М.А. Об использовании априорной информации в коэффициентных обратных задачах для гиперболических уравнений // Тр. ИММ УрО РАН. 2012. Т. 18, № 1. с. 147–164.

[10] Дучков А. А., Карчевский А. Л. Определение глубинного теплового потока по данным мониторинга температуры донных осадков //Сибирский журнал индустриальной математики. – 2013. – Т. 16. – №. 3. – С. 61–85.

[11] Shishlenin M.A. Matrix method in inverse source problems. Сибирские электронные математические известия. 2014. Т. 11. С. 161.

[12] Belonosov A., Shishlenin M. Regularization methods of the continuation problem for the parabolic equation. Lecture Notes in Computer Science. Vol. 10187 LNCS, 2017, Pages 220–226.

[13] Yaparova N., Gavrilova T. Mathematical modeling and method for solving an inverse heat conduction problem via the Volterra equation //AIP Conference Proceedings. – AIP Publishing, 2018. – Т. 2025. – №. 1. – С. 100013.

[14] Шишленин М.А. Микросейсмический мониторинг гидроразрыва пласта. Труды Международной конференции "Вычислительная математика и математическая геофизика" посвященная 90-летию со дня рождения академика А. С. Алексеева. 2018. С. 407-413.

[15] Япарова Н. М. Метод прогнозирования температурного состояния цилиндра при термообработке в условиях неполной исходной информации //Компьютерные технологии, управление, радиоэлектроника. – 2019. – Т. 19. – №. 2. – С. 54–65.

[16] Kabanikhin S. I. Inverse and ill-posed problems: theory and applications. – Walter De Gruyter, 2011. – Т. 55.

[17] Belonosov A., Shishlenin M., Klyuchinskiy D. A comparative analysis of numerical methods of solving the continuation problem for 1D parabolic equation with the data given on the part of the boundary //Advances in Computational Mathematics. – 2019. – Т. 45. – №. 2. – С. 735–755.

Приходько Алексей Юрьевич — магистрант ММФ,
Новосибирский государственный университет;
e-mail: a.prikhodko@g.nsu.ru;
Шишленин Максим Александрович — д.ф.-м.н.,
заведующий лабораторией Института вычислительной математики
и математической геофизики СО РАН;
Новосибирский государственный университет;
e-mail: mshishlenin@ngs.ru.
Дата поступления — 30 апреля 2017 г.

ЗАДАЧА ОЦЕНИВАНИЯ ПОЛЯ КОНЦЕНТРАЦИИ ПРИМЕСИ ОТ ПЛОЩАДНОГО ИСТОЧНИКА

В. Ф. Рапута[1,2], Т. В. Ярославцева[2], Р. А. Амикишиева[1,3]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Новосибирский институт гигиены, 630108, Новосибирск*
[3] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 551.511.61:541.182.3:543.2
DOI: 10.24411/9999-016A-2019-10064

Рассматриваются постановки задач оценивания по данным наблюдений атмосферного загрязнения территорий площадными источниками. Основу этих задач оценивания составляют различные асимптотики решений полуэмпирического уравнения турбулентной диффузии в приземном и пограничном слоях атмосферы. Обсуждаются модели восстановления полей загрязнения городских и около городских территорий, обусловленных совокупностью выбросов примеси от большого числа распределённых источников. При построении алгоритмов оценивания используются асимптотические методы теории потенциала и общие закономерности атмосферной диффузии примеси в атмосфере. Проведена апробация предложенных алгоритмов применительно к данным мониторинга загрязнения снегового покрова г. Новосибирска.

Ключевые слова: атмосфера, загрязнение города, площадной источник, мониторинг, асимптотика, параметр, модель оценивания.

Введение

Широкое применение расчётных методов для определения полей концентраций не исключают экспериментальные исследования состояния загрязнения городской среды. Необходимость их проведения связана со сложностями описания локальных климатических условий распространения примесей в атмосфере, с характером размещения промышленных предприятий и автомагистралей на территории города, типом застройки, рельефом местности, размером города и другими факторами [1–4]. Следует также отметить наличие неопределённости в количественном составе поступающих в атмосферу газовых и аэрозольных примесей, процессах их дальнейшей химической трансформации. В связи с этим результаты расчётов полей концентраций примесей необходимо сравнивать с экспериментальными данными, анализировать возможные расхождения, проводить поиск дополнительных источников и устанавливать соответствие расчётных значений максимумов концентраций реальным [5–9].

Для решения поставленных задач необходимы результаты мониторинговых исследований о содержании примесей в атмосферном воздухе. В связи с этим в городе требуется размещать весьма значительное количество станций контроля, что далеко не всегда возможно, исходя из экономических и технических ограничений. Одним из выходов в создавшейся ситуации является использование природных планшетов. В частности, проведения в зимнее время мониторинга загрязнения снежного покрова города [10–12]. Этот подход позволяет обеспечить непрерывность измерений и создать необходимую плотность сети наблюдений.

Работа выполнена при финансовой поддержке РФФИ и Правительства Новосибирской области в рамках научного проекта № 19-47-540008, в рамках Госзадания (№ 0315-2019-0004) и программы РАН № 51 (№ 0315-2018-0016).

ISBN 978-5-901548-42-4

1 Асимптотические приближения полей концентраций на территориях города

Поле концентрации лёгкой примеси от точечного источника за длительный промежуток времени может быть описано следующим соотношением [7, 8].

$$q(x,y) = \frac{\theta P(\gamma + 180^\circ)}{r^2 e^{\frac{2r_{max}}{r}}}, \tag{1}$$

где x, y — координаты расчетной точки с началом в месте расположения источника, θ — агрегированный параметр, зависящий от мощности эмиссии примеси и метеорологических величин, $P(\gamma)$ — повторяемость направлений ветра за рассматриваемый промежуток времени, $\gamma = \arctan\frac{y}{x}$, $r = \sqrt{x^2+y^2}$, r_{max} — расстояние от источника, на котором достигается максимальная приземная концентрация.

Тогда с учётом (1) концентрацию примеси на территории площадного источника S в точке $M(x,y) \in S$ можно представить в виде

$$Q(x,y) = \theta \cdot \iint_S \frac{m(\mu,\sigma)P(\varphi + 180^\circ)}{d^2 e^{\frac{2r_{\max}}{d}}} d\mu d\sigma, \tag{2}$$

где (μ,σ) — текущие координаты источника в точке M_1 $(\mu,\sigma) \in S$, $m(\mu,\sigma)$ — функция, описывающая эмиссию примеси в области S, $\varphi(\mu,\sigma,x,y) = \arctan\frac{y-\sigma}{x-\mu}$, $d = |M_1M| = \sqrt{(x-\mu)^2+(y-\sigma)^2}$.

В соотношении (2) содержится особая точка $(x = \mu, y = \sigma)$. Для того, чтобы локализовать эту особенность, представим функцию (2) в следующем виде

$$Q(x,y) = Q_1(x,y) + Q_2(x,y), \tag{3}$$

где

$$Q_1(x,y) = \theta \cdot \iint_{S-O(\varepsilon)} \frac{m(\mu,\sigma)P(\varphi + 180^\circ)}{d^2 e^{\frac{2r_{\max}}{d}}} d\mu d\sigma, \tag{4}$$

$$Q_2(x,y) = \theta \cdot \iint_{O(\varepsilon)} \frac{m(\mu,\sigma)P(\varphi + 180^\circ)}{d^2 e^{\frac{2r_{\max}}{d}}} d\mu d\sigma, \tag{5}$$

$O(\varepsilon)$ — круг радиуса ε с центром в точке (x,y).

Из разложения в степенной ряд знаменателя подынтегральной функции в (5) следует, что

$$d^2 e^{\frac{2r_{\max}}{d}} \geq d^2 + 2d \cdot r_{\max} + 2r_{\max}^2, \tag{6}$$

Тогда из (5), (6) вытекает неравенство

$$Q_2(x,y) \leq \theta \cdot \iint_{O(\varepsilon)} \frac{m(\mu,\sigma)P(\varphi + 180^\circ)}{d^2 + 2d \cdot r_{\max} + 2r_{\max}^2} d\mu d\sigma \equiv Q_3(x,y), \tag{7}$$

Из соотношений (3), (7) следует, что

$$Q(x,y) \leq F(x,y) \equiv Q_1(x,y) + Q_3(x,y), \tag{8}$$

Функция $F(x,y)$ уже не имеет в области S особых точек и является мажорантой полю концентрации примеси, задаваемому соотношением (2).

С другой стороны, согласно соотношению (3)

$$Q_1(x,y) \leq Q(x,y), \tag{9}$$

Функция $Q_1(x,y)$ является минорантой полю концентрации $Q(x,y)$ и стремится к нему при $\varepsilon \to 0$. В частности, при $\varepsilon = r_{\max}$ и малых $r_{\max}$ функция

$$G(x,y) = \theta \cdot \iint_{S-O(\varepsilon)} \frac{m(\mu,\sigma)P(\varphi + 180^\circ)}{d^2 e^2} d\mu d\sigma, \tag{10}$$

будет достаточно близка к функции $Q(x,y)$.

На соотношении (10) основана следующая интерполяционная формула [6]

$$Q_{\text{int}}(x,y) = \sum_{i=1}^{n} \frac{R_i}{(x-x_i)^2+(y-y_i)^2} \Big/ \sum_{i=1}^{n} \frac{1}{(x-x_i)^2+(y-y_i)^2}, \tag{11}$$

где R_i — измеренное значение концентрации примеси в точке $(x_i, y_i) \in S$, $(x,y) \neq (x_i, y_i)$, $i = 1, \ldots, n$.

Соотношение (11) получено, исходя из учёта общих закономерностей атмосферной диффузии лёгкой примеси от большого числа низких источников, рассредоточенных на определённой территории. В частности, к таким источникам следует отнести автомагистрали, частный сектор, пылящие поверхности на территории города и т.д. Следует также учитывать, что соотношение (11) пригодно для достаточно ограниченных территорий. Поэтому к числу важных характеристик относятся зоны репрезентативности точек наблюдений, которые зависят от пространственной изменчивости поля концентрации примеси [6].

В качестве примера использования формулы (11) на рис. 1 приведено проинтерполированное с редкой сети точек отбора проб снега поле концентрации бенз(а)пирена на территории г. Новосибирска. Точки отбора проб снега расположены вблизи стационарных пунктов Росгидромета контроля состояния атмосферного воздуха города, что создаёт возможности установления корреляционных связей межсредового загрязнения (атмосферный воздух — снежный покров) и выявления зон повышенных концентраций примесей.

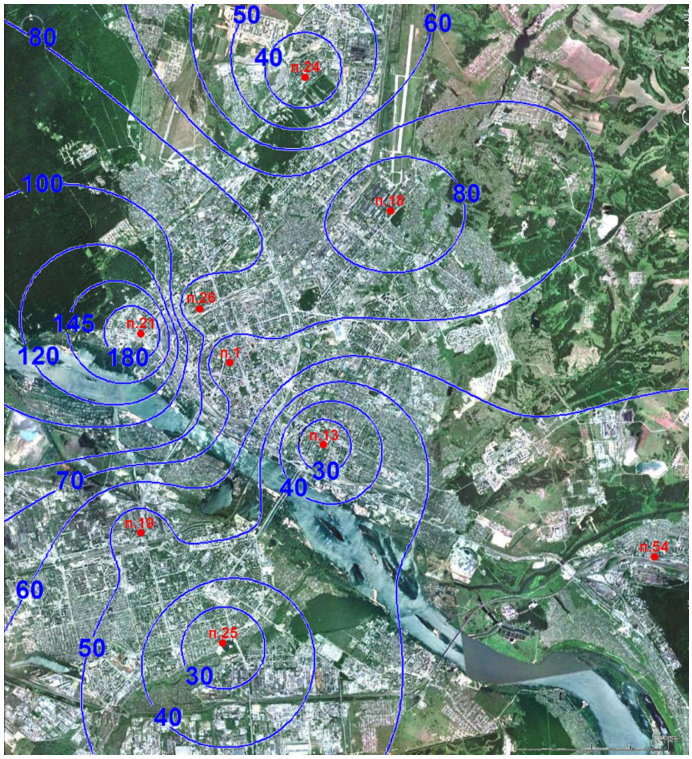


Рис. 1: Интерполированное поле значений концентрации бенз(а)пирена (нг/л) в снежном покрове г. Новосибирска по девяти точкам наблюдений

Анализ рис. 1 показывает наличие высоких уровней загрязнения снежного покрова в окрестностях постов №№ 21, 18, 1, 26. Сравнительно умеренные уровни концентраций зафиксированы на севере города (пост № 24) и в Кировском районе Новосибирска (пост № 25).

2 Асимптотическое поведение полей концентраций в окрестностях площадного источника

Прямое численное моделирование поля концентрации примеси за пределами площадного источника является не простой задачей. Одна из основных причин возникающих затруднений — отсутствие адекватного описания мощности эмиссии $m(\mu, \sigma)$ площадного источника S. В этом случае целесообразно использование как

данных мониторинга загрязнения, так и определённых асимптотик решений полуэмпирического уравнения турбулентной диффузии в приземном и пограничном слоях атмосферы [7].

Известно, что на удалениях от точечного источника примеси порядка 7–10 км, убывание концентрации по оси шлейфа вполне удовлетворительно описывается соотношением [7]

$$q(r,\theta) = \frac{\theta}{r}, \tag{12}$$

где ось r ориентирована в направлении ветра, $\theta = \frac{Q}{\sqrt{2\pi}Uhk_0}$, Q — мощность источника примеси, U, h — средняя скорость ветра и толщина слоя перемешивания, k_0 — коэффициент поперечного расширения шлейфа.

С учётом (12), по аналогии с (2) получим

$$Q(x,y) = \theta \cdot \iint_S \frac{m(\mu,\sigma)P(\varphi+180^\circ)}{d} d\mu d\sigma, \tag{13}$$

Используя свойства производящей функции для полиномов Лежандра [13], в первом приближении приходим к следующему описанию поля концентрации в окрестностях площадного источника

$$Q(r,\varphi,\theta) \approx \frac{\theta_1 P(\varphi+180^\circ)}{r}, \tag{14}$$

где

$$\theta_1 = M\theta, M = \iint_S m(\mu,\sigma)d\mu d\sigma. \tag{15}$$

Соотношения (14), (15) позволяют в агрегированном виде выразить неизвестные величины и провести их оценивание по данным мониторинга.

На рис. 2 представлена схема отбора проб снега в окрестностях Новосибирска в конце зимнего сезона 2009 г.

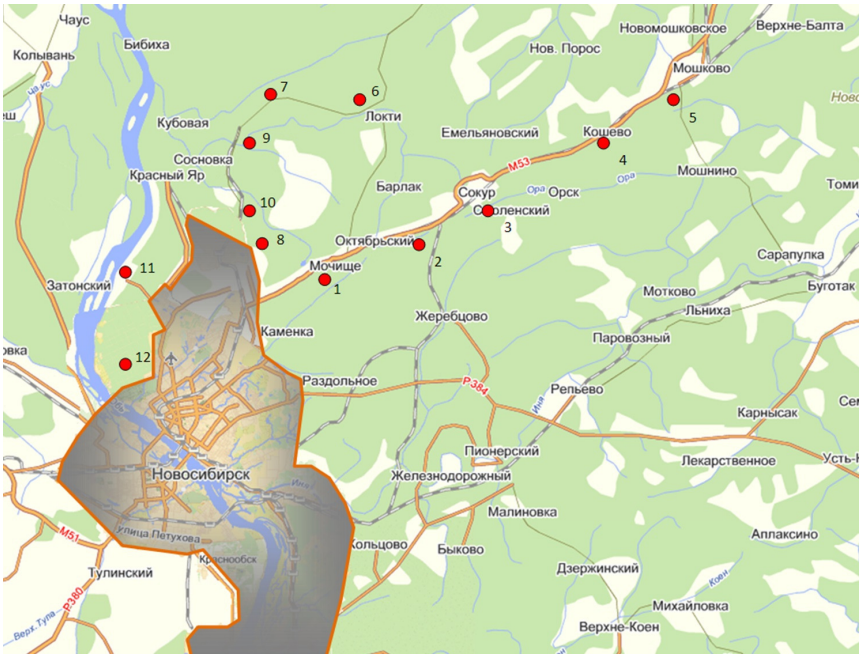


Рис. 2: Схема отбора проб снега в окрестностях г. Новосибирска, • — обозначены положения точек пробоотбора

Маршруты отбора проб выбирались с учётом направлений преобладающих выносов городских примесей и системы дорог. В данном случае этим требованиям в наибольшей степени удовлетворяет томско-кемеровская трасса, расположенная северовосточнее города. Зимняя повторяемость ветров южного и юго-западного направления составляет около 60 % [14].

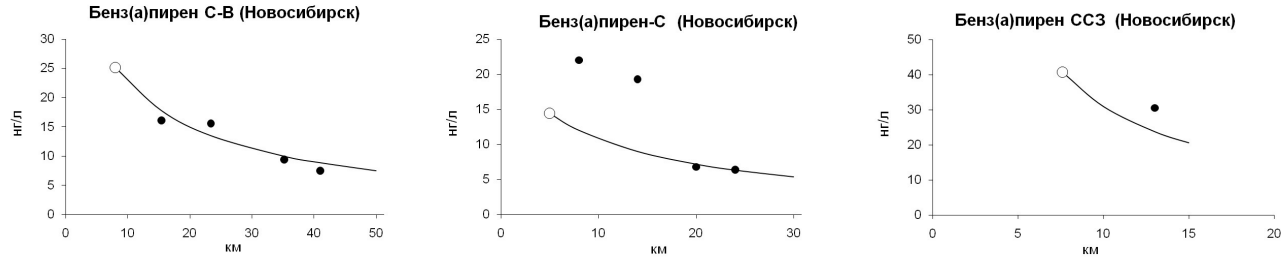


Рис. 3: Измеренные и численно реконструированные концентрации бенз(а)пирена в снеге в северо-восточном, северном и северо-северо-западном направлениях выноса от г. Новосибирска. ○ — опорные точки, • — контрольные точки наблюдений; ——— — восстановленная по модели (14) концентрация

Результаты численной интерпретации данных измерений бенз(а)пирена в зимнем сезоне 2008/09 гг. представлены на рис. 3,

Анализ рис. 3 показывает удовлетворительное согласие расчета наблюдениям в контрольных точках (чёрные кружки). Несмотря на значительный разброс в некоторых точках сравниваемых значений концентраций, в целом, преобладает вклад в загрязнение снегового покрова от совокупности источников, расположенных на территории Новосибирска.

Заключение

Предложены модели оценивания атмосферного загрязнения территорий площадными источниками. Их основу составляют различные асимптотики решений полуэмпирического уравнения турбулентной диффузии в приземном и пограничном слоях атмосферы. Выполнен численный анализ результатов мониторинговых исследований загрязнения снежного покрова в г. Новосибирске и его окрестностях. Проведённые исследования показали, что уровни загрязнения на преобладающих направлениях сноса вредных примесей весьма значительны даже на расстояниях нескольких десятков километров от городских территорий. Выявленные зависимости позволяют ограниченными средствами выполнить интегральные оценки выносов рассматриваемых примесей от Новосибирска за длительный промежуток времени и оценить степень антропогенной нагрузки на окружающие территории.

Список литературы

[1] Безуглая Э.Ю. Мониторинг состояния загрязнения атмосферы в городах. Л.: Гидрометеоиздат, 1986. 199 с.

[2] Берлянд М.Е., Безуглая Э.Ю., Генихович Е.Л., Зашихин М.Н., Оникул Р.И. О методах определения фонового загрязнения атмосферы в городах // Труды ГГО, 1984. Вып. 479. С.17–30.

[3] Безуглая Э.Ю., Смирнова И.В. Воздух городов и его изменения. СПб.: Астерион, 2008. 254 с.

[4] Аналитический обзор: Качество воздуха в крупнейших городах России за десять лет. СПб.: Астерион, 2009. 133 с.

[5] Ревокатова А.П., Суркова Г.В., Кислов А.В. О методе оперативного расчета эмиссии загрязняющих веществ в атмосферу на примере Москвы и Московской области // Вестн. Моск. ун-та. Сер. 5. География. 2013. № 1. С. 33–41.

[6] Грачёва И.Г., Оникул Р.И., Яковлева Е.А. Об интерполяции по территории города фоновых концентраций атмосферных примесей // Труды ГГО, 1998. — Вып. 549 — С. 98–107.

[7] Бызова Н. Л., Гаргер Е. К., Иванов В. Н. Экспериментальные исследования атмосферной диффузии и расчеты рассеяния примеси. Л. : Гидрометеоиздат, 1991. 278 с.

[8] Берлянд М.Е. Прогноз и регулирование загрязнения атмосферы. Л.: Гидрометеоиздат, 1985. 272 с.

[9] Konovalov I.B., Beekmann M., Kuznetsova I.N. Atmospheric impacts of the 2010 Russian wildfires: integrating modelling and measurements of an extreme air pollution episode in the Moscow region // Atmos. Chem. Phys. 2011. Vol. 11. P. 10 031-10 056.

[10] Лабузова О.М., Носкова Т.В., Лысенко М.С., Подчуфарова Д.П., Ильина Е.Г., Папина Т.С. Оценка влияния автотранспорта на загрязнение снежного покрова в городе Барнауле // Известия Алтайского отделения РГО. 2018. № 3 (50). С. 53–56.

[11] Демиденко Г.А., Напесочный Н.С. Оценка загрязнения снежного покрова в городе Красноярске // Вестник ОмГАУ. 2016. № 2 (22). С. 115–120.

[12] Василенко В.Н., Назаров И.М., Фридман Ш.Д. Мониторинг загрязнения снежного покрова. Л: Гидрометеоиздат, 1985. 182 с.

[13] Тихонов А.Н., Самарский А.А. Уравнения математической физики. М.: Наука, 1972. 735 с.

[14] Лучицкая И.О., Белая Н.И., Арбузов С.А. Климат Новосибирска и его изменения. Новосибирск: Изд-во СО РАН, 2014. 224 с.

Рапута Владимир Федотович — д.ф.-м.н., гл. науч.сотр. Института вычислительной математики и математической геофизики СО РАН; e-mail: raputa@sscc.ru;
Ярославцева Татьяна Владимировна — к.т.н., ст. науч. сотр. Новосибирского НИИ гигиены Роспотребнадзора; e-mail: tani-ta@list.ru;
Амикишиева Руслана Александровна — магистрант первого курса ФИТ Новосибирского государственного университета; e-mail: ruslana215w@mail.ru
Дата поступления — 30 апреля 2019 г.

ПОСТРОЕНИЕ РЕГУЛЯРИЗИРУЮЩЕГО ОПЕРАТОРА ПРИ ОБРАТНОМ АНАЛИЗЕ ОШИБОК

А. Н. Рогалев

Институт вычислительной моделирования СО РАН, 660036, Красноярск

УДК 517.968
DOI: 10.24411/9999-016A-2019-10065

В статье исследуется обратный анализ ошибок численных решений. Причина этого в том, что эффект сильного роста оценок границ ошибок проявляется для большого числа используемых методов оценки ошибки численного решения. Это означает отсутствие корректности алгоритмов оценки точности численных решений в силу не выполнения условий устойчивости относительно возмущений правой части. Поэтому разумно исследовать регуляризацию алгоритма оценки ошибки численного решения. Результатом регуляризации является алгоритм обратного анализа ошибок, его использование показало свою эффективность при оценке ошибок численных решений как систем линейных алгебраических уравнений (СЛАУ), также при оценке ошибок численных решений систем обыкновенных дифференциальных уравнений (ОДУ). Приводятся примеры численных расчетов.

Ключевые слова: регуляризация, обратный анализ ошибок, невязка.

Введение

Часто требуется оценить точность некоторого численного алгоритма по известным ошибкам вычисления решений поставленной задачи, то есть выполнить прямую задачу теории ошибок. Для большинства методов прямой задачи оценки ошибки численных (приближенных) решений наблюдается слишком большой рост границ оценок. В статье [1] отмечается, что "кто-то из классиков вычислительной математики (кажется Рихтмайер) около 1950 года заметил, что все разностные схемы некорректны в определенном смысле: при $h \to 0$ необходимо неограниченно увеличивать размерность сетки". Это пессимистичное утверждение, высказанное многими крупными специалистами по численному моделированию, подтверждает, что для большей части методов оценивания ошибки величина ошибки численного решения растет, хотя результаты расчетов часто оказываются более оптимистичными.

В силу этого началось развитие обратных задач теории ошибок: по заданной ошибке функции требуется определить ошибки ее аргументов. В результате возникает уравнение с $n-$ неизвестными, что допускает бесконечное множество решений. Для однозначного решения поставленной задачи требуется хотя бы одно дополнительное условие. Поэтому суть обратного анализа ошибок численного решения состоит в том, что при оценке точности численного решения, это решение рассматривается как точное решение задачи, близкой к исходной задаче [2,3,9]. Такой метод анализа ошибок был предложен Уилкинсоном [2] для численного решения задач линейной алгебры, и распространен на другие области численного анализа. В.В. Воеводин в рамках прямого и обратного анализа [3,4]получил мажорантные оценки влияния ошибок округления во всех наиболее важных методах линейной алгебры, что существенно усилило его результаты. Для прямых методов линейной алгебры ошибки округления результатов промежуточных вычислений рассматривались как функции случайных входных данных. Было доказано, что ошибки округления асимптотически (по числу разрядов представления чисел) являются независимыми, равномерно распределёнными случайными величинами. Были проведены исследования влияния малых возмущений входных данных на решение многих задач линейной алгебры, в том числе некорректно поставленных.

Общую идею обратного анализа ошибок (Backward Error Analysis) можно понять, рассматривая поставленную задачу как отображение F из пространства данных D в пространство решений S . Обратный анализ ошибок показал свою эффективность при оценке ошибок решений систем линейных алгебраических уравнений (СЛАУ) [5], [6] и для многих других задач.

ISBN 978-5-901548-42-4

1 Регуляризация оценок ошибки численных решений

Пусть рассматривается операторное уравнение

$$Az = u, z \in Z, u \in U, \tag{1}$$

используются обозначения $Z, U-$ некоторые метрические пространства. Задача(1) называется корректно поставленной, если выполняются следующие условия: 1) операторное уравнение (1) разрешимо для любой правой части $u \in U$; 2) решение операторного уравнения (1) устойчиво при возмущении правой части уравнения, это означает непрерывность обратного оператора A^{-1}, определенного на всем пространстве U . Если хотя бы одно из данных условий не выполнено, то задача называется некорректно поставленной [7]. Класс некорректных задач очень широк, к ним относятся задачи суммирования рядов Фурье с коэффициентами, заданными с ошибками, некоторые задачи оптимального управления, некоторые задачи линейной алгебры, минимизация функционалов. Опыт применения обратного анализа ошибок дает основание считать, что этот анализ также является некорректной задачей, обосновано применение регуляризации поставленной задачи.

Для решения некорректных задач вводится новая постановка некорректных задач, основанная на понятие регуляризации алгоритма [7]. Предположим, что решается операторное уравнение (1) с правой частью $\overline{u}$, имеющее единственное решение $\overline{z} \in Z$. На практике вместо $\overline{u}$ известно его приближенное значение $u_\delta \in U$, такое, что в метрике пространства U выполнено неравенство $\rho(u_\delta, \overline{u}) \leq \delta$. Величина δ погрешностью элемента u_δ.

Оператор R_δ, отображающий любую пару (u_δ, δ) в элемент $z_\delta \in Z$, удовлетворяющий условию сходимости $z_\delta \xrightarrow{Z} \overline{z}$ при $\delta \to 0$ называется регуляризирующим алгоритмом.Для регуляризируемой задачи можно построить семейство R_δ таких операторов, зависящих от параметра δ, таких, что $z_\delta = R_\delta u_\delta \xrightarrow{Z} \overline{z}$ при $\delta to 0$.

Возможен следующий подход построения регуляризирующего алгоритма для некорректной задачи. Предположим, что существует априорно заданное компактное множество корректности M, также полагаем, что отображение A на множество A_M взаимно однозначно. В этом случае обратное отображение A^{-1}, определенное на A_M, непрерывно, тогда за приближенное решение можно принять произвольный элемент $z_\delta \in M$, что $\rho(Az_\delta, u_\delta) \leq \delta$.

Известно [7], что элемент $z_\delta \in M$, минимизирующий на M невязку $\rho(Az, u_\delta)$ называется квазирешением уравнения (1):

$$\rho(Az_\delta, u_\delta) = \inf_{z \in M} \rho(Az, u_\delta). \tag{2}$$

В этой схеме используется дополнительная информация об искомом решении, иногда этого достаточно для выбора параметра регуляризации. в качестве основного алгоритма выбора параметра регуляризации часто используется обобщенный принцип невязки.

В этой работе мы проведем анализ регуляризации оператора оценки ошибок численных решений систем дифференциальных уравнений

$$\frac{dy}{dt} = f(t, y), \tag{3}$$

$$y(t_0) = y_0 \tag{4}$$

где $y \in R^n$, $t \in R$ и $f : R^n \times R \to R^n$..

Определение 1. Глобальной ошибкой интерполянта $u(t)$ численного решения системы ОДУ (4) называется разность $E(t) = y(t) - u(t)$ между точным решением $y(t)$ системы (4) и интерполянтом численного решения. В некоторых случаях термин глобальная ошибка используется также для обозначения нормы $\|y(t) - u(t)\|$ разности этих величин.

Определение 2. Дефектом, или невязкой называется величина

$$\delta(t) := \frac{du}{dt}(t) - f(t, u). \tag{5}$$

Обратный анализ ошибок численных решений можно выполнить применяя дефект для вычисления границы глобальной ошибки $\|y(t) - u(t)\|$, где $y(t)$ является точным решением системы ОДУ (4) и $u(t)$ является интерполянтом, построенным по численному решению y_h.

Так как это означает, что $\frac{du}{dt} = f(t, u) + \delta(t)$, то дефект равен величине, измеряющей величину того, насколько численное решение не удовлетворяет дифференциальному уравнению (4). Понятие дефекта (невязки) можно сравнить также с разностью между исходным уравнением (4) и уравнением, точным решением которого является численное решение.

2 Изменение параметра регуляризации

В некоторых случаях, включая представленный случай системы ОДУ, уместно рассмотреть пространство данных D как пространство задач (или уравнений); в таком случае под возмущением входа $\widehat{x}$ понимается видоизмененная задача (или уравнение), которая решается точно численным методом. В контексте задачи ОДУ отображению f соответствует задача "решить систему ОДУ, ' и диагональное отображение $\widehat{f}$ означает задачу "решить возмущенную систему ОДУ". Это происходит потому, что нас интересует возмущенная система $x + \Delta x$ как модифицированная модель, и размер возмущения Δx как возмущение указанной модели, в силу этого мы рассматриваем пространство данных в виде пространства задач. Одно из главных преимуществ возмущенной задачи обратного анализа ошибок состоит в том, что с ее помощью возможно оценить качество решения поставленной задачи, не зная, что ничего о точном решении, и если ошибка при обратном анализе мала, и соответствующие величины модели непрерывно изменяются при ее возмущении, то мы нашли достаточно точное решение. Для того, чтобы найти полезную информацию о численном решении, вычисляют число обусловленности задачи, то есть коэффициент чувствительность решения задачи к возмущению задачи.

Определение 3 (обусловленность). Пусть f задача, имеющая данные x и решение y. Тогда задача f называется хорошо обусловленной, если небольшие изменения данных x приводят лишь к небольшим изменениям решения y, и f плохо обусловлена, если небольшие изменения в данных x приводят к большим изменениям в решении y. Обусловленность задачи, как правило, характеризуется числом обусловленности, описывающим насколько возмущения данных усиливаются в решении. Можно определить число обусловленности χ задачи f

$$\sup \lim_{x \in D} \frac{\dfrac{\|f(\widehat{x}) - f(x)\|}{\|f(x)\|}}{\dfrac{\|\widehat{x} - x\|}{\|x\|}} \tag{6}$$

Таким образом, мы видим, что число обусловленности подобно производной измеряет (максимально в некотором диапазоне данных D скорость изменения решений относительно изменений данных).

Приведем в качестве примера линейную систему $Ax = B$, хорошо известно, что число обусловленности равняется произведению нормы матрицы A и норму обратной матрицы A^{-1}, что делает возможным его оценку или вычисление. При оценке числа обусловленности χ, известно, что если оценка ошибки обратным методом (δ_{back}) мала, то есть, невязка мала и число обусловленности невелико, то прямая оценка ошибки (Δ) также мала, так как

$$\|\Delta\| \leq \chi \, \|\delta_{back}\| .$$

Ситуация в случае систем ОДУ аналогична. Получение числа обусловленности для систем ОДУ является более сложным, возможно использовать для этого результаты вариационного исчисления.

Можно выделить три основные разновидности регуляризации анализа ошибки [9–11], которые изменяют либо вид дифференциального уравнения, либо дифференциального представления разностной схемы, либо условий, налагаемых на уравнение. Эти преобразования описаны в нескольких работах, наша задача связать их с регуляризацией оценки ошибки. Регуляризация оператора решаемой задачи, заключается в замене исходной системы на систему более простого вида, это помогает получить более точные оценки ошибок.

Первый вид анализа ошибок — анализ дефекта, который оперирует путем модификации уравнения и наложением фиксированных условий на уравнение. Это самый естественный вид обратного анализа для ОДУ. Численный метод применяется для решения уравнения, точное решение которого известно, это уравнение отличается от исходного введением дефекта, рассматриваемого как неавтономное возмущение ОДУ.

Дефект можно рассматривать как автономное возмущение ОДУ за счет увеличения размерности системы, используется стандартный прием преобразования неавтономной системы ОДУ в эквивалентную автономную систему, полагаем $y_{n+1} = t$ и вводим новое уравнение системы. Дефект наиболее часто используется в контексте задач управления дефектами, в которых величина нормы дефекта, для подходящей нормы, (чаще всего нормы максимум), контролируется численным методом. Численные методы, использующие управление дефектами, чаще всего содержат интерполянт, формирующийся на основе численного решения численного метода для оценки или вычисления дефекта, затем используется для управления размер шага численного метода.

Вторую разновидность обратного анализа ошибок для ОДУ составляют методы конструирования теневых решений ОДУ, эта методика выполняется путем изменения начальных условий задачи, оставляя систему

ОДУ без изменений. Основная задача слежки за теневым решением — показать, что численный метод вычисляет решение, расположенное вдоль точного решения исходной задачи с возмущенными начальными условиями в течение некоторого интервала времени. Анализ ошибок такого рода реализуется гораздо сложнее, чем в метода управления дефектами. Так как дифференциальное уравнение остается без изменений, метод больше подходит для задач, в которых уравнения движения очень хорошо известны и / или ошибки физических моделей и ошибки математических моделей ничтожно малы, поэтому мы заинтересованы в точных решений исходной задачи и так рассмотрение вопроса о внесении изменений в дифференциальные уравнения менее оправдана.

Третий вид анализа глобальной ошибки — это метод модифицированных уравнений, который реализуется как путем изменения исходной системы ОДУ и условий для этой системы. Этот подход использует как заданное уравнение и уравнения, определяющие численный метод, конструирующие формы уравнений и условия модифицированной задачи, решения, которые ближе к численным решениям, чем решения поставленной задачи. Таким образом, метод модифицированных уравнений может быть, лучше рассматривать как обобщение метода анализа дефектов и, следовательно, в этом методе очень много от формы обратного анализа ошибок BEA.

Различные методы могут дополнять друг друга. Например, в [10,11] применяются как метод модифицированных уравнений, так и метод контроля дефектов при анализе численного решения хаотических динамических систем. Частью обоснования процесса управления дефектом является то, что отношение дефекта к глобальной ошибке гораздо менее чувствительно к выбору используемого численного метода нежели локальная ошибка .Это позволяет пользователям проще интерпретировать результаты вычислений, а обеспечивать величину дефекта меньше заданного допуска. Еще одно преимущество управления дефектами, в частности, при использовании методов, которые строят границы дефекта, а не оценки его, состоит в том, что эта технология позволяет рассматривать решение, генерируемое численным методом, как просто еще один шаг в процессе упрощения процесса вычисления точного решения.

Самым полезным в численных методах решения ОДУ представляет глобальная ошибка $E(t)$, где

$$E(t_{n+1}) = y_{n+1} - y(t_{n+1}).$$

В общем случае подбор величины глобальной ошибки за счет выбора параметров численного метода не возможен. Вместо этого обычно выполняется контроль за величиной локальной ошибки $LEPS$ $\varepsilon(t)$, где

$$\varepsilon(t_{n+1}) = y_{n+1} - z_n(t_{n+1}),$$

напомним, что $z_n(t)-$ это локальное точное решение исходной системы ОДУ с начальным условием $y(t_n) = y_n$. Глобальная ошибка может быть представлена в виде

$$E(t_{n+1}) = [y_{n+1} - z_n(t_{n+1})] + [z_n(t_{n+1}) - y(t_{n+1})] \tag{7}$$

Первый член в квадратных скобках равен локальной ошибке, а второй член является величиной, которая зависит от устойчивости дифференциального уравнения, именно он зависит от того, насколько интегральные кривые, начинающиеся в точках y_n и $y(t_n)$ расходятся друг от друга в точке $t = t_{n+1}$. В частности, для малых по величине шагов второй член равен приблизительно

$$[I + h_n J_n] \cdot E(t_n),$$

где $h_n = t_{n+1} - t_n$ и $J_n = J_f(t_n, y_n)$.

Таким образом, выражение для глобальной ошибки распадается на сумму члена, зависящего от используемого численного метода, и члена, не зависящего от численного метода, а зависящего только от устойчивости самого дифференциального уравнения, так как собственные значения матрицы Якоби определяют (в линейном приближении) устойчивость самого уравнения.

Пытаясь управлять локальной ошибкой, мы хотя бы приблизительно управляем глобальной ошибкой, если решаемое дифференциальное уравнение не является динамически неустойчивым. Для неустойчивых уравнений локальная ошибка начинает расти при заданной величине шага, метод может обнаружить развивающуюся числовую неустойчивость и компенсировать за счет уменьшения размера шага, что сохраняет устойчивость метода.

В области анализа дефекта, существует строгий результат, соединяющий дефект и глобальную ошибку, а именно, следующая теорема:

Теорема Алексеева-Гребнера [12]. Пусть $y(t)$ это точное решение системы ОДУ (4) с начальным условием $y(t_0) = y_0$, и пусть $u(t)$ — это решение модифицированного уравнения

$$\frac{du}{dt} = f(t,u) + \delta(t,u), \quad u(t_0) = y_0,$$

в правую часть которого добавлен дефект (невязка) решения $\delta(t,u)$. В теореме Алексеева-Гребнера дефект может зависеть от решения u.Тогда, если $\phi_{t,t_0,f}$ представляет собой поток векторного поля f системы ОДУ и производная $\dfrac{\partial \phi_{t,t_0,f}}{\partial y_0}$ непрерывна, то

$$E(t) = y(t) - u(t) = \int_{t_0}^{t} \frac{\partial \phi_{t,\tau,f}(u(\tau))}{\partial y_0} \delta\left(u\left(\tau\right),\tau\right) d\tau. \tag{8}$$

3 Численные результаты

Для уравнения колебаний (записанного в виде системы двух дифференциальных уравнений первого порядка) на интервале $[0, 10000]$ были просчитаны значения численного решения одношаговым методом Рунге Кута 4 порядка и многошаговым методом Адамса 4 порядка.Учитывая, что точное решение системы известно и оно не вызывает необходимостей применять численные методы, направленные на компенсацию сложных областей решений (например, сингулярностей или жесткостей) для системы была получена оценка ошибки (глобальной ошибки). Использовалась отложенная разностная коррекция и экстраполяция по Ричардсону [13], а также обратный анализ ошибки.

Оценки, вычисленные с помощью обратного анализа ошибок выглядят так:

$$err(10) = 0.456148247 \cdot 10^{-6}, err(100) = 0.6465776709 \cdot 10^{-5}, \tag{9}$$
$$err(1000) = 0.78572562324 \cdot 10^{-4}, err(10000) = 0.187345329827 \cdot 10^{-3} \tag{10}$$

Оценки глобальной ошибки, вычисленные отложенной разностной коррекцией или экстраполяцией по Ричардсону (из пары чисел выбиралось минимальное значение)

$$err(10) = 0.56416401709 \cdot 10^{-4}, err(100) = 0.2383050218037 \cdot 10^{-2}, \tag{11}$$
$$err(1000) = 0.16208587246471 \cdot 10^{-1}, err(10000) = 0.41515538495543 \cdot 10^{-1} \tag{12}$$

Зная точное решение уравнения колебаний можно убедиться, что оба вида оценок являются верхними границами оценки ошибок, но оценки, вычисленные с помощью обратного анализа ошибок, значительно точнее.

Приведем пример количественных оценок влияния на точность решений ошибок округления в связи с шагом регулярной сетки *esize* и размерностью n матрицы жесткости, полученные для конечно-элементных моделей силовых конструкций при численном решении ряда модельных и прикладных задач [6].

В таблицах 1-2 показаны оценки ошибок узловых перемещений (мм), обусловленных влиянием ошибок округления в задачах, полученных тремя методами для рефлектора наземной зеркальной антенны спутниковой связи, блока механического устройства регулировки положения рефлектора зеркальной антенны космического базирования.

В этих таблицах латинскими цифрами I, II обозначены два варианта формирования систем, для которых применим обратный анализ ошибок. Пусть заданы две СЛАУ, имеющие одну матрицу коэффициентов A (матрицы жесткости), первая система

$$Au = b_{init} \tag{13}$$

имеет вектор b_{init} правых частей, для второй системы

$$Az = b_{new} \tag{14}$$

сформируем вектор правых частей b_{new} таким образом, чтобы полученная система имела известное решение.

В первом варианте I, решив систему линейных алгебраических уравнений численным методом M, мы находим решение u_{num}, затем, подставив это вектор решений в левую часть системы, вычисляем вектор $b_{new} = Au_{num}$. Таким образом, мы строим систему линейных алгебраических уравнений вида (14), точным

решением которой является вектор численного решения системы (13) u_{num}, в чем можно убедиться простой подстановкой. Далее решим полученную систему, используя тот же численный метод K на том же компьютере во время одного сеанса работы с программой, реализующей МКЭ.

В итоге мы вычислим вектор численного решения z_{num} системы вида (14) $Az_{num} = b_{new}$, где $b_{new} = Au_{num}$. Норма вектора разности $\|u_{num} - z_{num}\| \approx \|u_{num} - u\|$, то есть вектор разности $u_{num} - z_{num}$ приближает покомпонентно вектор ошибки численного метода K).

Далее второй вариант обратного анализа ошибки изменяет способ построения правой части СЛАУ. Сформулируем основы этого способа построения. Используя оценки нормы вектора решений СЛАУ, легко получить неравенство $\|u\| \geq \dfrac{\|b\|}{\|A\|}$. Будем конструировать решение $z = (z_k)_{k=1,n}$, компоненты которого $z_k = \dfrac{\|b\|\, r_k(A)}{\|A\|}$, либо $z_k = \dfrac{\|\phi(b)\|\, r_k(A)}{\|\psi(A)\|}$. Вид функций ϕ, ψ, r_k выбирается так, чтобы после подстановки этих функций в СЛАУ норма правой части была либо равна норме правой части, либо мажорировала ее. Например, если компоненты вектора решений положить равными

$$z_k = \left(\frac{\max\limits_i |b_i|}{\max\limits_i \sum_{j=1}^{n} a_{ij}} \right),$$

то после подстановки в систему (14) компоненты вектора правой части этой системы примут вид

$$b_{new,k} = \max_i |b_{init,i}| \left(\frac{\sum_{j=1}^{n} a_{ij}}{\max\limits_i \sum_{j=1}^{n} a_{ij}} \right). \tag{15}$$

Тогда для систем (13), (14) совпадут матрицы коэффициентов и будут равны нормы правых частей. Все тестовые примеры подтвердили оценки для ошибок численных решений этих систем

$$\|u - u_{num}\| \leq \|z - z_{num}\|. \tag{16}$$

Таблица 1: Ошибки при численном анализе рефлектора наземной антенны, мм

esize, мм	n	Ошибка по методу		
		Уилкинсона	I	II
200	108540	$2,235 \cdot 10^{6}$	$3,761 \cdot 10^{-11}$	$9,212 \cdot 10^{-10}$
150	169560	$5,436 \cdot 10^{6}$	$1,117 \cdot 10^{-10}$	$1,185 \cdot 10^{-9}$
100	339804	$1,993 \cdot 10^{7}$	$1,322 \cdot 10^{-10}$	$3,289 \cdot 10^{-9}$
80	513576	$6,453 \cdot 10^{7}$	$1,532 \cdot 10^{-10}$	$5,358 \cdot 10^{-9}$
60	869328	$2,708 \cdot 10^{8}$	$4,429 \cdot 10^{-10}$	$9,829 \cdot 10^{-9}$
40	1909548	$8,938 \cdot 10^{8}$	$1,238 \cdot 10^{-9}$	$2,321 \cdot 10^{-8}$

Таблица 2: Ошибки при численном анализе блока механического устройства регулировки, мм

esize, мм	n	Ошибка по методу		
		Уилкинсона	I	II
20	180875	$1,925 \cdot 10^{10}$	$6,342 \cdot 10^{-5}$	$1,031 \cdot 10^{-2}$
10	4221 456	$4,439 \cdot 10^{11}$	$8,345 \cdot 10^{-5}$	$1,500 \cdot 10^{-2}$
5	312212	$1,051 \cdot 10^{11}$	$1,144 \cdot 10^{-4}$	$2,504 \cdot 10^{-2}$
3	472130	$9,568 \cdot 10^{11}$	$4,821 \cdot 10^{-4}$	$3,071 \cdot 10^{-2}$

Из полученных результатов следует, что оценка ошибки по методу Уилкинсона является весьма завышенной по сравнению с оценками по методам I и II. Для рассматриваемых примеров метод Уилкинсона дает заведомо неадекватные результаты. Что касается оценок по методам I и II, они правдоподобны и непротиворечивы. Дальнейшую их сравнительную оценку выполним на базе разницы численного и аналитического решений (величины фактической ошибки).

Заключение

Опыт оценивания ошибок численных решений подтверждает, что одно из преимуществ обратного анализа ошибок состоит в том, что, применение его к хорошо обусловленной задаче с небольшой по величине обратной ошибкой приводит к малой величине прямой ошибки (глобальная ошибка для численных решений ОДУ). Нахождение некоторого множества конечного диаметра, которому принадлежит точное решение (хотя бы при достаточно малых погрешностях) и оценка расстояния от приближенного решения до его границы позволяет уменьшить влияние некорректности обратного анализ ошибок, проявляющуюся в отсутствии устойчивости. Можно охарактеризовать анализ ошибок численных решений как некорректно поставленную задачу. Регуляризация оператора решаемой задачи, заключающаяся в замене исходной системы на системы более простого вида, помогает получить более точные оценки ошибок.

Список литературы

[1] Калиткин Н.Н., Юхно Л.Ф., Кузьмина Л.В. Количественный критерий обусловленности систем линейных алгебраических уравнений //Математическое моделирование, 2011. Т. 23, № 2. С. 3–26.

[2] Wilkinson J. Rounding errors in algebraic processes. London. Her Majesty's Stationary Office. 1963. 161 p.

[3] Воеводин В. В. Об асимптотическом распределении ошибок округления при линейных преобразованиях // Ж. вычисл. матем. и матем. физ., 1967, том 7, номер 5. с. 965–976

[4] Воеводин В.В. Ошибки округления и устойчивость в прямых методах линейной алгебры. Изд-во МГУ, 1969. 153 стр.

[5] Рогалев А.Н. Обратные задачи оценки ошибки численных решений// Труды Международной конференции по вычислительной и прикладной математике "ВПМ'17"в рамках "Марчуковских научных чтений Новосибирск, 25 июня – 14 июля [Электрон. ресурс]. http://conf.nsc.ru/cam17/ru/proceedings, с. 739–743.

[6] Рогалев А.Н. Оценка точности численного анализа деформированного состояния силовых конструкций технических объектов / А.Н. Рогалев, С.В. Доронин, В.В. Москвичев // Вычислительные технологии. 2018. т.23, №2. с. 88–101.

[7] Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. Регуляризирующие алгоритмы и априорная информация. М.: Наука, 1983. 200 стр.

[8] Королев Ю.М., Ягола А. Г. Оценка погрешности в линейных обратных задачах при наличии априорной информации// Вычислительные методы и программирование. 2012. Т. 13. с 14–18.

[9] Reich S. Backward error analysis for numerical integrators // SIAM Journal on Numerical Analysis, 1999. V.36, No.5. pp. 1549–1570

[10] Corless R.M. Defect-controlled numerical methods and shadowing for chaotic differential equations // Physica D: Nonlinear Phenomena. 1992. 60(1-4). p. 323–334.

[11] Corless R.M. What good are numerical simulations of chaotic dynamical systems // Computers & Mathematics with Applications, 1994. 28(10-12). p. 107–121.

[12] Алексеев В.М. Об одной оценке возмущений решений обыкновенных дифференциальных уравнений. I. /В.М. Алексеев // Вестник Московского университета. Серия математика и механика — 1961 — №2. — с.28–36

[13] Skeel R. Thirteen ways to estimate global error. Numer. Math., 48 (1986), pp. 1–20.

Рогалев Алексей Николаевич — к.ф.-м.н., ст. науч.сотр. Института вычислительного моделирования СО РАН;
e-mail: rogalyov@icm.krasn.ru.
Дата поступления — 30 апреля 2019 г.

ИСПОЛЬЗОВАНИЕ ЕСТЕСТВЕННОГО ЯЗЫКА ПРИ СОЗДАНИИ ИМИТАЦИОННЫХ МОДЕЛЕЙ

С. В. Рудометов[1], В. В. Окольнишников[2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

УДК 519.685.1
DOI: 10.24411/9999-016A-2019-10066

В работе воспроизводятся действия человека, который пытается анализировать полностью незнакомый текст, в котором ему известно, какие слова и словесные конструкции обозначают сущности (предметы, существ), какие — действия, а какие — те или иные свойства сущностей или действий. При таком анализе, мысленно, человек начинает строить схему, в которой сначала «создает» сущности, затем «снабжает» их свойствами, указанными в тексте, и затем — выясняет их взаимодействие и его свойства.

Специалист в предметной области, разрабатывающий имитационную модель в среде имитационного моделирования MTSS [1], действует по похожей схеме. Сначала он принимает решение об использовании некоторых элементарных моделей (ЭМ), затем создает и размещает (визуально) некоторое количество экземпляров ЭМ (ЭЭМ), затем изменяет их свойства, затем графически соединяет их друг с другом. Использование естественного языка выглядит в таком случае как применение скриптового языка. Но есть и важное отличие: человеку не нужно учить этот новый язык. Он будет использовать хорошо известные ему и понятные языковые конструкции.

Ключевые слова: имитационное моделирование, естественный язык.

Введение

В последнее время активно развивается использование технологий искусственного интеллекта, таких как семантический анализ текстов на естественных языках.

Как мы уже отмечали ранее [2], «семантический анализ текста на естественном языке имеет смысл только в контексте какой-то задачи, сам по себе он не может существовать». Там же приводились наиболее известные задачи семантического анализа: машинный перевод, поиск смысловой информации в тексте (semantic web, определение эмоционального окраса текста, реферирование информации и т.д.), управление роботами или технологическим оборудованием.

Для предметного контекста можно использовать сущности, уже существующие в системах имитационного моделирования. Имеются в виду прежде всего субъекты (как они определены в DEVS-формализме) и их свойства.

Любое достаточно подробное и осмысленное описание проблемы, задачи и.т.д. может быть использовано для выделения подмножеств *понятий*, их *свойств* и их *активностей*. Эти три связанных подмножества затем могут быть использованы например для генерации исходного кода элементарных моделей (ЭМ) для создания новой библиотеки системы MTSS.

Сгенерированные таким образом ЭМ скорее всего будут неполными. Для их финального наполнения, отладки и верификации все равно потребуется участие программиста — специалиста в имитационном моделировании.

ISBN 978-5-901548-42-4

1 DEVS-формализм и элементы системы MTSS

В основу системы имитационного моделирования MTSS положен так называемый DEVS-формализм ([3]). Он является развитием дискретно-событийного подхода. DEVS описывает поведение *системы* на двух уровнях. В своей основе (первый уровень) система в нотации DEVS-формализма является дискретно-событийной имитационной моделью. Система входных и выходных портов, а также способ соединения простых *субъектов* для образования имитационной модели является DEVS-формализмом (второй уровень).

Более формально, имитационная модель представляется в виде наборов

$$M = \{s_j^i, i \in Class, j \in Inst^i\}, \{o^k, k \in Object\} \tag{1}$$

экземпляров субъектов разных классов. Здесь *Class* — множество классов объектов, $Inst^i$- множество экземпляров объектов класса i. Другими словами, множество M является набором экземпляров субъектов разных классов. Верхний индекс — номер класса, нижний индекс — номер экземпляра. Множество *Object* — это объекты, то есть сущности в имитационной модели, которые не участвуют в ее работе но с которыми могут взаимодействовать субъекты — например стенки, различные ограничители, разного рода указатели. Все эти классы определяются на основе декомпозиции имитируемой системы. s_j^i могут соединяться друг с другом и взаимодействовать с o^k.

Каждый субъект в нотации DEVS представляется в виде набора своих важных свойств:

$$s_j^i = \{S, ta, \delta_{int}, X, \delta_{ext}, Y, \lambda\} \tag{2}$$

Здесь S — фиксированный набор состояний (как определено в Discrete Event Simulation, или сокращенно DES), ta (или time advance) — время следующего исполнения субъекта, δ_{int} — набор входных портов, X — функция переключения состояний, как это определено для DES, δ_{ext} — набор выходных портов, Y- функция реакции на события во входных портах, λ — функция отображения внутреннего состояния из множества S в набор выходных сигналов.

2 Системы распознавания текста

Наиболее очевидным и как следствие — популярным способом синтаксического анализа фраз на естественном языке является т.н. «грамматика составляющих». Эта грамматика пытается разбить предложение на более мелкие структуры, затем их — на еще более мелкие структуры, пока не дойдет до отдельных слов или минимальных синтаксических структур. Каждая такая структура отмечается своим признаком, обозначающим ее роль в предложении (POS-tagging). Из такого определения следует, что результат синтаксического разбора исходного предложения в частности зависит от порядка слов в предложении. Этот подход успешно применяется для языков, в которых правилами жестко задан порядок слов (например — английский язык).

Другой возможный подход — грамматика зависимостей. В отличие от грамматики составляющих, результатом синтаксического разбора предложения в данной грамматике будет набор зависимостей между частями предложения (в частности, это подразумевает, что для предложения уже выполнен синтаксический разбор с использованием грамматики составляющих). Грамматика зависимостей выделяет в предложении значимые части (субъект, объект), и зависимости между ними. В данном случае порядок слов не так важен, например, как в русском языке.

И наконец третий, наиболее сложный подход — это системы на основе машинного обучения.

Сегодня системы распознавания текстов доступны как в виде библиотек, на и в виде онлайн-сервисов.

Примеры библиотек — Stanford NLP [4], Apache OpenNLP [5]. Библиотека должна быть включена в состав проекта в виде файлов библиотеки и файлов ресурсов. Существенной частью библиотеки является реализация языковой модели для какого-то естественного языка. Положительным моментом является то, что например обе эти библиотеки разрабатываются большим количеством инициативных групп, которые представляют разные естественные языки. И поэтому теоретически можно найти языковую библиотеку для своего языка. Но на практике получается следующая картина:

- хорошая языковая модель, как правило, очень большая. Она требует существенных ресурсов для своей работы — как времени загрузки, так и количества памяти. И если с временем загрузки еще можно (с трудом) смириться, то большой расход памяти в системе имитационного моделирования, которая сама по себе требует существенного количества памяти, является очень существенным ограничением.

- языковая модель оказывается «запаянной» в текущей версии MTSS. А значит для изменения или улучшения языковой модели требуется выпускать новые версии.

Другой вариант — использовать онлайн-сервисы, например Google NLP [6]. Такие сервисы абсолютно лишены недостатков библиотечных реализаций (загрузка, расход памяти), но требуют постоянного internet-соединения. Интерфейс этих сервисов достаточно прост — он выглядит как единственный метод «распознай текст», с посылкой запроса и получением ответа по технологии Representational State Transfer (REST). Также стоит отметить, что такие сервисы используют наиболее развитые языковые модели. Детали их реализации не доступны для конечного пользователя, но для системы имитационного моделирования существующей языковой модели вполне достаточно. Из недостатков онлайн — сервисов можно еще отметить, что они, как правило, являются платными, либо имеют триальные ограничения (количество распознаваний в сутки, ограниченный набор языков).

Стоит отметить, что каждый из рассмотренных подходов (статистический или на основе машинного обучения) имеет несколько неожиданное ограничение. Их создатели обучали (или анализировали) очень правильные с языковой точки зрения тексты (Шекспир, Толстой, Некрасов...). Поэтому и тексты для распознавания тоже должны быть написаны безупречно с точки зрения исходного естественного языка. Как правило, технические тексты таковыми не являются, и поэтому для более точного распознавания будет обязательно требоваться коррекция входного текста на естественном языке. Если же язык спецификации не поддерживается, то потребуется идеально перевести ее на поддерживаемый язук (например английский).

3 Пример задачи для имитационного моделирования с использованием естественного языка

Система MTSS предполагает наличие специализированной библиотеки ЭМ для решения определенных прикладных задач. На данный момент написано уже порядка 10 таких библиотек, в областях складской логистики, угледобычи, исследования компьютерных сетей. Каждая из этих библиотека создавалась программистами-специалистами в имитационном моделировании для решения конкретного круга задач, в кооперации со специалистами в предметной области.

При создании библиотеки требуется написать достаточно много программного кода. Он достаточно однотипный, и поэтому можно попробовать автоматизировать его написание.

А именно, из описания новой библиотеки, сделанного на естественном языке, потребуется выделить (согласно формулам 1 и 2) набор *Class* классов элементарных моделей, и для каждого класса определить внутренние состояния S, функции переключения состояний (алгоритмы) X и Y.

При выделении понятий (NOUN), алгоритмов (VERB) и свойств требуется явно указать, чему будут соответствовать выделенные сущности. Другими словами, произвести *теггирование*.

Существует простой и эффективный способ выделения нужных понятий из размеченного текста на естественном языке, схожий с обычными регулярными выражениями. Реализации можно найти, например, в Stanford NLP, TokenSequencePattern [7] или библиотеке Natural Language Toolkit для языка Python [8]. После выделения набора сущностей мы сопоставляем найденные слова с набором строк, которым идентифицирует (именует) себя каждая ЭМ в конкретной библиотеке MTSS. Похожим образом выделяются затем свойства найденных ЭЭМ, и затем – их взаимодействия. Такой процесс требует нескольких (как минимум трех) итераций в анализе одного и того же текста.

Для предложения (из [2]) «The combine moves from an insert point at an angle A to the horizon up, and after achievement of a roof of layer — at an angle bedding of layer B — deep into layer» после анализа сервисом google nlp (далее просто анализатор) и после теггирования внутри приложения:

- субъект (NOUN): «Combine»

- объекты (NOUN, сущности с которыми взаимодействуют субъекты): непосредственно из фразы: «insert point», «angle», «horizon», «roof», «layer», «bedding of layer».

- алгоритмы (VERB): ***moves*** (<субъект>**combine**, <точка>**от** (insert point, speed, angle(A, horizon)), <точка>**до**(не определено), <направление> **направление**(up, angle(A, horizon))), ***moves*** (<субъект>**combine**, <точка>**от**(не определено), <точка>**до**(не определено, <направле-ние> **направление**(forward, угол(B, layer))).

Здесь значения, заключенные в угловые скобки < и >, обозначают стереотипы данных и служат здесь для описания типов параметров.

4 Результат генерации

После теггирования полученные данные могут естественным образом быть применены для генерации кода ЭМ для новой библиотеки.

Алгоритмы функционирования после генерации будут представлены существующими, являющимися частью MTSS, опробованными реализациями модели движения с ускорением и торможением. Для моделей технологических процессов этого может оказаться достаточно. Но если это не так — данный код можно изменить.

Существенной проблемой остается визуальное представление элементов. При начальной генерации они будут представлены только в виде параллелепипедов, (форм, имеющих как положение в пространстве, так и линейные размеры). Если потребуется, в дальнейшем им можно придать любое графическое представление.

5 Дальнейшее развитие ЭМ

Как правило, такие сгенерированные ЭМ не удовлетворяют требованиям в постановке задач (хотя и являются вполне функциональными). Программист-специалист в имитационном моделировании может добавлять новые свойства ЭМ, реализовывать новые функциональные состояния ЭМ, изменять их графическое представление (2D и 3D). Результат генерации может рассматриваться в этом случае как достаточно полный «подстрочный перевод» того что хочет заказчик в программный код системы имитационного моделирования. А доработки ЭМ связаны с тем, что в работе специалиста в имитационном моделировании есть пункт «взаимодействие с заказчиком для выяснения очевидных для него но неочевидных для программиста деталей имитируемой системы».

6 Отличие от предыдущих работ

Работа [9] использовала фразы на естественном языке для конфигурирования конкретной имитационной модели. Работа [2] рассматривает аспекты использования естественного языка вообще, и предлагает использовать представление выделенных сущностей как простые, базовые элементы онтологии. Также она использует библиотеку Stanford NLP.

В данной работе используется сервис Google [6], и по сути своей результат является отдельным программным продуктом, не входящим в основную систему MTSS. Работа [2], может, и является более изящной в научном смысле, но с точки зрения практики лучше иметь именно такой генератор, как описан в данной статье. К сожалению, такой подход все-таки не позволяет полностью сократить «дистанцию программирования» (см. [2]), и является компромиссным, промежуточным между чистым использованием естественного языка и чистым программированием.

Список литературы

[1] Рудометов С. В. Визуально-интерактивная система имитационного моделирования технологических систем. Вестник СибГУТИ. 2011. No 3. С. 14–27.

[2] Рудометов С. В., Окольнишников В.В. Обобщенная онтология имитационного моделирования для создания имитационных моделей технологического процесса с использованием естественного языка. "Вычислительная и прикладная математика 2017"(ВПМ 2017). 25 — 30 июня 2017, Академгородок, Новосибирск, Россия. С. 744–749.

[3] Bernard Zeigler (1987). "Hierarchical, modular discrete-event modelling in an object-oriented environment". Simulation. 49 (5): pp 219–230.

[4] The Stanford NLP Group / [Электронный ресурс]. – Режим доступа: https://nlp.stanford.edu/ (дата обращения: 29.04.2018).

[5] Apache OpenNLP / [Электронный ресурс]. – Режим доступа: https://opennlp.apache.org/ (дата обращения: 29.04.2018).

[6] Cloud Natural Language / [Электронный ресурс]. – Режим доступа: https://cloud.google.com/natural-language/ (дата обращения: 29.04.2018).

[7] Stanford Log-linear Part-Of-Speech Tagger. URL: https://nlp.stanford.edu/software/tagger.shtml

[8] E. K. Steven Bird, Edward Loper, Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit. O'Reilly Media, 2009, p. 504

[9] Рудометов С. В., Окольнишников В.В. Использование онтологии технологических систем добычи угля для задания имитационных моделей на естественном языке. "Распределенные информационно-вычислительные ресурсы"(DICR-2014). 2–5 декабря 2014, Новосибирск, Россия.

Рудометов Сергей Валерьевич — к.т.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: sergersw@gmail.com;
Окольнишников Виктор Васильевич — д.т.н., вед. науч.сотр. Института вычислительных технологий СО РАН;
e-mail: okoln@mail.ru.
Дата поступления — 31 мая 2019 г.

МОДЕЛИРОВАНИЕ РАБОТЫ ОЧИСТНОГО ЗАБОЯ УГОЛЬНОЙ ШАХТЫ

С. В. Рудометов[1], В. В. Окольнишников[2], А. А. Ордин[2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*
[2] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

УДК 519.876.5
DOI: 10.24411/9999-016A-2019-10067

В системе имитационного моделирования MTSS реализована специализированная библиотека моделей горных машин, используемых при добыче угля в длинном очистном забое угольной шахты. С использованием специализированной библиотеки имитационных моделей горных машин разработана комплексная модель технологических процессов подземной добычи угля в очистном забое угольной шахты. Исследованы подходы повышения производительности очистного забоя в зависимости от технических характеристик горных машин, размеров лавы, технологических схем добычи угля в условиях изменяющихся горногеологических и геомеханических характеристик угольного пласта.

Ключевые слова: имитационное моделирование, шахта, очистной комбайн, производительность.

Введение

В настоящее время для многих угольных шахт возникли проблемы обоснованного принятия решений по повышению производительности, улучшению планирования добычи угля, использованию новых горных машин и новых перспективных технологий добычи угля. Для обоснования принятия решений наиболее подходящим способом является имитационное моделирование. На важность использования имитационного моделирования для поддержки принятия решений по проектированию, разработке и оптимизации угольных шахт указывает большое число публикаций [1–5].

Для решения этих проблем в Институте вычислительных технологий СО РАН была использована система имитационного моделирования MTSS [6, 7]. Система MTSS является визуально-интерактивной процессно-ориентированной системой дискретного имитационного моделирования, предназначенной для разработки и исполнения моделей технологических процессов. Отличительной чертой системы моделирования является ориентация на пользователей, являющихся специалистами конкретной предметной области (инженеров-технологов, горных инженеров), не имеющих опыта использования универсальных систем имитационного моделирования. Быстрое построение моделей обеспечивается наличием визуально-интерактивного интерфейса и специализированных библиотек моделей технологического оборудования конкретных предметных областей. Система моделирования MTSS предоставляет пользователю следующие возможности: визуально-интерактивное построение модели, задание параметров модели, различные режимы выполнения модели, 2D и 3D визуализацию выполнения модели.

Для моделирования технологических процессов в угольных шахтах в системе моделирования MTSS разработана специализированная библиотека технологического оборудования для таких подсистем угольной шахты, как конвейерная подсистема, подсистема водоотлива и подсистема электроснабжения. С использованием специализированной библиотеки разработан ряд моделей этих подсистем для угольных шахт Кузбасса [8]. В статье рассматривается использование системы MTSS для моделирования технологических процессов добычи угля в очистном забое.

1 Математическая модель

В [9,10] предложены формулы (1) и (2) для расчета скорости V и производительности A очистного комбайна.

ISBN 978-5-901548-42-4

$$V = \frac{30N\eta n_1 K_1}{fP\cos\alpha \pm P\sin\alpha + Smn_2K_0(m))K} \tag{1}$$

Здесь N — мощность электродвигателя рабочего органа, η — КПД редуктора исполнительного органа, n_1 — количество резцов в одной линии резания, K_1 — коэффициент, учитывающий затраты мощности на перемещение комбайна, f — коэффициент трения между комбайном и скребковым конвейером, P — вес комбайна, α — угол наклона угольного пласта. Знаки + (−) соответствуют движению комбайна вверх (вниз) по лаве, соответственно, S — сопротивляемость угольного пласта резанию, m — вынимаемая мощность угольного пласта, n_2 — количество резцов, одновременно разрушающих забой, K_0 — коэффициент отжима, учитывающий уменьшение сил резания вследствие горного давления, K — коэффициент, учитывающий угол резания, ширину резца, затупление и форму резцов.

$$A = \gamma rmV = \frac{30\gamma rN\eta n_1 K_1}{\frac{P}{m}(f\cos\alpha \pm \sin\alpha) + Sn_2K_0(m)K} \tag{2}$$

Здесь γ — плотность угля, r — ширина захвата комбайна.

В формулах (1) и (2) используются средние горногеологические и геомеханические характеристики. Для более корректного моделирования технологических процессов при добыче угля в очистном забое в этой работе были использованы не средние горногеологические и геомеханические характеристики, а их функциональные зависимости от текущих координат (x, y) очистного комбайна. Скорость V и производительность A очистного комбайна вычисляются по модифицированным формулам (3) и (4).

$$V(x,y) = \frac{30N\eta n_1 K_1}{fP\cos\alpha(x,y) \pm P\sin\alpha(x,y) + S(x,y)m(x,y)n_2K_0(m(x,y))K} \tag{3}$$

$$A(x,y) = \frac{30\gamma(x,y)rN\eta n_1 K_1}{\frac{P}{m(x,y)}(f\cos\alpha(x,y) \pm \sin\alpha(x,y)) + S(x,y)n_2K_0(m(x,y))K} \tag{4}$$

где $x \le L$, $y \le Lc$, L — длина лавы, Lc — длина выемочного столба.

Для вычисления распределенных горногеологических и геомеханических характеристик угольного пласта $(\gamma, \alpha, S, m, K_0(m))$ был использован метод обратных расстояний IDW(Inverse Distance Weighting) [11], формула (5).

$$F(x,y) = \begin{cases} \dfrac{\sum\limits_{i=1}^{n} d_i^{-2} F_i}{\sum\limits_{i=1}^{n} d_i^{-2}}, & \text{если } d_i \neq 0 \\ F_i, & \text{если } d_i = 0 \end{cases} \tag{5}$$

где $F(x, y)$ — расчетное значение геомеханического показателя пласта в точке с координатами (x, y) очистного комбайна, n — количество геологоразведочных скважин, пересекающих выемочный столб, F_i — значение геомеханического показателя в i-ой скважине, d_i — расстояние между i–ой геологоразведочной скважиной и местоположением очистного комбайна в точке с координатами (x, y).

2 Имитационная модель

Формула (4) определяет максимальную производительность очистного комбайна при условии непрерывного прямолинейного движения очистного комбайна. В общем случае на производительность очистного комбайна в очистном забое влияют следующие факторы: горногеологическое и геомеханическое состояние угольного пласта, технические параметры очистного комбайна, технологическая схема работы очистного комбайна, передвижение секций механизированной крепи, работа скребкового конвейера, содержание газа метана.

Существует две наиболее используемые технологические схемы работы очистного комбайна: челночная и односторонняя. При челночной схеме цикл работы очистного комбайна состоит из движения комбайна в прямом и обратном ходе в рабочем режиме. При односторонней схеме цикл работы очистного комбайна состоит из движения комбайна в прямом ходе в рабочем режиме, а в обратном ходе очистной комбайн движется с маневровой скоростью и производит зачистку отбитого угля с почвы пласта.

Поскольку очистной комбайн не может “развернуться”, то на границах лавы осуществляется так называемый “косой заезд”. Косой заезд — это сложный технологический процесс, связанный с маневрированием комбайна, занимающий значительное время, приводящий к потере угля. При челночной схеме за один цикл работы комбайна осуществляется два косых заезда, при односторонней схеме — один. Кроме того, при косом заезде происходит “искривление” скребкового конвейера, что приводит к увеличению времени выполнения косого заезда.

Особенностью реализованной имитационной модели является детальная имитация движения очистного комбайна, включая косой заезд, взаимодействие движения очистного комбайна с другими горными машинами (механизированными крепями и скребковым конвейером), распределенное вычисление горногеологических и геомеханических состояний угольного пласта, включая газоносность, на основании данных геологоразведочных скважин. Имитационные эксперименты с реализованной моделью позволяют получать теоретически и модельно обоснованные результаты для сравнения и решения оптимизационных задач.

Перед запуском модели на исполнение пользователь с помощью системы меню выбирает состав горных машин, участвующих в имитационном эксперименте (очистной комбайн, механизированные крепи, конвейеры). При выборе некоторой горной машины все технические характеристики этой машины загружаются в модель. Далее пользователь задает число и координаты геологоразведочных скважин, а также значения горногеологических и геомеханических характеристик угольного пласта в этих скважинах. После этого пользователь задает линейные размеры забоя, например, длину лавы, выбирает технологическую схему работы очистного комбайна, задает режимы исполнения модели, формы представления выходных данных модели и запускает модель на исполнение.

На рис. 1 представлено главное окно имитационной модели технологических процессов подземной добычи угля в очистном забое, реализованной с помощью системы имитационного моделирования MTSS.

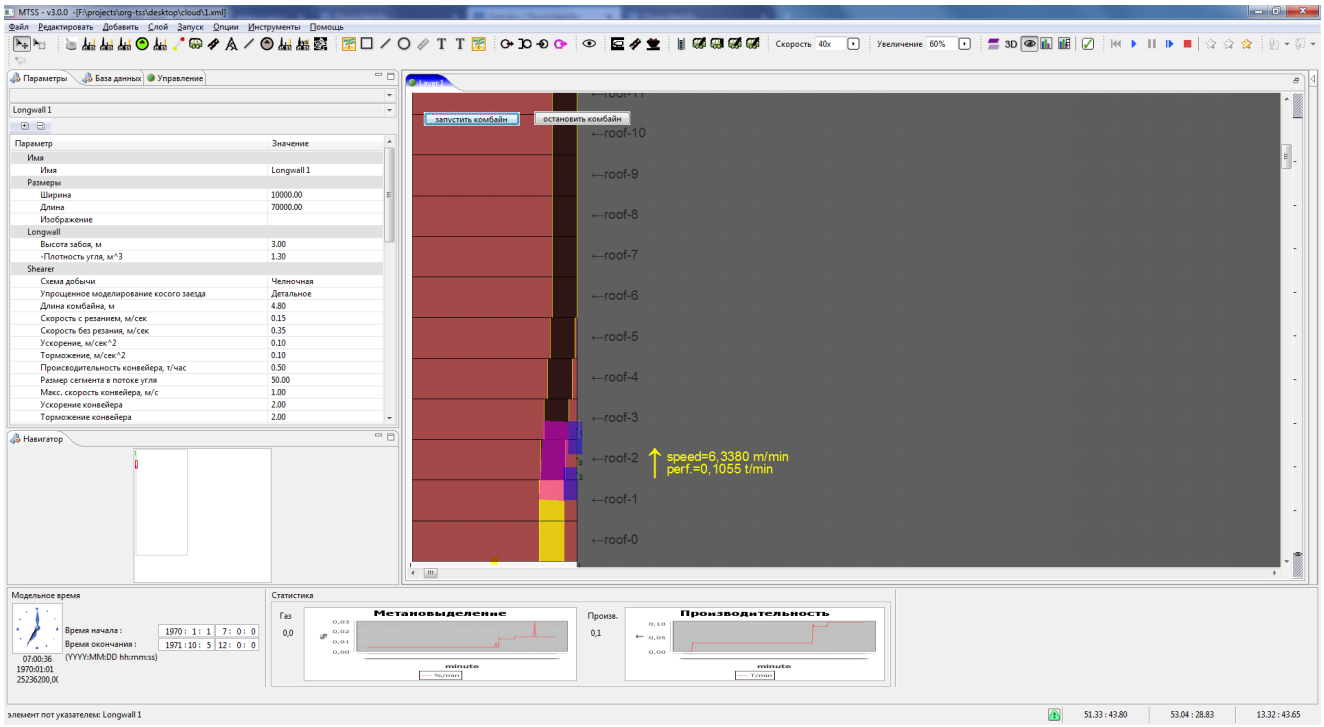


Рис. 1: Главное окно имитационной модели

При исполнении имитационной модели вычисляются текущие значения всех выходных параметров. Эти текущие значения могут выводиться в указанные пользователем индикаторы в главном окне модели или в виде графиков в главном окне модели или в новых окнах. Например, на рис. 2 приведен график суммарного количества добытого угля. “Неравномерности” графика соответствуют изменению скорости или остановке очистного комбайна.

Визуализация работы очистного комбайна увеличивает время исполнения модели. Для получения интегральных выходных данных визуализация может быть отключена.

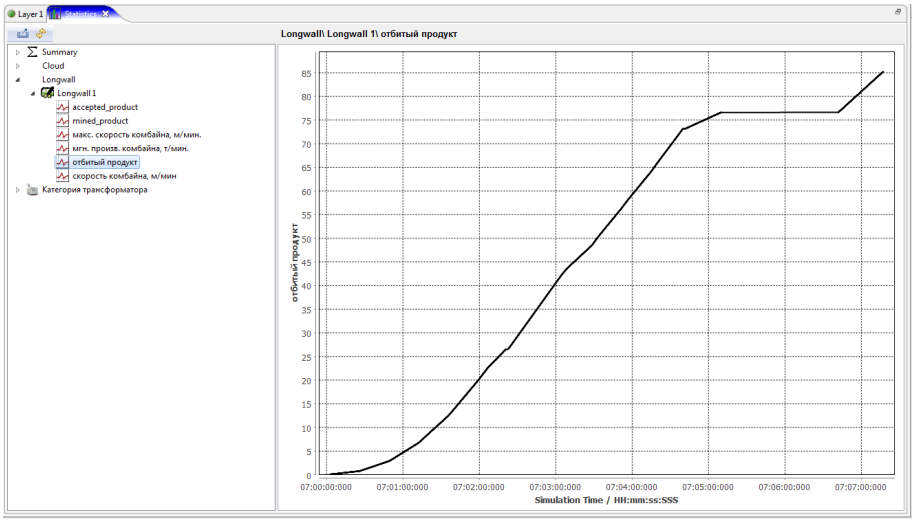


Рис. 2: Суммарное количество добытого угля

Заключение

При выполнении работы были получены следующие основные результаты:

- В рамках системы моделирования MTSS разработана комплексная модель технологических процессов подземной добычи угля в очистном забое угольной шахты, учитывающая распределенные значения горногеологических и геомеханических состояний угольного пласта, технические параметры используемых горных машин, технологические схемы добычи угля.
- С использованием комплексной модели выполнен ряд имитационных экспериментов для оценки производительности очистного комбайна в зависимости от перечисленных выше факторов.

Список литературы

[1] Salama A., Greberg J., Schunnesson H. The use of discrete event simulation for underground haulage mining equipment selection // International Journal of Mining and Mineral Engineering. 2014. V. 5, iss. 3. P. 256–271.

[2] Fioroni M., Santos L., Franzese L., Santana I., Telles G., Seixas J., Penna B., Alkmim G. Logistic Evaluation of an Underground mine Using Simulation // Proceedings of Winter Simulation Conference. 2014. P. 1855–1865.

[3] Michalakopoulos T. N., Roumpos C. P., Galetakis M. J., Panagiotou G. N. Discrete-Event Simulation of Continuous Mining Systems in Multi-layer Lignite Deposits // Lecture Notes in Production Engineering. Proceedings of the 12th International Symposium Continuous Surface Mining. 2015. P. 225–239.

[4] Gospodarczyk P. Modeling and Simulation of Coal Loading by Cutting Drum in Flat Seams //Archives of Mining Sciences. 2016. V. 61, iss. 2. P. 385–379.

[5] Gao Y., Liu D., Zhang X., He M. Analysis and Optimization of Entry Stability in Underground Longwall Mining // Sustainability. 2017. V. 9, iss. 11. P. 2079–2088.

[6] Рудометов С. В. Визуально-интерактивная система имитационного моделирования технологических систем // Вестник СибГУТИ. 2011. № 3. С. 14–26.

[7] Okolnishnikov V. V., Rudometov S. V. A System for Computer Simulation of Technological Pro-cesses. // St. Petersburg State Polytechnic University Journal. Computer Science. Telecommunications and Control Systems. 2014. V. 181, iss. 1. P. 62–68.

[8] Okolnishnikov V., Rudometov S., Zhuravlev S. Simulating the Various Subsystems of a Coal Mine. // Engineering, Technology & Applied Science Research. 2016. V. 6, iss. 3. P. 993–999.

[9] Ordin A. A., Metel'kov A. A. Analysis of longwall face output in screw-type cutter–loader-and-scraper conveyor system in underground mining of flat-lying coal beds // Journal of Mining Science. 2015. V. 51, iss. 6. P. 1173–1179.

[10] Ордин А. А., Никольский А. М. Оптимизация ширины захвата и производительности шнекового комбайна при отработке пологого угольного пласта длинным очистным забоем // Физикотехнические проблемы разработки полезных ископаемых. 2018. № 1. С. 79–86.

[11] Дэвис Дж. Статистический анализ данных в геологии. М.: Недра. 1990. 319 с.

Рудометов Сергей Валерьевич — к.т.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: rsw@inbox.ru;
Окольнишников Виктор Васильевич — д.т.н., вед.науч.сотр. Института вычислительных технологий СО РАН;
e-mail: okoln@mail.ru;
Ордин Александр Александрович — д.т.н., вед.науч.сотр. Института вычислительных технологий СО РАН;
e-mail: ordin@misd.ru.
Дата поступления — 30 апреля 2019 г.

СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ ОПЕРАТОРОВ ДИФФЕРЕНЦИРОВАНИЯ И ИНТЕГРИРОВАНИЯ ОТНОСИТЕЛЬНО ОРТОГОНАЛЬНЫХ ФИНИТНЫХ ФУНКЦИЙ. I. СИММЕТРИЧНЫЕ СПЛАЙНЫ

К. А. Рыбаков, В. В. Рыбин

Московский авиационный институт (национальный исследовательский университет), 125993, Москва

УДК 519.711.3
DOI: 10.24411/9999-016A-2019-10068

В первой части работы сформированы алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования относительно системы ортогональных финитных функций, которая порождается с помощью масштабирования и сдвига симметричного кусочно-линейного сплайна Леонтьева.

Ключевые слова: ортогональные финитные функции, симметричный сплайн Леонтьева, спектральная форма математического описания, спектральный метод, финитная базисная система.

Введение

Спектральная форма математического описания и спектральный метод применяются в различных задачах теории автоматического управления для систем как с сосредоточенными, так и распределенными параметрами при детерминированных и случайных воздействиях [3, 4, 7, 9]. Истоки спектрального метода для расчета линейных нестационарных систем автоматического управления лежат в представлении сигналов и их временных характеристик в виде ортогональных рядов по выбранной базисной системе функций. Коэффициенты этих рядов, отделенные от самих рядов, рассматриваются как характеристики сигналов и систем управления [8, 9]. Отличительной особенностью спектральной формы математического описания и спектрального метода является то, что все операции при решении задач производятся не с ортогональным рядом, а с упорядоченным набором коэффициентов при базисных функциях, который в общем случае представляется в виде бесконечной матрицы-столбца и называется спектральной характеристикой.

Для линейных преобразований спектральных характеристик функций (сигналов) используются спектральные характеристики линейных операторов, которые представляются бесконечными матрицами. В первую очередь применяются операторы умножения, дифференцирования и интегрирования. Они соответствуют элементарным звеньям линейных систем: усилительному, дифференцирующему и интегрирующему. Для эффективного применения спектральной формы математического описания разработаны пакеты прикладных программ в виде расширений систем компьютерной математики (*Mathcad*, *Matlab*, *Maple* и *Mathematica*), а также в виде отдельных приложений (*Спектр*, *Spectrum*) для аналитического и численного расчета спектральных характеристик [5]. При приближенном решении задач теории автоматического управления спектральным методом все спектральные характеристики «усекаются», т.е. осуществляется переход от ортогональных рядов к частичным суммам, либо изначально используются базисные системы с конечным числом функций.

С момента появления спектрального метода формировалось соответствующее алгоритмическое обеспечение: алгоритмы расчета спектральных характеристик функций и линейных операторов для различных

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-08-00530).

ISBN 978-5-901548-42-4

базисных систем [5, 9, 10]. В этой работе, состоящей из двух частей, сформированы алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования относительно ортогональных кусочно-линейных финитных функций, определенных в [1]. В первой части рассмотрен симметричный сплайн, а во второй части — несимметричный сплайн. Такие сплайны будем называть сплайнами Леонтьева.

1 Ортонормированные базисные системы

Пусть система функций $\{q(i,t)\}_{i=0}^{\infty}$, определенных на отрезке $\mathrm{T} = [0,T]$ со значениями в $\mathbb{R}$, является ортонормированной:

$$\big(q(i,t),q(j,t)\big)_{L_2(\mathrm{T})} = \int_{\mathrm{T}} q(i,t)q(j,t)dt = \delta_{ij}, \quad i,j = 0,1,2,\ldots,$$

где δ_{ij} — символ Кронекера, и $x(t)\colon \mathrm{T} \to \mathbb{R}$ — квадратично интегрируемая функция:

$$\int_{\mathrm{T}} x^2(t)dt < \infty \quad \big(x(t) \in L_2(\mathrm{T})\big).$$

Тогда

$$x(t) = \sum_{i=0}^{\infty} X_i q(i,t), \quad \sum_{i=0}^{\infty} X_i^2 < \infty,$$

где

$$X_i = \big(q(i,t),x(t)\big)_{L_2(\mathrm{T})} = \int_{\mathrm{T}} q(i,t)x(t)dt, \quad i = 0,1,2,\ldots,$$

при этом функциональный ряд сходится к $x(t)$ по норме $L_2(\mathrm{T})$, если система функций $\{q(i,t)\}_{i=0}^{\infty}$ полна (такие системы функций будем называть базисными).

Упорядоченный набор X коэффициентов разложения X_i функции $x(t)$ по базисной системе $\{q(i,t)\}_{i=0}^{\infty}$ представляется в виде бесконечной матрицы-столбца и называется нестационарной спектральной характеристикой функции $x(t)$. Этот термин был введен в [9], нестационарность состояла в том, что границы отрезка T в общем случае могли быть функциями времени, однако здесь ограничимся фиксированным отрезком и будем называть X и аналогичные ей характеристики спектральными характеристиками.

Наряду с полными базисными системами широкое применение нашли финитные базисные системы $\{q(i,t)\}_{i=0}^{L-1}$, для них в приведенных выше формулах индексы i и j ограничены: $i,j = 0,1,\ldots,L-1$, а представление функции $x(t)$ частичной суммой в общем случае выполняется лишь приближенно. В частности финитные базисные системы можно определять с помощью масштабирования и сдвига некоторой порождающей финитной функции (сочетание проекционного-сеточного [2] и спектрального методов [7]). Спектральная характеристика в этом случае — L-мерный вектор.

2 Симметричные кусочно-линейные финитные функции

Определим две вспомогательные финитные функции:

$$B_0^*(t) = \begin{cases} 1, & t \in \left[-\frac{1}{2},\frac{1}{2}\right), \\ 0, & t \notin \left[-\frac{1}{2},\frac{1}{2}\right), \end{cases} \qquad \Phi_0^*(\alpha,\theta,t) = \begin{cases} -\frac{\alpha}{\theta}, & t \in \left[-\frac{1}{2},-\frac{1}{2}+\theta\right), \\ \frac{1+2\alpha}{1-2\theta}, & t \in \left[-\frac{1}{2}+\theta,\frac{1}{2}-\theta\right], \\ -\frac{\alpha}{\theta}, & t \in \left(\frac{1}{2}-\theta,\frac{1}{2}\right], \\ 0, & t \notin \left[-\frac{1}{2},\frac{1}{2}\right], \end{cases}$$

где α и θ — параметры: $\alpha > 0$, $0 < \theta < \frac{1}{2}$.

Симметричный сплайн степени p можно выразить через функции $\Phi_0^*(\alpha,\theta,t)$ и $B_0^*(t)$ (B-сплайн нулевой степени) следующим образом [1]:

$$\Phi_p^*(\alpha,\theta,t) = \Phi_{p-1}^*(\alpha,\theta,t) * B_0^*(t) = \int_{-\infty}^{+\infty} \Phi_{p-1}^*(\alpha,\theta,\tau)B_0^*(t-\tau)d\tau, \quad p = 1,2,3,\ldots$$

Ограничимся случаем $p = 1$. Тогда

$$\Phi_1^*(\alpha,\theta,t) = \begin{cases} \frac{\alpha}{\theta}t + 1, \ t \in [0,\theta], \\ -\frac{1+2\alpha}{1-2\theta}\left(t - \frac{1}{2}\right) + \frac{1}{2}, \ t \in [\theta, 1-\theta], \\ \frac{\alpha}{\theta}\left(t-1\right), \ t \in [1-\theta, 1], \\ 0, \ t > 1, \end{cases}$$

при дополнительном условии четности, т.е. $\Phi_1^*(\alpha,\theta,-t) = \Phi_1^*(\alpha,\theta,t)$.

Кусочно-линейный сплайн $\Phi_1^*(\alpha,\theta,t)$ порождает финитную базисную систему на отрезке $\mathrm{T} = [0,T]$ с помощью масштабирования и сдвига согласно правилу

$$\Phi_1(i,t) = \frac{1}{\sqrt{h}}\,\Phi_1^*\left(\alpha,\theta,\frac{t}{h} - i\right), \quad h = \frac{T}{L-1}, \quad t \in \mathrm{T}, \quad i = 0,1,\ldots,L-1.$$

В общем случае матрица Грама для системы функций $\{\Phi_1(i,t)\}_{i=0}^{L-1}$ является трехдиагональной, так как носители функций этой системы с номерами, отличающимися на единицу, имеют непустое пересечение, однако при условии

$$\theta = \frac{1}{4} \quad \text{и} \quad \alpha = \frac{\sqrt{41}-5}{8} \tag{1}$$

скалярные произведения функций $\{\Phi_1(i,t)\}_{i=0}^{L-1}$ с номерами, отличающимися на единицу, равны нулю и $\{\Phi_1(i,t)\}$ — ортогональная система функций. Свойство ортогональности позволяет не использовать процедуры ортогонализации или построения двойственной базисной системы [4].

Отметим некоторые свойства функции $\Phi_1^*(\alpha,\theta,t)$, которые будут использованы в дальнейшем, в частности

$$\int_{-1}^{1}\left(\Phi_1^*(\alpha,\theta,t)\right)^2 dt = 1, \quad \int_{0}^{1}\left(\Phi_1^*(\alpha,\theta,t)\right)^2 dt = \int_{-1}^{0}\left(\Phi_1^*(\alpha,\theta,t)\right)^2 dt = \frac{1}{2},$$

т.е. условие нормировки выполнено для функций $\{\Phi_1(i,t)\}_{i=0}^{L-1}$ с номерами $i = 1,2,\ldots,L-2$, а для функций с номерами $i = 0$ и $i = L-1$ нужен дополнительный нормировочный коэффициент $\sqrt{2}$. Таким образом, функции

$$\hat{\Phi}_1(i,t) = \frac{\gamma_i}{\sqrt{h}}\,\Phi_1^*\left(\alpha,\theta,\frac{t}{h} - i\right) = \gamma_i\Phi_1(i,t), \quad i = 0,1,\ldots,L-1,$$

где $\gamma_0 = \gamma_{L-1} = \sqrt{2}$ и $\gamma_1 = \ldots = \gamma_{L-2} = 1$, образуют ортонормированную систему на отрезке $\mathrm{T} = [0,T]$.

3 Спектральные характеристики операторов дифференцирования и интегрирования

При использовании спектральной формы математического описания возникает необходимость использования спектральных характеристик операторов дифференцирования и интегрирования (двумерных нестационарных передаточных функций дифференцирующего и интегрирующего звеньев). При выборе финитных базисных систем они представляют собой квадратные матрицы порядка L, столбцы этих матриц — спектральные характеристики производных и первообразных соответственно функций $\{q(i,t)\}_{i=0}^{L-1}$.

Спектральную характеристику оператора дифференцирования будем обозначать $\mathcal{P}$, а спектральную характеристику оператора интегрирования — P^{-1}, их элементы определяются формулами

$$\mathcal{P}_{ij} = \int_{\mathrm{T}} q(i,t)q'(j,t)dt, \quad P_{ij}^{-1} = \int_{\mathrm{T}} q(i,t)\int_0^t q(j,\tau)d\tau dt, \quad i,j = 0,1,\ldots,L-1. \tag{2}$$

Найдем производную функции $\Phi_1^*(\alpha,\theta,t)$ на интервале $(-1,1)$, из которого исключены $0, \pm\theta, \pm(1-\theta)$:

$$\Phi_1^{*\prime}(\alpha,\theta,t) = \begin{cases} \frac{\alpha}{\theta}, \ t \in (0,\theta), \\ -\frac{1+2\alpha}{1-2\theta}, \ t \in (\theta, 1-\theta), \\ \frac{\alpha}{\theta}, \ t \in (1-\theta, 1), \\ 0, \ t > 1, \end{cases}$$

при дополнительном условии нечетности, т.е. $\Phi_1^{*\prime}(\alpha,\theta,-t) = -\Phi_1^{*\prime}(\alpha,\theta,t)$. То, что производная функции $\Phi_1^*(\alpha,\theta,t)$ не определена в точках $0, \pm\theta, \pm(1-\theta)$, не имеет принципиального значения для нахождения элементов $\mathcal{P}_{ij}$, поскольку интегралы в (2) не зависят от значения подынтегральной функции на множестве нулевой меры.

Отметим, что при выводе соотношений для расчета элементов спектральных характеристик линейных операторов, в частности операторов дифференцирования и интегрирования, как правило, используются различные свойства функций базисной системы. Например, для ортогональных полиномов это рекуррентные формулы, связывающие базисные функции с их производными и первообразными; для тригонометрических функций это тригонометрические тождества и т.п. Для финитных базисных систем целесообразно использовать правило, согласно которому они определяются: масштабирование и сдвиг порождающей финитной функции.

Запишем соотношения, необходимые для расчета спектральной характеристики оператора дифференцирования (здесь и далее интегрирование по отрезку $[0,1]$ используется для учета функции финитной базисной системы с номером $i=0$, по отрезку $[-1,1]$ — с номерами $i=1,2,\ldots,L-2$ и по отрезку $[-1,0]$ — с номером $i=L-1$), а именно

— для элементов главной диагонали:

$$\int_{-1}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t)dt = 0, \quad \int_{0}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t)dt = -\int_{-1}^{0} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t)dt = -\frac{1}{2};$$

— для элементов наддиагонали и поддиагонали:

$$\int_{-1}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t-1)dt = \int_{0}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t-1)dt = \frac{1}{2},$$
$$\int_{-1}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t+1)dt = \int_{-1}^{0} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t+1)dt = -\frac{1}{2};$$

— для остальных элементов ($k>1$):

$$\int_{-1}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t-k)dt = \int_{0}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t-k)dt = 0,$$
$$\int_{-1}^{1} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t+k)dt = \int_{-1}^{0} \Phi_1^*(\alpha,\theta,t)\Phi_1^{*\prime}(\alpha,\theta,t+k)dt = 0,$$

поскольку функции $\Phi_1^*(\alpha,\theta,t)$ и $\Phi_1^{*\prime}(\alpha,\theta,t\pm k)$ имеют непересекающиеся носители.

Так так

$$\hat{\Phi}_1'(i,t) = \frac{\gamma_i}{h\sqrt{h}}\,\Phi_1^{*\prime}\left(\alpha,\theta,\frac{t}{h}-i\right), \quad i=0,1,\ldots,L-1,$$

получаем $\mathcal{P}_{ij} = (1/h)\,C_{ij}$, где

$$C_{00} = -1, \quad C_{L-1,L-1} = 1; \quad C_{11} = \ldots = C_{L-2,L-2} = 0 \quad (L>2);$$
$$C_{01} = 1 \quad (L=2); \quad C_{01} = C_{L-2,L-1} = \frac{\sqrt{2}}{2} \quad (L>2); \quad C_{12} = \ldots = C_{L-3,L-2} = \frac{1}{2} \quad (L>3);$$
$$C_{10} = -1 \quad (L=2); \quad C_{10} = C_{L-1,L-2} = -\frac{\sqrt{2}}{2} \quad (L>2); \quad C_{21} = \ldots = C_{L-2,L-3} = -\frac{1}{2} \quad (L>3);$$
$$C_{i-k,i} = C_{i,i-k} = 0, \quad i=2,3,\ldots,L-1, \quad k=2,3,\ldots,i \quad (L>2),$$

следовательно,

$$\mathcal{P} = \frac{1}{h}\cdot\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \quad (L=2); \quad \mathcal{P} = \frac{1}{h}\cdot\begin{pmatrix} -1 & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & 1 \end{pmatrix} \quad (L=3);$$

$$\mathcal{P} = \frac{1}{h} \cdot \begin{pmatrix} -1 & \frac{\sqrt{2}}{2} & 0 & 0 & \dots & 0 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{1}{2} & 0 & \dots & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \ddots & -\frac{1}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 0 & 0 & \dots & 0 & -\frac{\sqrt{2}}{2} & 1 \end{pmatrix} \quad (L > 3).$$

Помимо спектральной характеристики $\mathcal{P}$ оператора дифференцирования в спектральной форме математического описания важное значение имеет спектральная характеристика P оператора дифференцирования с учетом начального значения функции [4, 9]. Она определяется в виде $P = \mathcal{P} + q(0) \cdot q^{\mathrm{T}}(0)$, где $q(0)$ — матрица столбец значений функций базисной системы в точке $t = 0$, т.е. P — квадратная матрица порядка L с элементами

$$P_{ij} = \int_{\mathrm{T}} q(i,t)q'(j,t)dt + q(i,0)q(j,0) = \mathcal{P}_{ij} + q(i,0)q(j,0), \quad i,j = 0,1,\dots,L-1.$$

Все функции финитной базисной системы $\{\hat{\Phi}_1(i,t)\}_{i=0}^{L-1}$ принимают нулевое значение при $t = 0$ за исключением $\hat{\Phi}_1(0,t)$. Значение последней определяется величиной $\Phi_1^*(\alpha,\theta,0) = 1$ и нормировочным коэффициентом $\sqrt{2/h}$. Таким образом, в матрице $q(0) \cdot q^{\mathrm{T}}(0)$ первый элемент первой строки равен $2/h$, остальные элементы — нулевые, следовательно,

$$P = \frac{1}{h} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (L = 2); \quad P = \frac{1}{h} \cdot \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & 1 \end{pmatrix} \quad (L = 3);$$

$$P = \frac{1}{h} \cdot \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} & 0 & 0 & \dots & 0 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{1}{2} & 0 & \dots & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \ddots & -\frac{1}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 0 & 0 & \dots & 0 & -\frac{\sqrt{2}}{2} & 1 \end{pmatrix} \quad (L > 3).$$

Далее приведем две первообразных функции $\Phi_1^*(\alpha,\theta,t)$:

$$\int_0^t \Phi_1^*(\alpha,\theta,\tau)d\tau = \begin{cases} \frac{t(\alpha t+2\theta)}{2\theta},\ t\in[0,\theta), \\ \frac{\alpha\theta}{2}+\theta+\frac{(t-\theta)((t+\theta-1)(2\alpha+1)-2(\alpha-\theta+1))}{2(2\theta-1)},\ t\in[\theta,1-\theta), \\ \frac{1}{2}+\frac{\alpha\theta}{2}+\frac{\alpha(t+\theta-1)(t-\theta-1)}{2\theta},\ t\in[1-\theta,1), \\ \frac{1}{2},\ t\geqslant 1, \end{cases} \qquad t\geqslant 0;$$

$$\int_{-1}^t \Phi_1^*(\alpha,\theta,\tau)d\tau = \begin{cases} -\frac{\alpha(t+1)}{2\theta},\ t\in[-1,\theta-1), \\ -\frac{\alpha\theta}{2}-\frac{(t-\theta+1)(t-\theta+1+2\alpha(t+\theta))}{2(2\theta-1)},\ t\in[\theta-1,-\theta), \\ \frac{1-\alpha\theta}{2}-\theta-\frac{(t+\theta)(\alpha(t-\theta)-2\theta)}{2\theta},\ t\in[-\theta,0), \\ \frac{1}{2}+\frac{t(\alpha t+2\theta)}{2\theta},\ t\in[0,\theta), \\ \frac{1+\alpha\theta}{2}+\theta+\frac{(t-\theta)((t+\theta-1)(2\alpha+1)-2(\alpha-\theta+1))}{2(2\theta-1)},\ t\in[\theta,1-\theta), \\ 1+\frac{\alpha\theta}{2}+\frac{\alpha(t+\theta-1)(t-\theta-1)}{2\theta},\ t\in[1-\theta,1), \\ 1,\ t\geqslant 1, \end{cases} \qquad t\geqslant -1.$$

С учетом (1) получаем соотношения, необходимые для расчета спектральной характеристики оператора интегрирования, а именно

— для элементов главной диагонали:

$$\int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_{-1}^t \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{2},$$

$$\int_0^1 \Phi_1^*(\alpha,\theta,t)\int_0^t \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \int_{-1}^0 \Phi_1^*(\alpha,\theta,t)\int_{-1}^t \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{8};$$

— для элементов наддиагонали и поддиагонали за исключением первого столбца и последней строки:

$$\int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t-1} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \int_0^1 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t-1} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{17}{192}-\frac{\sqrt{41}}{64},$$

$$\int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t+1} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = 1-\frac{17}{192}+\frac{\sqrt{41}}{64} = \frac{175}{192}+\frac{\sqrt{41}}{64};$$

— для остальных элементов за исключением первого столбца и последней строки (здесь и далее $k>1$):

$$\int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t-k} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = 0, \quad \int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t+k} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = 1;$$

— для оставшихся элементов первого столбца:

$$\int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_0^{t+1} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{2}-\frac{17}{192}+\frac{\sqrt{41}}{64} = \frac{79}{192}+\frac{\sqrt{41}}{64},$$

$$\int_{-1}^0 \Phi_1^*(\alpha,\theta,t)\int_0^{t+1} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{4}-\frac{17}{192}+\frac{\sqrt{41}}{64} = \frac{31}{192}+\frac{\sqrt{41}}{64},$$

$$\int_{-1}^1 \Phi_1^*(\alpha,\theta,t)\int_0^{t+k} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{2}, \quad \int_{-1}^0 \Phi_1^*(\alpha,\theta,t)\int_0^{t+k} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{4};$$

— для оставшихся элементов последней строки:

$$\int_{-1}^0 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t+k} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{1}{2}, \quad \int_{-1}^0 \Phi_1^*(\alpha,\theta,t)\int_{-1}^{t+1} \Phi_1^*(\alpha,\theta,\tau)d\tau dt = \frac{79}{192}+\frac{\sqrt{41}}{64}.$$

Так так

$$\int_0^t \hat{\Phi}_1(i,\tau)d\tau = \begin{cases} \frac{\gamma_0 h}{\sqrt{h}}\int_0^t \Phi_1^*\Big(\alpha,\theta,\frac{\tau}{h}\Big)d\Big(\frac{\tau}{h}\Big), & i=0, \\ \frac{\gamma_i h}{\sqrt{h}}\int_{-1}^t \Phi_1^*\Big(\alpha,\theta,\frac{\tau}{h}-i\Big)d\Big(\frac{\tau}{h}\Big), & i=1,2,\ldots,L-1, \end{cases}$$

находим $P_{ij}^{-1} = hC_{ij}$, где

$$C_{00} = C_{L-1,L-1} = \frac{1}{4}; \quad C_{11} = \ldots = C_{L-2,L-2} = \frac{1}{2} \quad (L > 2);$$
$$C_{01} = 2\upsilon \quad (L = 2); \quad C_{01} = C_{L-2,L-1} = \sqrt{2}\upsilon \quad (L > 2); \quad C_{12} = \ldots = C_{L-3,L-2} = \upsilon \quad (L > 3);$$
$$C_{i-k,i} = 0, \quad i = 2, 3, \ldots, L-1, \quad k = 2, 3, \ldots, i \quad (L > 2);$$
$$C_{10} = C_{L-1,L-2} = \sqrt{2}\left(\frac{1}{2} - \upsilon\right) \quad (L > 2); \quad C_{21} = \ldots = C_{L-2,L-3} = 1 - \upsilon \quad (L > 3);$$
$$C_{i,i-k} = 1, \quad i = 2, 3, \ldots, L-2, \quad k = 1, 2, \ldots, i-1 \quad (L > 3);$$
$$C_{20} = \ldots = C_{L-2,0} = \frac{\sqrt{2}}{2} \quad (L > 3); \quad C_{10} = 2\left(\frac{1}{4} - \upsilon\right) \quad (L = 2); \quad C_{L-1,0} = \frac{1}{2} \quad (L > 2);$$
$$C_{L-1,1} = \ldots = C_{L-1,L-3} = \frac{\sqrt{2}}{2} \quad (L > 3),$$

где $\upsilon = \frac{17}{192} - \frac{\sqrt{41}}{64}$, т.е.

$$P^{-1} = h \cdot \begin{pmatrix} \frac{1}{4} & 2\upsilon \\ 2\left(\frac{1}{4} - \upsilon\right) & \frac{1}{4} \end{pmatrix} \quad (L = 2); \quad P^{-1} = h \cdot \begin{pmatrix} \frac{1}{4} & \sqrt{2}\upsilon & 0 \\ \sqrt{2}\left(\frac{1}{2} - \upsilon\right) & \frac{1}{2} & \sqrt{2}\upsilon \\ \frac{1}{2} & \sqrt{2}\left(\frac{1}{2} - \upsilon\right) & \frac{1}{4} \end{pmatrix} \quad (L = 3);$$

$$P^{-1} = h \cdot \begin{pmatrix} \frac{1}{4} & \sqrt{2}\upsilon & 0 & 0 & \ldots & 0 & 0 \\ \sqrt{2}\left(\frac{1}{2} - \upsilon\right) & \frac{1}{2} & \upsilon & 0 & \ldots & 0 & 0 \\ \frac{\sqrt{2}}{2} & 1 - \upsilon & \frac{1}{2} & \upsilon & \ddots & 0 & 0 \\ \frac{\sqrt{2}}{2} & 1 & 1 - \upsilon & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \upsilon & 0 \\ \frac{\sqrt{2}}{2} & 1 & 1 & \ddots & 1 - \upsilon & \frac{1}{2} & \sqrt{2}\upsilon \\ \frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \ldots & \frac{\sqrt{2}}{2} & \sqrt{2}\left(\frac{1}{2} - \upsilon\right) & \frac{1}{4} \end{pmatrix} \quad (L > 3).$$

Заключение

В первой части работы рассмотрена система ортогональных финитных функций, порожденная симметричным кусочно-линейным сплайном Леонтьева. Сформированы алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования относительно этой системы. Система ортогональных финитных функций, порожденная несимметричным кусочно-линейным сплайном Леонтьева, рассмотрена во второй части этой работы [6].

Список литературы

[1] Леонтьев В. Л. Ортогональные финитные функции и численные методы. Ульяновск: УлГУ, 2003.

[2] Марчук Г. И., Агошков В. И. Введение в проекционно-сеточные методы. М.: Наука, 1981.

[3] Рыбаков К. А. Построение множества допустимых управлений в спектральной форме математического описания // Вычислительные технологии. 2015. Т. 20. № 3. С. 58–74.

[4] Рыбаков К. А., Рыбин В. В. Моделирование распределенных и дробно-распределенных процессов и систем управления спектральным методом. М.: Изд-во МАИ, 2016.

[5] Рыбаков К. А., Рыбин В. В. Алгоритмическое и программное обеспечение расчета систем автоматического управления в спектральной форме математического описания / В кн. Современная наука: теоретические, практические и инновационные аспекты развития. Т. 2. Ростов-на-Дону: Изд-во Международного исследовательского центра «Научное сотрудничество», 2018. С. 171–199.

[6] Рыбаков К. А., Рыбин В. В. Спектральные характеристики операторов дифференцирования и интегрирования относительно ортогональных финитных функций. II. Несимметричные сплайны // Актуальные проблемы вычислительной и прикладной математики (АПВПМ-2019). Международная конференция, Новосибирск, 1–5 июля 2019 г.: Материалы конф. Новосибирск: Изд-во ИВМиМГ СО РАН, 2019.

[7] Рыбин В. В. Моделирование нестационарных систем управления целого и дробного порядка проекционно-сеточным спектральным методом. М.: Изд-во МАИ, 2013.

[8] Семенов В. В. Формы математического описания линейных систем. М.: МАИ, 1980.

[9] Солодовников В. В., Семенов В. В. Спектральная теория нестационарных систем управления. М.: Наука, 1974.

[10] Таблицы и математическое обеспечение спектрального метода теории автоматического управления / Под ред. В. В. Семенова. М.: МВТУ им. Н. Э. Баумана, 1973.

Рыбаков Константин Александрович — к.ф.-м.н., доцент Московского авиационного института (национального исследовательского университета)
e-mail: rkoffice@mail.ru;
Рыбин Владимир Васильевич — к.т.н., доцент Московского авиационного института (национального исследовательского университета)
e-mail: vv-ribin@mail.ru.
Дата поступления — 22 апреля 2019 г.

СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ ОПЕРАТОРОВ ДИФФЕРЕНЦИРОВАНИЯ И ИНТЕГРИРОВАНИЯ ОТНОСИТЕЛЬНО ОРТОГОНАЛЬНЫХ ФИНИТНЫХ ФУНКЦИЙ. II. НЕСИММЕТРИЧНЫЕ СПЛАЙНЫ

К. А. Рыбаков, В. В. Рыбин

Московский авиационный институт (национальный исследовательский университет), 125993, Москва

УДК 519.711.3
DOI: 10.24411/9999-016A-2019-10069

Во второй части работы сформированы алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования относительно системы ортогональных финитных функций, которая порождается с помощью масштабирования и сдвига несимметричного кусочно-линейного сплайна Леонтьева.

Ключевые слова: ортогональные финитные функции, несимметричный сплайн Леонтьева, спектральная форма математического описания, спектральный метод, финитная базисная система.

Введение

В первой части представленной работы сформированы алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования, или согласно терминологии, принятой в теории управления, двумерных нестационарных передаточных функций дифференцирующего и интегрирующего звеньев [7]. Эти характеристики были определены относительно системы ортогональных финитных функций, которая порождается с помощью масштабирования и сдвига симметричного кусочно-линейного сплайна. Во второй части рассмотрена система ортогональных финитных функций, построенная на основе несимметричного кусочно-линейного сплайна.

Отметим, что базисные системы функций, которые формируются с помощью масштабирования и сдвига некоторой порождающей финитной функции, часто применяются в вычислительной математике для аппроксимации решений уравнений разных типов. Однако во многих случаях такие функции не образуют ортогональную систему (относительно скалярного произведения пространства квадратично интегрируемых функций). Матрица Грама для такой базисной системы имеет ленточную структуру: трехдиагональную, пятидиагональную и т.д. за счет того, что носители функций базисной системы, номера которых соответственно отличаются на 1, 2 и т.д. имеют непустое пересечение. Такая ситуация возникает при выборе порождающей финитной функции в виде сплайнов Шёнберга, сплайнов Эрмита [2] и др. Однако, как показано в [1], можно определить порождающую финитную функцию таким образом, что результирующая базисная система будет обладать свойством ортогональности. В этой работе именно такие функции выбраны в качестве базисной системы: симметричные и несимметричные кусочно-линейные сплайны Леонтьева.

Сформированные алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования дополняют библиотеку элементарных алгоритмов спектрального метода анализа и синтеза систем автоматического управления [5,6,9,10]. Они могут применяться для различных задач, таких как моделирование систем управления, анализ выходных процессов, синтез управления динамическими системами, идентификация и оценивание параметров [3,8,9].

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 17-08-00530).

ISBN 978-5-901548-42-4

Первая часть работы содержит необходимые определения базисных систем [7], поэтому в этой части изложение начинается сразу с определения используемого далее несимметричного кусочно-линейного сплайна.

1 Несимметричные кусочно-линейные финитные функции

Рассмотрим кусочно-линейную функцию

$$\Psi_1^*(\alpha,\theta,t)=\begin{cases}1+t,\ t\in[-1,-1+\theta)\cup[-\theta,0),\\ -\alpha+\frac{\theta+\alpha}{\theta-\frac12}\big(t+\frac12\big),\ t\in\big[-1+\theta,-\frac12\big),\\ -\alpha+\frac{\theta-\alpha-1}{\theta-\frac12}\big(t+\frac12\big),\ t\in\big[-\frac12,-\theta\big),\\ 1-t,\ t\in[0,\theta)\cup[1-\theta,1),\\ 1+\alpha-\frac{\theta+\alpha}{\theta-\frac12}\big(t-\frac12\big),\ t\in\big[\theta,\frac12\big),\\ 1+\alpha-\frac{\theta-\alpha-1}{\theta-\frac12}\big(t-\frac12\big),\ t\in\big[\frac12,1-\theta\big),\\ 0,\ t\notin[-1,1],\end{cases}$$

где α и θ — параметры: $\alpha>0$, $0<\theta<\frac12$.

Несимметричный сплайн степени p можно выразить через функции $\Psi_1^*(\alpha,\theta,t)$ и $B_0^*(t)$ (B-сплайн нулевой степени [7]), используя свертку [1]:

$$\Psi_p^*(\alpha,\theta,t)=\Psi_{p-1}^*(\alpha,\theta,t)*B_0^*(t)=\int_{-\infty}^{+\infty}\Psi_{p-1}^*(\alpha,\theta,\tau)B_0^*(t-\tau)d\tau,\quad p=2,3,4,\ldots$$

Как и в первой части работы [7], ограничимся случаем $p=1$. Кусочно-линейный сплайн $\Psi_1^*(\alpha,\theta,t)$ порождает финитную базисную систему на отрезке $\mathrm{T}=[0,T]$ с помощью масштабирования и сдвига согласно правилу

$$\Psi_1(i,t)=\frac{1}{\sqrt{h}}\,\Psi_1^*\bigg(\alpha,\theta,\frac{t}{h}-i\bigg),\quad h=\frac{T}{L-1},\quad t\in\mathrm{T},\quad i=0,1,\ldots,L-1.$$

Матрица Грама для системы функций $\{\Psi_1(i,t)\}_{i=0}^{L-1}$ в общем случае является трехдиагональной так же, как и для системы функций $\{\Phi_1(i,t)\}_{i=0}^{L-1}$ из [7], однако при условии

$$\theta=\frac14\quad\text{и}\quad\alpha=\frac12 \tag{1}$$

скалярные произведения функций $\{\Psi_1(i,t)\}_{i=0}^{L-1}$ с номерами, отличающимися на единицу, равны нулю и $\{\Psi_1(i,t)\}$ — ортогональная система функций.

Вывод соотношений для расчета спектральных характеристик операторов дифференцирования и интегрирования здесь аналогичен случаю симметричных ортогональных финитных функций, рассмотренных в первой части работы [7]. Поэтому часть комментариев будет для краткости опущена.

Для функции $\Psi_1^*(\alpha,\theta,t)$ справедливы следующие свойства:

$$\int_{-1}^{1}\big(\Psi_1^*(\alpha,\theta,t)\big)^2dt=1,\quad\int_{0}^{1}\big(\Psi_1^*(\alpha,\theta,t)\big)^2dt=\frac34,\quad\int_{-1}^{0}\big(\Psi_1^*(\alpha,\theta,t)\big)^2dt=\frac14,$$

т.е. условие нормировки выполнено для функций $\{\Psi_1(i,t)\}_{i=0}^{L-1}$ с номерами $i=1,2,\ldots,L-2$, а для функций с номерами $i=0$ и $i=L-1$ нужны дополнительные нормировочные коэффициенты $2/\sqrt{3}$ и 2 соответственно. Таким образом, функции

$$\hat\Psi_1(i,t)=\frac{\nu_i}{\sqrt{h}}\,\Psi_1^*\bigg(\alpha,\theta,\frac{t}{h}-i\bigg)=\nu_i\Psi_1(i,t),\quad i=0,1,\ldots,L-1,$$

где $\nu_0=2/\sqrt{3}$, $\nu_1=\ldots=\nu_{L-2}=1$ и $\nu_{L-1}=2$, образуют ортонормированную систему на отрезке $\mathrm{T}=[0,T]$.

2 Спектральные характеристики операторов дифференцирования и интегрирования

Найдем производную функции $\Psi_1^*(\alpha,\theta,t)$ на интервале $(-1,1)$, из которого исключены $0, \pm\theta, \pm 1/2, \pm(1-\theta)$:

$$\Psi_1^{*\prime}(\alpha,\theta,t) = \begin{cases} 1,\ t \in (-1,-1+\theta) \cup (-\theta,0), \\ \frac{\theta+\alpha}{\theta-\frac{1}{2}},\ t \in \left(-1+\theta,-\frac{1}{2}\right), \\ \frac{\theta-\alpha-1}{\theta-\frac{1}{2}},\ t \in \left(-\frac{1}{2},-\theta\right), \\ -1,\ t \in (0,\theta) \cup (1-\theta,1), \\ -\frac{\theta+\alpha}{\theta-\frac{1}{2}},\ t \in \left(\theta,\frac{1}{2}\right), \\ -\frac{\theta-\alpha-1}{\theta-\frac{1}{2}},\ t \in \left(\frac{1}{2},1-\theta\right), \\ 0,\ t \notin (-1,1). \end{cases}$$

Соотношения, необходимые для расчета спектральной характеристики оператора дифференцирования (здесь, как и ранее в [7], интегрирование по отрезку $[0,1]$ используется для учета функции финитной базисной системы с номером $i=0$, по отрезку $[-1,1]$ — с номерами $i=1,2,\ldots,L-2$ и по отрезку $[-1,0]$ — с номером $i=L-1$), имеют вид

— для элементов главной диагонали:

$$\int_{-1}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t)dt = 0, \quad \int_{0}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t)dt = -\int_{-1}^{0} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t)dt = -\frac{1}{2};$$

— для элементов наддиагонали и поддиагонали:

$$\int_{-1}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t-1)dt = \int_{0}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t-1)dt = \frac{1}{2},$$
$$\int_{-1}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t+1)dt = \int_{-1}^{0} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t+1)dt = -\frac{1}{2};$$

— для остальных элементов ($k>1$):

$$\int_{-1}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t-k)dt = \int_{0}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t-k)dt = 0,$$
$$\int_{-1}^{1} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t+k)dt = \int_{-1}^{0} \Psi_1^*(\alpha,\theta,t)\Psi_1^{*\prime}(\alpha,\theta,t+k)dt = 0,$$

поскольку функции $\Psi_1^*(\alpha,\theta,t)$ и $\Psi_1^{*\prime}(\alpha,\theta,t\pm k)$ имеют непересекающиеся носители.

Так так

$$\hat{\Psi}_1'(i,t) = \frac{\gamma_i}{h\sqrt{h}}\,\Psi_1^{*\prime}\left(\alpha,\theta,\frac{t}{h}-i\right), \quad i=0,1,\ldots,L-1,$$

получаем $\mathcal{P}_{ij} = (1/h)\,C_{ij}$, где

$$C_{00} = -\frac{2}{3}, \quad C_{L-1,L-1} = 2; \quad C_{11} = \ldots = C_{L-2,L-2} = 0 \quad (L>2);$$
$$C_{01} = \frac{2\sqrt{3}}{3} \quad (L=2); \quad C_{01} = \frac{\sqrt{3}}{3}, \quad C_{L-2,L-1} = 1 \quad (L>2); \quad C_{12} = \ldots = C_{L-3,L-2} = \frac{1}{2} \quad (L>3);$$
$$C_{10} = -\frac{2\sqrt{3}}{3} \quad (L=2); \quad C_{21} = -\frac{\sqrt{3}}{3}, \quad C_{L-1,L-2} = -1 \quad (L>2); \quad C_{21} = \ldots = C_{L-2,L-3} = -\frac{1}{2} \quad (L>3);$$
$$C_{i-k,i} = C_{i,i-k} = 0, \quad i=2,3,\ldots,L-1, \quad j=2,3,\ldots,i \quad (L>2),$$

следовательно,

$$\mathcal{P} = \frac{1}{h} \cdot \begin{pmatrix} -\frac{2}{3} & \frac{2\sqrt{3}}{3} \\ -\frac{2\sqrt{3}}{3} & 2 \end{pmatrix} \quad (L=2); \quad \mathcal{P} = \frac{1}{h} \cdot \begin{pmatrix} -\frac{2}{3} & \frac{\sqrt{3}}{3} & 0 \\ -\frac{\sqrt{3}}{3} & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} \quad (L=3);$$

$$\mathcal{P} = \frac{1}{h} \cdot \begin{pmatrix} -\frac{2}{3} & \frac{\sqrt{3}}{3} & 0 & 0 & \dots & 0 & 0 \\ -\frac{\sqrt{3}}{3} & 0 & \frac{1}{2} & 0 & \dots & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \ddots & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{pmatrix} \quad (L>3).$$

Кроме того, все функции финитной базисной системы $\{\hat{\Psi}_1(i,t)\}_{i=0}^{L-1}$ принимают нулевое значение при $t=0$ за исключением $\hat{\Psi}_1(0,t)$. Ее значение выражается через значение $\Psi_1^*(\alpha,\theta,0)=1$ и нормировочный коэффициент $2/\sqrt{3}$, поэтому

$$P = \frac{1}{h} \cdot \begin{pmatrix} \frac{2}{3} & \frac{2\sqrt{3}}{3} \\ -\frac{2\sqrt{3}}{3} & 2 \end{pmatrix} \quad (L=2); \quad P = \frac{1}{h} \cdot \begin{pmatrix} \frac{2}{3} & \frac{\sqrt{3}}{3} & 0 \\ -\frac{\sqrt{3}}{3} & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} \quad (L=3);$$

$$P = \frac{1}{h} \cdot \begin{pmatrix} \frac{2}{3} & \frac{\sqrt{3}}{3} & 0 & 0 & \dots & 0 & 0 \\ -\frac{\sqrt{3}}{3} & 0 & \frac{1}{2} & 0 & \dots & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \ddots & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{pmatrix} \quad (L>3).$$

Запишем две первообразных функции $\Psi_1^*(\alpha,\theta,t)$:

$$\int_0^t \Psi_1^*(\alpha,\theta,\tau)d\tau = \begin{cases} -\frac{(1-t)^2}{2} + \frac{(2\alpha+1)(2\theta-1)+3}{4}, \ t \in [0,\theta), \\ -\frac{(\alpha+\theta)(t-\theta)^2}{2\theta-1} - (\theta-1)(t-\theta) + \frac{(2\alpha-\theta+3)(2\theta-1)-\theta+3}{4}, \ t \in \left[\theta, \frac{1}{2}\right), \\ \frac{(\alpha-\theta+1)(2t-1)^2}{4(2\theta-1)} + \frac{(\alpha+1)(2t-1)}{2} + \frac{\alpha(2\theta-1)+\theta}{4} + \frac{1}{2}, \ t \in \left[\frac{1}{2}, 1-\theta\right), \\ \frac{3}{4} - \frac{(1-t)^2}{2}, \ t \in [1-\theta, 1), \\ \frac{3}{4}, \ t \geqslant 1, \end{cases} \qquad t \geqslant 0;$$

$$\int_{-1}^{t}\Psi_1^*(\alpha,\theta,\tau)d\tau=\begin{cases}\frac{(1+t)^2}{2},\ t\in[-1,-1+\theta),\\ \frac{(1+t)^2}{2}+\frac{(2\alpha+1)(t-\theta+1)^2}{2(2\theta-1)},\ t\in\left[-1+\theta,-\frac12\right),\\ \frac{(1+t)^2}{2}+\frac{(2\alpha+1)(t-\theta+1)^2}{2(2\theta-1)}-\frac{(2\alpha+1)(2t+1)^2}{4(2\theta-1)},\ t\in\left[-\frac12,-\theta\right),\\ \frac{(1+t)^2}{2}+\frac{(2\alpha+1)(2\theta-1)}{4},\ t\in[-\theta,0),\\ -\frac{(1-t)^2}{2}+\frac{(2\alpha+1)(2\theta-1)}{4}+1,\ t\in[0,\theta),\\ -\frac{(\alpha+\theta)(t-\theta)^2}{2\theta-1}-(\theta-1)(t-\theta)+\frac{(2\alpha-\theta+3)(2\theta-1)-\theta}{4}+1,\ t\in\left[\theta,\frac12\right),\\ \frac{(\alpha-\theta+1)(2t-1)^2}{4(2\theta-1)}+\frac{(\alpha+1)(2t-1)}{2}+\frac{\alpha(2\theta-1)+\theta+3}{4},\ t\in\left[\frac12,1-\theta\right),\\ 1-\frac{(1-t)^2}{2},\ t\in[1-\theta,1),\\ 1,\ t\geqslant1,\end{cases}\qquad t\geqslant-1.$$

Следовательно, с учетом (1) получаем соотношения, необходимые для расчета спектральной характеристики оператора интегрирования, а именно

— для элементов главной диагонали:

$$\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac12,$$
$$\int_{0}^{1}\Psi_1^*(\alpha,\theta,t)\int_{0}^{t}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac{9}{32},\quad\int_{-1}^{0}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac{1}{32};$$

— для элементов наддиагонали и поддиагонали за исключением первого столбца и последней строки:

$$\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t-1}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\int_{0}^{1}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t-1}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac{1}{96},$$
$$\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t+1}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=1-\frac{1}{96}=\frac{95}{96};$$

— для остальных элементов за исключением первого столбца и последней строки (здесь и далее $k>1$):

$$\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t-k}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=0,\quad\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t+k}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=1;$$

— для оставшихся элементов первого столбца:

$$\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{0}^{t+1}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac34-\frac{1}{96}=\frac{71}{96},$$
$$\int_{-1}^{0}\Psi_1^*(\alpha,\theta,t)\int_{0}^{t+1}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac{3}{16}-\frac{1}{96}=\frac{17}{96},$$
$$\int_{-1}^{1}\Psi_1^*(\alpha,\theta,t)\int_{0}^{t+k}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac34,\quad\int_{-1}^{0}\Psi_1^*(\alpha,\theta,t)\int_{0}^{t+k}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac{3}{16};$$

— для оставшихся элементов последней строки:

$$\int_{-1}^{0}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t+k}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac14,\quad\int_{-1}^{0}\Psi_1^*(\alpha,\theta,t)\int_{-1}^{t+1}\Psi_1^*(\alpha,\theta,\tau)d\tau dt=\frac14-\frac{1}{96}=\frac{23}{96}.$$

Так так

$$\int_{0}^{t}\hat\Psi_1(i,\tau)d\tau=\begin{cases}\dfrac{\gamma_0h}{\sqrt h}\displaystyle\int_0^t\Psi_1^*\left(\alpha,\theta,\frac{\tau}{h}\right)d\left(\frac{\tau}{h}\right),\quad i=0,\\ \dfrac{\gamma_ih}{\sqrt h}\displaystyle\int_{-1}^t\Psi_1^*\left(\alpha,\theta,\frac{\tau}{h}-i\right)d\left(\frac{\tau}{h}\right),\quad i=1,2,\ldots,L-1,\end{cases}$$

имеем $P_{ij}^{-1} = hC_{ij}$, где

$$C_{00} = \frac{3}{8}; \quad C_{11} = \ldots = C_{L-2,L-2} = \frac{1}{2} \quad (L > 2); \quad C_{L-1,L-1} = \frac{1}{8};$$

$$C_{01} = \frac{4}{\sqrt{3}}\upsilon \quad (L = 2); \quad C_{01} = \frac{2}{\sqrt{3}}\upsilon \quad (L > 2); \quad C_{12} = \ldots = C_{L-3,L-2} = \upsilon \quad (L > 3);$$

$$C_{L-2,L-1} = 2\upsilon \quad (L > 2); \quad C_{i-k,i} = 0, \quad i = 2,3,\ldots,L-1, \quad k = 2,3,\ldots,i \quad (L > 2);$$

$$C_{10} = \frac{2}{\sqrt{3}}\left(\frac{3}{4} - \upsilon\right), \quad C_{L-1,L-2} = 2\left(\frac{1}{4} - \upsilon\right) \quad (L > 2); \quad C_{21} = \ldots = C_{L-2,L-3} = 1 - \upsilon \quad (L > 3);$$

$$C_{i,i-k} = 1, \quad i = 2,3,\ldots,L-2, \quad k = 1,2,\ldots,i-1 \quad (L > 3);$$

$$C_{20} = \ldots = C_{L-2,0} = \frac{\sqrt{3}}{2} \quad (L > 3); \quad C_{10} = \frac{4\sqrt{3}}{4}\left(\frac{3}{16} - \upsilon\right) \quad (L = 2); \quad C_{L-1,0} = \frac{\sqrt{3}}{4} \quad (L > 2);$$

$$C_{L-1,1} = \ldots = C_{L-1,L-3} = \frac{1}{2} \quad (L > 3),$$

где $\upsilon = \frac{1}{96}$, т.е.

$$P^{-1} = h \cdot \begin{pmatrix} \frac{3}{8} & \frac{\sqrt{3}}{72} \\ \frac{17\sqrt{3}}{72} & \frac{1}{8} \end{pmatrix} \quad (L = 2); \quad P^{-1} = h \cdot \begin{pmatrix} \frac{3}{8} & \frac{\sqrt{3}}{144} & 0 \\ \frac{71\sqrt{3}}{72} & \frac{1}{2} & \frac{1}{48} \\ \frac{\sqrt{3}}{4} & \frac{23}{48} & \frac{1}{8} \end{pmatrix} \quad (L = 3);$$

$$P^{-1} = h \cdot \begin{pmatrix} \frac{3}{8} & \frac{\sqrt{3}}{144} & 0 & 0 & \ldots & 0 & 0 \\ \frac{71\sqrt{3}}{72} & \frac{1}{2} & \frac{1}{96} & 0 & \ldots & 0 & 0 \\ \frac{\sqrt{3}}{2} & \frac{95}{96} & \frac{1}{2} & \frac{1}{96} & \ddots & 0 & 0 \\ \frac{\sqrt{3}}{2} & 1 & \frac{95}{96} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{1}{96} & 0 \\ \frac{\sqrt{3}}{2} & 1 & 1 & \ddots & \frac{95}{96} & \frac{1}{2} & \frac{1}{48} \\ \frac{\sqrt{3}}{4} & \frac{1}{2} & \frac{1}{2} & \ldots & \frac{\sqrt{2}}{2} & \frac{23}{48} & \frac{1}{8} \end{pmatrix} \quad (L > 3).$$

Заключение

В представленной работе, состоящей из двух частей, рассмотрены две системы ортогональных финитных функций, порожденные симметричным и несимметричным кусочно-линейными сплайнами Леонтьева. Сформированы алгоритмы расчета спектральных характеристик операторов дифференцирования и интегрирования относительно этих систем. Полученные результаты включены в библиотеку элементарных алгоритмов спектрального метода анализа и синтеза систем автоматического управления [6]. Они могут применяться для различных задач, таких как моделирование систем управления, анализ выходных процессов, синтез управления динамическими системами, идентификация и оценивание параметров [3, 8, 9]. Представляет интерес моделирование случайных процессов с непрерывными траекториями на основе спектральной формы математического описания и спектрального метода [4] с помощью рассмотренных ортогональных финитных функций.

Список литературы

[1] Леонтьев В. Л. Ортогональные финитные функции и численные методы. Ульяновск: УлГУ, 2003.

[2] Марчук Г. И., Агошков В. И. Введение в проекционно-сеточные методы. М.: Наука, 1981.

[3] Рыбаков К. А. Спектральные аналоги множества допустимых управлений для финитных базисных систем // Дифференциальные уравнения и процессы управления. 2016. № 2. С. 40–71.

[4] Рыбаков К. А. Применение спектрального метода для моделирования, анализа и оптимизации линейных стохастических систем управления // Мехатроника, автоматика и робототехника. III Международная научно-практическая конференция, Новокузнецк, 22 февраля 2019 г.: Материалы конф. Новокузнецк: НИЦ МС, 2019. С. 173–175.

[5] Рыбаков К. А., Рыбин В. В. Алгоритмическое и программное обеспечение расчета спектральной характеристики оператора дробного интегродифференцирования относительно функций Хаара // Машиностроение и компьютерные технологии. 2018. № 1. С. 31–51.

[6] Рыбаков К. А., Рыбин В. В. Алгоритмическое и программное обеспечение расчета систем автоматического управления в спектральной форме математического описания / В кн. Современная наука: теоретические, практические и инновационные аспекты развития. Т. 2. Ростов-на-Дону: Изд-во Международного исследовательского центра «Научное сотрудничество», 2018. С. 171–199.

[7] Рыбаков К. А., Рыбин В. В. Спектральные характеристики операторов дифференцирования и интегрирования относительно ортогональных финитных функций. I. Симметричные сплайны // Актуальные проблемы вычислительной и прикладной математики (АПВПМ-2019). Международная конференция, Новосибирск, 1–5 июля 2019 г.: Материалы конф. Новосибирск: Изд-во ИВМиМГ СО РАН, 2019.

[8] Рыбин В. В. Моделирование нестационарных непрерывно-дискретных систем управления спектральным методом в системах компьютерной математики. М.: Изд-во МАИ, 2011.

[9] Солодовников В. В., Семенов В. В. Спектральная теория нестационарных систем управления. М.: Наука, 1974.

[10] Таблицы и математическое обеспечение спектрального метода теории автоматического управления / Под ред. В. В. Семенова. М.: МВТУ им. Н. Э. Баумана, 1973.

Рыбаков Константин Александрович — к.ф.-м.н., доцент Московского авиационного института (национального исследовательского университета)
e-mail: rkoffice@mail.ru;
Рыбин Владимир Васильевич — к.т.н., доцент Московского авиационного института (национального исследовательского университета)
e-mail: vv-ribin@mail.ru.
Дата поступления — 22 апреля 2019 г.

АНАЛИТИЧЕСКАЯ И ЧИСЛЕННАЯ РЕГУЛЯРИЗАЦИЯ РЕШЕНИЯ ЗАДАЧИ ИНТЕГРАЛЬНОЙ ГЕОМЕТРИИ НА СЕМЕЙСТВАХ ПАРАБОЛ И ЛОМАНЫХ

А. К. Сеидуллаев[1], Г. М. Джайков[2]

[1] *Каракалпакский государственный университет, 230112, Нукус, Узбекистан*

[2] *Нукусский филиал Ташкентского университета информационных технологий, 230102, Нукус, Узбекистан*

УДК 517.946
DOI: 10.24411/9999-016A-2019-10070

В этой работе изучается задача восстановления функции в полосе по известным интегралам от нее с заданной весовой функцией на семействе ломаных и параболах. Для двух классов семействах получены явные формулы обращения, на их основе доказаны теоремы единственности и существования решения. Получены оценки устойчивости решения задач в пространствах Соболева, откуда вытекает слабая некорректность задач. Для задач интегральной геометрии на семействе парабола получены регуляризованное решение. На основе этих результатах представлены результаты численных экспериментов.

Ключевые слова: некорректные задачи, задачи интегральной геометрии, интегральные преобразования, формула обращения, единственность решения, теорема существования, регуляризация.

Введение

В работе рассматривается задача восстановления функции, если известны интегралы от нее по ломаным с заданной весовой функцией и на семействе парабол. Задачи такого типа называются задачами интегральной геометрии [1]. Такие задачи на линейных многообразиях и других явно заданных кривых и поверхностях имеют многочисленные приложения в компьютерной, сейсмической и ультразвуковой томографии, задачах восстановления изображения [1–3].

В работах [4, 5] рассматривались новые постановки слабо некорректных задач интегральной геометрии на параболических кривых со специальными весовыми функциями. В частности, в работе [4] приведены теоремы единственности и оценка устойчивости, из которого вытекает утверждение о сильной некорректности решения задачи.

В работе [6] получено аналитическое представление для образа Фурье по первой переменной от искомой функции, из которого вытекает утверждение о сильной некорректности решения задачи. Важные результаты по обращению преобразования Радона и приложениям в сейсмической и компьютерной томографии представлены в [7–9].

В первом параграфе мы рассматриваем задачу интегральной геометрии на семействе ломаных с кусочно-постоянной весовой функцией. Получена явная формула обращения, на основе которой доказана теорема единственности и получены оценки устойчивости ее решения. Из этих оценок вытекает слабая некорректность задачи.

Во втором параграфе рассмотрена задача интегральной геометрии на семейство парабол. В работе исходя из идеи А. Н. Тихонова о регуляризации некорректных задач, построена последовательность решений, приближенных к точному решению задачи.

Введем обозначения, которые будем использовать далее:

$$(x,\, y) \in R^2, \quad (\xi,\, \eta) \in R^2, \quad \lambda \in R^1, \quad \mu \in R^1,$$

$$L_H = \left\{(x,\, y):\ \ x \in R^1,\ y \in [0,\, H],\ \ H < \infty\right\}.$$

ISBN 978-5-901548-42-4

Задача 1. Восстановить функцию двух переменных $u(x,y)$, если в полосе L_H известны интегралы от нее по кривым семейства $\Gamma(x,y)=\{(\xi,\eta): y-\eta=|x-\xi|,\ 0\le y\le H\}$ с весовой функции $g(x,\xi)=sgn(x-\xi)$:

$$\int\limits_{\Gamma(x,y)} g(x,\xi)\,u(x,y)\,d\xi = f(x,y). \tag{1}$$

Задача 2. В полосе L_H рассмотрим функцию двух переменных $u(x,y)$, если известны интегралы от нее по кривым семейства $\Psi(x,y)=\left\{y-\eta=(x-\xi)^2,\ 0<\eta<y\right\}$ с весовой функции $g(x,\xi)=1$:

$$\int\limits_{\Psi(x,y)} g(x,\xi)\,u(x,y)\,d\xi = f(x,y). \tag{2}$$

1 Задача интегральной геометрии с кусочно-постоянной весовой функцией

Теорема 1. Пусть функция $f(x,y)$ известна для всех $(x,y)\in L_H$, весовая функция имеет вид $g(x,\xi)=sign(x-\xi)$. Тогда решение задачи в классе дважды непрерывно дифференцируемых финитных с носителем в полосе L_H функций единственно, выражается через функцию $f(x,y)$ по формуле

$$\frac{\partial}{\partial x}u(x,y)=\frac{1}{2}\left(\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}\right)f(x,y) \tag{3}$$

и удовлетворяет неравенства

$$\|u(x,y)\|_{W_2^{1,0}}\le C_1\|f(x,y)\|_{W_2^{2,2}},$$

здесь C_1 — некоторая константа.

Доказательство. Уравнение (1) можно представить в следующем виде

$$\int\limits_0^y [u(x-h,\eta)-u(x+h,\eta)]\,d\eta = f(x,y), \tag{4}$$

где $h=y-\eta$. Применим к этому уравнению преобразование Фурье по первой переменной:

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{i\lambda x}f(x,y)\,dx=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{i\lambda x}\int\limits_0^y [u(x-h,\eta)-u(x+h,\eta)]\,d\eta dx.$$

Получим уравнение:

$$\int\limits_0^y \hat{u}(\lambda,\eta)\sin(\lambda h)\,d\eta=\phi(\lambda,y), \tag{5}$$

где $\phi(\lambda,y)=-\frac{i}{2\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{i\lambda x}f(x,y)\,dx$ — преобразование Фурье по переменной x от функций $f(x,y)$. Применим к уравнению (5) преобразование Лапласа по переменной y:

$$\int\limits_0^{\infty} e^{-py}\phi(\lambda,y)\,dy=\int\limits_0^{\infty} e^{-py}\int\limits_0^y \hat{u}(\lambda,\eta)\sin(\lambda(y-\eta))\,d\eta dy.$$

Сделав замену $t=y-\eta$, имеем

$$\int\limits_0^{\infty} e^{-p\eta}\hat{u}(\lambda,\eta)\,d\eta\int\limits_0^{\infty} e^{-pt}\sin(\lambda t)\,dt=\int\limits_0^{\infty} e^{-py}\phi(\lambda,y)\,dy.$$

Таким образом, из уравнения (4) получаем:

$$\tilde{\hat{u}}(\lambda,p)\cdot J(\lambda,p)=\tilde{\phi}(\lambda,p), \tag{6}$$

где $\phi(\lambda,p)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}e^{i\mu y}\hat{f}(\lambda,y)\,dy$,

$$J(\lambda,p)=\int\limits_{0}^{\infty}e^{-pt}\sin(\lambda t)\,dt. \tag{7}$$

Используя формулы Эйлера, уравнение (7) приведем к виду

$$J(\lambda,p)=\frac{\lambda}{p^2+\lambda^2}. \tag{8}$$

Из уравнения (6), учитывая (8), получим

$$\lambda\tilde{\hat{u}}(\lambda,p)=\left(p^2+\lambda^2\right)\tilde{\phi}(\lambda,p). \tag{9}$$

Применим к уравнению (9) обратное преобразование Лапласа по p и обратное преобразование Фурье по λ. Учитывая свойства преобразования Лапласа и Фурье, приходим к формуле обращения

$$\frac{\partial}{\partial x}u(x,y)=\frac{1}{2}\left(\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}\right)f(x,y). \tag{10}$$

Из последнего равенства (10), получаем

$$\left\|\frac{\partial}{\partial x}u(x,y)\right\|_{L_2(L_H)}=\left\|\frac{1}{2}\left(\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}\right)f(x,y)\right\|_{L_2(L_H)}. \tag{11}$$

Используя свойства преобразования Лапласа и Фурье, а также учитывая (11) получим оценку

$$\|u(x,y)\|_{W_2^{1,0}}\le C_1\|f(x,y)\|_{W_2^{2,2}},$$

где C_1 — некоторая константа.

2 Аналитическая регуляризация задачи интегральной геометрии на семейство парабол

Теорема 2. Пусть функция $f(x,y)$ известна в полосе L_H, $g(x,\xi)=1$. Тогда решение задачи 2 в C_0^{∞} единственно и имеет место представление

$$u_\alpha(x,y)=\frac{1}{2\alpha\sqrt{\pi^3}}\frac{\partial}{\partial y}\int\limits_0^y\int\limits_{-\infty}^{\infty}e^{\frac{y-\eta-(x-\xi)^2}{4\alpha^2}}\cos\frac{(x-\xi)\sqrt{y-\eta}}{2\alpha^2}\frac{f(\xi,\eta)}{\sqrt{y-\eta}}d\xi d\eta,$$

где $\alpha>0$.

Доказательство. Уравнение (2) запишем в виде

$$\int\limits_0^y\left[u(x-h,\eta)+u(x+h,\eta)\right]\frac{d\eta}{2\sqrt{y-\eta}}=f(x,y). \tag{12}$$

Теперь применим к обеим частям уравнения (12) преобразование Фурье по переменной x:

$$\int\limits_0^y\hat{u}(\lambda,\eta)\frac{\cos(\lambda\sqrt{y-\eta})}{\sqrt{y-\eta}}d\eta=\hat{f}(\lambda,y). \tag{13}$$

Подействуем на (13) вольтеровским оператором с ядром

$$\frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}},$$

и применим теорему Фубини к левой части получившегося уравнения:

$$\int_0^y \frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}}\hat{f}\left(\lambda,\eta\right)d\eta = \int_0^y\int_0^\eta \frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}}\hat{u}\left(\lambda,t\right)\frac{\cos\left(\lambda\sqrt{\eta-t}\right)}{\sqrt{\eta-t}}dtd\eta$$

$$= \int_0^y\left(\int_\eta^y \frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}\sqrt{\eta-t}}\cdot\cos\left(\lambda\sqrt{\eta-t}\right)d\eta\right)\hat{u}\left(\lambda,t\right)dt$$

Получим уравнение

$$\int_0^y \frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}}\hat{f}\left(\lambda,\eta\right)d\eta = \int_0^y K_1\left(\lambda,y,t\right)\hat{u}\left(\lambda,t\right)dt,$$

где

$$K_1\left(\lambda,y,t\right) = 2\int_0^1 \frac{ch\left(\lambda\sqrt{y-t}\sqrt{1-s^2}\right)}{\sqrt{1-s^2}}\cos\left(\lambda s\sqrt{y-t}\right)ds.$$

Сделав замену $s=\sqrt{\frac{\eta-t}{y-t}}$ и учитывая выражение для табличного интеграла получаем

$$\pi\int_0^y \hat{u}\left(\lambda,\eta\right)d\eta = \int_0^y \frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}}\hat{f}\left(\lambda,\eta\right)d\eta.$$

Продифференцировав обе части последнего уравнения, приходим к выражению:

$$\hat{u}\left(\lambda,y\right) = \frac{1}{\pi}\frac{\partial}{\partial y}\int_0^y \frac{ch\left(\lambda\sqrt{y-\eta}\right)}{\sqrt{y-\eta}}\hat{f}\left(\lambda,\eta\right)d\eta. \tag{14}$$

Используя метод регуляризации некорректных задач А.Н.Тихонова [10, стр.175], введем последовательность приближенных решений к точному решению задачи (2) в следующим образом

$$\hat{u}_\alpha\left(\lambda,y\right) = e^{-\alpha^2\lambda^2}\hat{u}\left(\lambda,y\right), \tag{15}$$

где $\alpha>0$ — параметр регуляризации. Вычислим обратное преобразование Фурье от функции $\hat{u}_\alpha(\lambda,y)$

$$F^{-1}[\hat{u}_\alpha(\lambda,y)] = u_\alpha(x,y), \tag{16}$$

это есть обычная функция, выражаемая в виде свертки

$$u_\alpha\left(x,y\right) = \int_{-\infty}^{\infty} W_\alpha\left(x-\xi\right)u\left(\xi,y\right)d\xi. \tag{17}$$

Тогда из (14), (15), (16) и (17) получим

$$u_\alpha\left(x,y\right) = \frac{1}{2\alpha\sqrt{\pi^3}}\frac{\partial}{\partial y}\int_0^y\int_{-\infty}^{\infty} e^{\frac{y-\eta-(x-\xi)^2}{4\alpha^2}}\cos\frac{\left(x-\xi\right)\sqrt{y-\eta}}{2\alpha^2}\frac{f\left(\xi,\eta\right)}{\sqrt{y-\eta}}d\xi d\eta. \tag{18}$$

3 Численное моделирование

Пример 1. Введем равномерную сетку в прямоугольной области $D = [-1;\, 1] \times [0;\, 2]$. Отыскиваем приближенные решения задачи на этом прямоугольнике. Функции $f(x, y)$ получаем из формулы (19).
Схема алгоритма решения задачи такова:
Шаг 1. Разобьем отрезок $[-1, 1]$ на оси Ox и $[0, 2]$ на оси Oy на $n_x - 1$ и $n_y - 1$ частей, соответственно. $x_i = -1 + (i-1)\, h_x,\ \ y_j = (j-1)\, h_y$, обозначим $N = n_x = n_y$.
Шаг 2. Приближения функций $u(x_i,\, y_j)$ будем обозначать u^A_{ij};

$$u^A(x_i, y_j) = \int_{-1}^{x_i} F(\xi, y_j)\, d\xi \tag{19}$$

где $F_{ij} = \frac{f_{i+1j} - 2f_{ij} + f_{i-1j}}{2\sqrt{2}{h_x}^2} - \frac{f_{ij+1} - 2f_{ij} + f_{ij-1}}{2\sqrt{2}{h_y}^2}$.
Для реализации описанного выше алгоритма восстановления был написан пакет программ на языке программирования C++. Тестирование алгоритмов происходило на широко применяемой модели – "фантом Sheep-Logan". Это виртуальный объект и в этой области функция $u(x, y)$ изменяется в диапазоне от -1 до 1. Данный фантом является моделью черепа человека. Исторически он был использован для тестирования методов медицинской томографии. Модель Sheep-Logan является чрезвычайно трудным объектом для восстановления, т.к. содержит множество пересекающихся областей, находящихся внутри замкнутой оболочки. Результаты численного расчета представим на рис. 1.

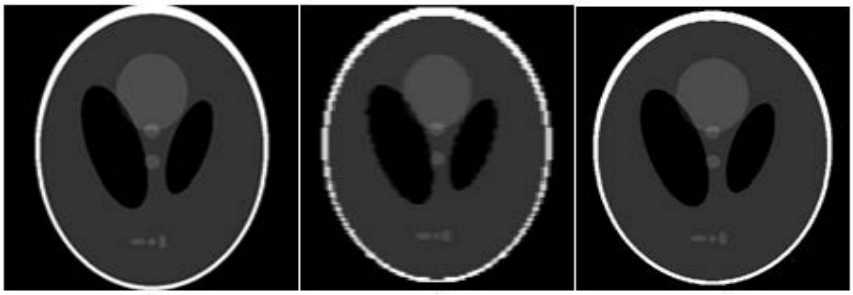


Рис. 1: Восстановление тестового изображения. a) оригинал фантом b) восстановление при $N = 128$, c) восстановление при $N = 512$

Пример 2. На втором примере в качестве тестового примера для сравнения с результатами численного расчета используем следующую функцию

$$f(x, y) = \frac{4}{105} y^{3/2} \left(35 + y\,(-63 + 8y) + 7x^2\,(-5 + 8y)\right)$$

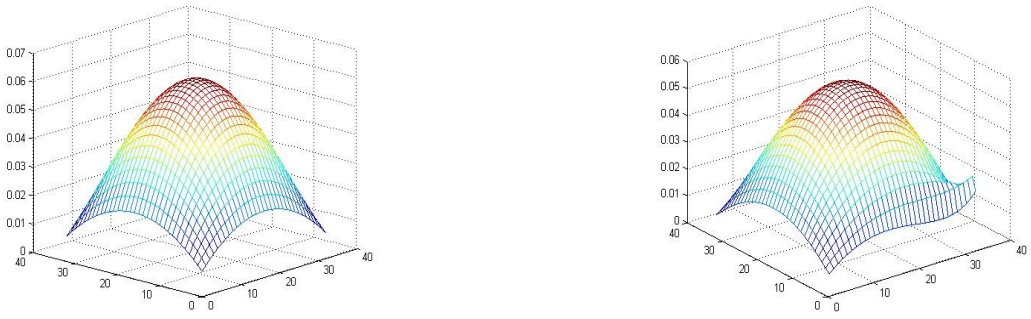
Как и ранее введем равномерную сетку в прямоугольной области $D = [-1,\, 1] \times [0; 0, 5]$.
Схема алгоритма решения задачи такова:
Шаг 1. Разобьем отрезок $[-1; 1]$ на оси Ox и $[0; 0, 5]$ на оси Oy на $n_x - 1$ и $n_y - 1$ частей соответственно. $x_i = a + (i-1)\, h_x,\ \ y_j = b + (j-1)\, h_y$.
Шаг 2. Приближения функций $u(x_i,\, y_j)$ будем обозначать u^A_{ij}.

$$u_\alpha(x_i, y_j) = \frac{F_{ij+1} - F_{ij-1}}{4\alpha h_y \sqrt{\pi^3}} \tag{20}$$

где

$$F_{ij} = \int_0^{y_j} \int_{-1}^{+1} e^{\frac{y_j - \eta - (x_i - \xi)^2}{4\alpha^2}} \cos \frac{(x_i - \xi)\sqrt{y_j - \eta}}{2\alpha^2} \frac{f(\xi, \eta)}{\sqrt{y_j - \eta}} d\xi d\eta,$$

$\alpha-$ параметр регуляризации. Подсчитаем искомые значения по формуле (20). Результаты вычисления представим на рис.2.



a) b)

Рис. 2: Результаты восстановления тестовой функции a) исходная функция, b) результат восстановления при $N = 40$

Список литературы

[1] Лаврентьев М. М., Романов В. Г., Шишатский С. П. Некорректные задачи математической физики и анализа. Москва: Наука, 1980.

[2] Natterer F. The mathematics of computerized tomography. 2nd ed., Classics Appl. Math. 32, SIAM, Philadelphia, 2001.

[3] Natterer F., Wubbeling F. Mathematical methods in image reconstruction. Philadelphia: SIAM, 2001.

[4] Бегматов А.Х., Пиримбетов А. О., Сеидуллаев А.К. Задачи интегральной геометрии в полосе на семействах параболических кривых // Доклады АН ВШ РФ, 2012, Том №2 (19), С. 6–15.

[5] Begmatov A.H., Pirimbetov A.O., Seidullaev A.K. Reconstruction stability in some problems of X-ray and seismic tomography // Proceedings of IFOST-2012, Tomsk Polytechnic University, Volume II, pp. 261–266.

[6] Бегматов А. Х., Джайков Г. М. О восстановлении функции по сферическим средним // Доклады АН ВШ РФ, 2013. Т.1, № 20. С. 6–16.

[7] Nowack R. L. Tomography and the Herglotz-Wiechert inverse formulation.Pure and Apllied Geophysics. 1990. Vol.133, pp.305–315.

[8] Natterer F. The Mathematics of Computerized Tomography. B. G. Teubner. 1986, 222 p.

[9] Кабанихин С. И. Обратные и некорректные задачи. Новосибирск: Сибирское научное издательство. 2009. 457 с.

[10] Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука 1979. 284 стр.

Сеидуллаев Абат Камалович — ассистент Каракалпакского государственного университета;
e-mail: abat_1984@inbox.ru;
Джайков Гафур Муратбаевич — старший преподаватель Нукусского филиала Ташкентского университета информационных технологий;
e-mail: gafur_djaykov@mail.ru.

Дата поступления — 30 апреля 2019 г.

МАТРИЧНЫЙ АЛГОРИТМ ВЫЧИСЛЕНИЯ ЦЕПНЫХ ДРОБЕЙ

В. А. Селезнев, Л. В. Пехтерева

Новосибирский государственный технический университет, 630073, Новосибирск

УДК 511.41
DOI: 10.24411/9999-016A-2019-10071

Рассматривается матричный алгоритм представления цепных дробей. Этот алгоритм, во-первых, позволяет вычислять цепные дроби как произведение двумерных матриц. Во-вторых, имеет геометрическую реализацию в виде морфизмов двумерной решетки с координатами из натуральных чисел, что дает некоторые его геометрические применения.

Ключевые слова: представления цепных дробей, вычисления цепных дробей.

Введение

Матричное представление конечных цепных дробей мы получаем, как алгебраическую формализацию алгоритма Евклида, представляющего рациональные числа в виде конечных цепных дробей. Полученное представление позволяет интерпретировать алгоритм Минковского «вытянутых носов» в виде действия композиции унимодулярных матриц, представляющих рациональные числа. В качестве приложений получено представление унимодулярных морфизмов целочисленной решетки, основанное на взаимно однозначном соответствии множества рациональных чисел и множества унимодулярных матриц второго порядка с элементами из натуральных чисел. Приводятся свойства полученного представления.

1 Матричное представление конечной цепной дроби

Напомним, [1], что любое рациональное число $\frac{p}{q} \geq 1$ разлагается в конечную цепную дробь $\frac{p}{q} = \{a_0; a_1, a_2, ..., a_n\}$, где $a_k \in N$, по правилу

$$\frac{p}{q} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + ... + \frac{1}{a_n}}}. \tag{1}$$

Заметим, что это представление является суперпозицией дробно-линейных функций $a + \frac{1}{z} = \frac{az+1}{z}$, каждую из которых можно представить как действие оператора $L_a(z)$ в матричном виде

$$L_a(z) = \frac{az+1}{z} \equiv \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} (z) \tag{2}$$

Используя (2), по индукции можно показать, что конечная цепная дробь (1) при условии $\frac{p}{q} \geq 1$ представляется в виде матричного произведения

$$\frac{p}{q} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} ... \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} (a_n) \tag{3}$$

Если последний элемент в этом представлении $a_n \neq 1$, то представим его в виде матричного действия

$$a_n = (a_n - 1) + 1 = (a_n - 1) + \frac{1}{1} \equiv L_{a_n - 1}(1).$$

ISBN 978-5-901548-42-4

Учитывая, что взятие обратной величины представляется в виде матричного действия $\frac{1}{z} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}(z)$, получаем следующее

Предложение 1. *Всякая конечная цепная дробь (1) при $\frac{p}{q} \geq 1$ единственным образом представляется в виде матричного произведения*

$$\frac{p}{q} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}(1) = \prod_{k=0}^{n}\begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix}(1),\ a_k \in N \tag{4}$$

а при $\frac{p}{q} < 1$

$$\frac{p}{q} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}(1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\prod_{k=0}^{n}\begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix}(1),\ a_k \in N \tag{5}$$

Перемножая матрицы в представлениях (4) и (5), получаем

Следствие 1. *Представления (4) и (5) редуцируются к виду, соответственно*

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}(1) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}(1) \equiv \frac{A+B}{C+D}$$

и

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}(1) = \begin{pmatrix} C & D \\ A & B \end{pmatrix}(1) \equiv \frac{C+D}{A+B}$$

где по индукции несложно проверить, что

$$det\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} = (-1)^{n+1}$$

$$det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} = (-1)^{n}.$$

Из предложения 1 и следствия 1 следует

Теорема 1. *(Об унимодулярном представлении рациональных чисел). Всякое рациональное число $\frac{p}{q}$ единственным образом представляется в виде действия унимодулярной матрицы на единицу по правилу*

$$\frac{p}{q} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}(1) = \frac{A+B}{C+D}$$

где $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \pm 1$

Единственность этого представления доказывается от противного.

Следствие 2. *Теорема 1 устанавливает взаимно однозначное соответствие между множеством положительных рациональных чисел и множеством унимодулярных 2×2 матриц, элементами которых являются натуральные числа.*

Следствие 3. *Всякая унимодулярная 2×2 матрица, элементы которой есть положительные натуральные числа, представляется при $A + B \geq C + D$ в виде*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}$$

и при $A + B < C + D$ в виде

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\ldots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}$$

Действительно, по теореме 1 такой матрице соответствует определенное рациональное число, а требуемое представление следует из следствия 1.

2 Представление унимодулярных морфизмов натуральной числовой решетки рациональными числами

Представление рациональных чисел, полученное в теореме 1, позволяет построить морфизмы натуральных числовых решеток. Натуральную числовую решетку $\{N\times N\}$ определим, как множество пар положительных натуральных чисел $(p,q)\in\{N\times N\}$.

Теорема 2. *Для любой унимодулярной матрицы* $\begin{pmatrix}A & B\\ C & D\end{pmatrix}$, *отображение*

$$(p,q)\in\{N\times N\}\to L\Big(\frac{p}{q}\Big)\to(P,Q)\in\{N\times N\}$$

определяемое как последовательность операций

$$(p,q)\to$$

$$\to L\Big(\frac{p}{q}\Big)=\begin{pmatrix}A & B\\ C & D\end{pmatrix}\Big(\frac{p}{q}\Big)=\frac{A\frac{p}{q}+B}{C\frac{p}{q}+D}=\frac{Ap+Bq}{Cp+Dq}\to$$

$$\to(Ap+Bq,Cp+Dq)\equiv(P,Q)\in\{N\times N\},$$

есть морфизм $L:\{N\times N\}\to\{N\times N\}$.

Из теорем 1 и 2 получаем

Следствие 4. *Всякое рациональное число* $\frac{p}{q}$ *определяет унимодулярный морфизм натуральной числовой решетки, переводящий точку плоскости* $(1,1)$ *в точку плоскости* (q,p).

Это следствие позволяет построить унимодулярный нормированный морфизм натуральной числовой решетки, отображающий точку (p_1,q_1) в точку (p_2,q_2) следующим образом.
Пусть $L_1:(1,1)\to(p_1,q_1)$, $L_2:(1,1)\to(p_2,q_2)$, тогда, если дроби $\frac{p_1}{q_1}$ и $\frac{p_2}{q_2}$ не сократимы, то искомый морфизм есть композиция $L_2\otimes L_1^{-1}$, для которой

$$\frac{p_2}{q_2}=L_2(L_1^{-1}\Big(\frac{p_1}{q_1}\Big)) \tag{6}$$

В случае точек решетки с взаимно сократимыми координатами (mp_1,mq_1) и (np_2,nq_2) имеем морфизм, отображающий точку (mp_1,mq_1) в точку (np_2,nq_2) в виде

$$nI\otimes L_2\otimes L_1^{-1}\otimes\frac{1}{m}I, \tag{7}$$

где nI и $\frac{1}{m}I$ I — это 2×2 — матрицы растяжения в n и $\frac{1}{m}$ раз, соответственно.
Заметим, что морфизмы натуральной числовой решетки, определяемые рациональными числами в теореме 2 построены не на аффинных, а на дробно линейных отображениях. Тем не менее, учитывая унимодулярность матриц в представлении (6) теоремы 1, несложно проверить, что образ и прообраз многоугольников с вершинами в точках решетки сохраняют площадь. Для проверки указанного свойства достаточно проверить, что площадь любого треугольника с вершинами в узлах решетки сохраняется при отображении задаваемом действием матриц вида $\begin{pmatrix}a & 1\\ 1 & 0\end{pmatrix}$, которые участвуют в представлениях следствия 3.

Заключение

1) О представлении бесконечных цепных дробей.
Известно три способа записи цепных дробей: Прингсгейма, Мюллера и Роджерса, [2]. Матричный способ представления (4), (5) позволяет записать бесконечную цепную дробь, представляющую иррациональные числа в виде

$$r=\prod_{k=1}^{\infty}\begin{pmatrix}a_k & 1\\ 1 & 0\end{pmatrix}(1),\ a_k\in N,\ r>1; \tag{8}$$

$$r = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \prod_{k=1}^{\infty} \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} (1),\ a_k \in N,\ r < 1; \tag{9}$$

В случае приближения бесконечной цепной дроби подходящей конечной цепной дробью мы возвращаемся к представлениям (4) и (5).

2) О морфизме решеток натуральных чисел.

Согласно следствию 2 рациональные числа генерируют унимодулярные морфизмы натуральной числовой решетки, сохраняющие площади многоугольных фигур на этой решетке.

Известный алгоритм Минковского «вытянутых носов», [3], стр.105, может быть реализован итерациями точки $(1, 1)$ натуральной числовой решетки последовательным действием унимодулярных матриц $\begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix}$ в представлениях (8) или (9). Есть другие геометрические реализации, связанные с построенными унимодулярными морфизмами.

3) О приложениях к представлениям рациональных выражений.

Построенное матричное представление цепных дробей остается верным на формальных структурах, допускающих применяемые в полученном представлении операции.

Список литературы

[1] Хинчин А. Я. Цепные дроби. М.: Наука, 1978. — 111с.

[2] Хованский А.Н. Приложение цепных дробей и их обобщений к вопросам приближенного анализа. – М.: государственное издательство технико-теоретической литературы, 1956. —203с.

[3] Арнольд В.И. Дополнительнвые главы теории обыкновенных дифференциальных уравнений. – М.: Наука, 1978. — 304с.

Селезнев Вадим Александрович — д.ф.-м.н., зав. кафедрой инженерной математики Новосибирского государственного технического университета;
e-mail: seleznev@corp.nstu.ru;
Пехтерева Лина Вадимовна — к.т.н., старший преподаватель кафедры высшей математики Новосибирского государственного технического университета;
e-mail: pekhtereva@corp.nstu.ru;
Дата поступления — 28 мая 2019 г.

АЛГОРИТМЫ АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ МЕДИЦИНСКИХ ДАННЫХ

К. В. Симонов[1], А. Г. Зотин[2], Ю. А. Хамад[3], М. А. Курако[3], Т. В. Черепанова[4]

[1] *Институт вычислительного моделирования СО РАН, 660036, Красноярск*
[2] *СибГУ им. М.Ф. Решетнёва, 660037, Красноярск*
[3] *Сибирский федеральный университет, 660041, Красноярск*
[4] *КрасГМУ им. проф. В.Ф. Войно-Ясенецкого, 660022, Красноярск*

УДК 550.8.05
DOI: 10.24411/9999-016A-2019-10072

Исследование посвящено разработке вычислительной методики для оценки показателей процесса регенерации тканей в эксперименте с применением сетчатых никелид-титановых имплантатов с памятью формы. Обработка и анализ данных наблюдений электронной микроскопии и классического гистологического исследования выполняются с применением авторских алгоритмов и их модификаций, которые позволяют существенно упростить процедуру анализа данных и повысить точность оценок (на 5–10%). Предлагаемая методика в качестве вычислительного инструментария для анализа динамики изучаемого процесса, а также для выделения внутренних геометрических черт экспериментальных изображений объектов интереса, содержит алгоритмы шиарлет- и вейвлет-преобразований, а также алгоритмы построения упругих карт для эффективной визуализации пространственных данных. Важным аспектом предлагаемой методики являются вычислительные средства предобработки визуальных данных для повышения контрастности и яркости анализируемых изображений на основе технологии Retinex, что существенным образом влияет на качество применения компьютерных оценочных средств, которые предлагаются в работе.

Ключевые слова: медицинское изображение, обработка изображений, шумоподавление, retinex, вейвлет-преобразование, шиарлет-преобразование, медианный фильтр, упругие карты.

Введение

В настоящее время сетчатые имплантаты широко применяются в различных областях пластической хирургии [1]. Одной из задач современной герниопластики, решаемой с помощью имплантатов, является восстановление дефектов передней брюшной стенки, т.е. хирургическое лечение грыжевых дефектов. В рамках этой задачи проводятся эксперименты с использованием различных материалов, а также различных методов гистологического анализа тканей [2].

Одним из наиболее распространённых решений в качестве вычислительного инструмента для анализа визуальных данных является вейвлет- и шиарлет-преобразование [3, 4]. Так, в работе [5] описаны методы комплексной обработки экспериментальных клинических данных. Шарлет-трансформация используется как основа многомерного анализа пространственных данных и медицинских изображений для решения гистологических задач.

С учётом опыта исследований в данной предметной области количественной гистологии и алгоритмических разработок в области спектрального разложения и визуализации экспериментальных клинических данных предложено новое алгоритмическое решение проблемы количественного морфологического (геометрического) представления визуальных данных. Он основан на использовании шиарлет-преобразования и цветового кодирования. Данное решение направлено на снижение сложности и трудоёмкости проведения медицинского эксперимента при анализе динамики роста тканей на имплантатах, а также на формирование более чёткого различения внутренних геометрических особенностей объектов, представляющих интерес на экспериментальных медицинских изображениях.

ISBN 978-5-901548-42-4

1 Описание медицинского эксперимента

Экспериментальные исследования проведены на 39 кроликах породы шиншилла обоих полов. Выбор кроликов в качестве экспериментальных животных обусловлен наилучшим соответствием биологической реакции тканей животных на имплантацию синтетического материала и реакции тканей человека и в то же время предоставлением возможности получить достаточное количество биопсийного материала относительно большого размера.

Для исследования отобраны животные без внешних признаков заболевания, помещённые на карантин в условиях вивария. В эксперименте участвовали 3 группы сравнения по 11 животных в каждой и одна контрольная группа, включающая 6 животных.

В лабораторных исследованиях использованы следующие материалы:

- нить PDS plus — моноволоконный синтетический рассасывающийся шовный материал, который представляет собой полидиоксаноновую нить, содержащую полидиоксанон и триклозан;
- стандарт «Эсфил» — классический сетчатый эндопротез из монофиламентного полипропилена;
- сетчатый имплантат на основе никелида титана (марка TN-10), который обладает биохимической и биомеханической совместимостью с подвижными анатомическими структурами организма.

Наибольший интерес для медицинских специалистов вызвали эксперименты на группе с использованием имплантата, сотканного по текстильной технологии из суперэластичной никель-титановой пряжи (марка TN-10) размером 5×4 см, толщиной 60 мкм, с размерами ячеек 120–240 мкм. Нить представляет собой композитный материал, содержащий сердцевину из наноструктурированного монолитного никелида титана и пористый поверхностный слой (5–7 мкм) из оксида титана.

Классический подсчёт проводили на микроскопе Bimam при увеличении $10 \times 1,5 \times 20$ и $10 \times 15 \times 40$ с использованием окулярного микрометра и сетки с равноотстоящими точками. Для стандартизации данных морфометрию тканевых структур проводили в соответствии с рекомендациями [6].

В ходе гистологического исследования оценивались следующие параметры: количество клеток соединительной ткани (фибробласты, фиброциты), клеточных элементов воспалительного ряда, аморфного компонента, размера сосудов, толщины волокнистых элементов, эластичных и коллагеновых волокон в исследуемой группе и группах сравнения. В качестве отдельного параметра были оценены признаки хронического воспаления — наличие гигантских клеток.

Дополнительные медицинские исследования проводились с помощью растрового (сканирующего) электронного микроскопа. Он имеет определенные преимущества по сравнению с просвечивающими электронными микроскопами. Одним из них является более простой метод подготовки образца. Определение степени прорастания тканей в сетчатых структурах имплантата на основе никелида титана проводили на электронном микроскопе TM 3500 с использованием программного обеспечения QUANTAX 70 EDS.

2 Методика анализа экспериментальных изображений

В случае каждого периода (дня) отбора проб было сделано 10–15 снимков. Это, в свою очередь, значительно увеличило объём работы для исследователей. Наиболее трудоёмкой оценкой прорастания тканей на имплантате является расчёт показателей, демонстрирующих динамику изучаемого процесса, таких как площадь коллагеновых волокон, количество эритроцитов и др. Для ускорения и повышения точности расчётов решено использовать обработку на основе шиарлет-преобразования и цветового кодирования.

За основу был взят метод [5], где для аппроксимации и спектрального разложения используется алгоритм FFST шиарлет-преобразования [3]. Поскольку результаты анализа могут быть подвержены искажениям из-за присутствия шума, возникающего в цифровой матрице, или характера области освещения тестового образца, мы предлагаем использовать следующие обобщённые этапы (вычислительные процедуры):

1. Определение деталей, представляющих интерес для исследуемого объекта-структуры, в рамках которой осуществляется определение характеристик эталонных объектов (структурных и текстурных особенностей).
2. Применение предварительной обработки изображения с целью подавления шумов, возникающих в результате формирования изображения, а также для коррекции яркости и локальной контрастности анализируемых изображений.

3. Преобразование изображения в форму, удобную для визуальной интерпретации и дальнейшего анализа. На этом этапе определяются параметры алгоритма вычисления шиарлет-преобразования и карты цветового кодирования.

4. Моделирование вычисленных изображений на основе шиарлет-преобразования и определённой шкалы цветового кодирования для получения оптимальной сегментации и цветовой визуализации конкретных объектов, представляющих интерес.

5. Количественная оценка геометрических, текстурных и морфологических характеристик выбранных объектов интереса для решения поставленной предметной задачи.

Принимая во внимание описанные этапы анализа медицинских изображений, предлагаются следующие шаги вычислительной методики для компьютерной обработки медицинских изображений (рис. 1):

- Применение фильтра шумоподавления;
- Проведение коррекции яркости и повышения контрастности;
- Формирование контурного представления;
- Цветовое кодирование объектов в контурном представлении и формирование карт краёв и областей;
- Проведение анализа данных.

Обобщённая форма процедуры обработки медицинских изображений показана на рисунке 1. Ввиду предлагаемого способа шаг 5 анализа данных зависит от данных (изображений) и задач.

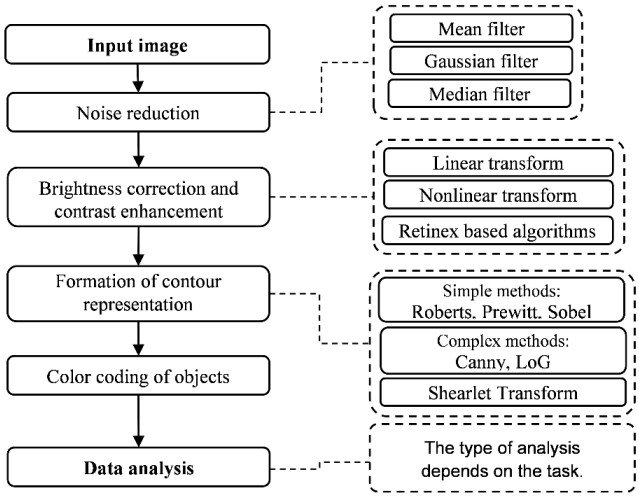


Рис. 1: Структурная схема предлагаемого способа анализа медицинских изображений

Анализ медицинских изображений для клинических экспериментов на имплантатах состоит в выявлении геометрических особенностей текстуры, выделении волокон и сосудов различных типов и их количественном определении, а также сегментации и очерчивании аморфных образований с оценкой их размеров и форм. При оценке прорастания тканей также учитывается зона охвата соответствующего имплантата.

На стадии предварительной обработки для подавления возникающих из-за технических особенностей оборудования шумов выполняется шумоподавление с использованием различных фильтров. Основное назначение этих фильтров — снижение уровня шума, но они также могут быть использованы для выделения объектов на изображении.

В качестве основного используется медианный фильтр, поскольку он вместе с высокой производительностью при подавлении импульсных помех позволяет сохранить чёткость границ контрастируемых объектов. Медианный фильтр заменяет значение интенсивности пикселя медианой значений интенсивности в окрестности этого пикселя. Для подавления небольшого аддитивного шума используется гауссов фильтр.

Медицинские изображения в зависимости от условий съёмки могут иметь различную яркость и контрастность, поэтому для улучшения качества изображений было решено использовать модификацию Multi-Scale

Retinex (MSR) в цветовом пространстве HSV, в которой вейвлет-преобразование использовалось для ускорения расчётов [7].

Контурное представление может быть сгенерировано с использованием разных алгоритмов. Мы можем использовать как классические простые методы, такие как перекрёстный оператор Робертса, операторы Собеля и Прюитт, так и более сложные, например, оператор Кэнни и LoG [8–10].

Для оконтуривания объектов интереса используется алгоритм быстрого конечного шиарлет-преобразования (FFST) [3]. Он основан на дискретном быстром прямом и обратном преобразованиях Фурье и использует вейвлет Мейера в качестве материнского вейвлета. Изучение алгоритма FFST показывает, что контуры объектов интереса могут быть получены как алгебраическая сумма коэффициентов для фиксированного масштаба и всех возможных значений параметров сдвига и смещения.

Чтобы лучше понять представленные медицинские данные и упростить дальнейший анализ, предлагается использовать метод контрастирования медицинских изображений, основанный на цветовом кодировании. Этот метод позволяет привлечь внимание медицинских специалистов к определенным областям анализируемого изображения. Одновременно проводится подготовка к дальнейшему анализу, в ходе которого решаются задачи, необходимые для количественной гистологии.

Для цветового кодирования (выбор цветов и распределение плотности контуров в изображении) используется алгоритмическая процедура с применением метода построения упругой карты из пространственных данных [11]. Предполагается, что анализируемый объект представляет собой ограниченное двумерное многообразие, встроенное в набор изучаемых данных. В этом случае карта объекта является частью криволинейной поверхности, расположенной внутри облака данных анализируемого изображения.

3 Результаты моделирования экспериментальных данных

Оценка методики проводилась с использованием данных, полученных в ходе медицинских экспериментов на кроликах породы шиншилла. Эксперименты проводились на кафедре хирургии Красноярского государственного медицинского университета. Целью медицинского эксперимента было выявление особенностей реакции тканей в зависимости от материала имплантата. Для экспериментального исследования использовалось 140 изображений, охватывающих по времени весь период медицинского эксперимента.

В случае гистологических изображений диапазон стандартного отклонения после коррекции освещённости уменьшался. Влияние коррекции яркости на карту краёв и карту областей после цветового кодирования продемонстрировано для гистологического изображения и для сканирующего электронного микроскопа. Пример улучшения качества одного изображения из набора растровой микроскопии с использованием модифицированного алгоритма Multi-Scale Retinex показан на рисунке 2. Более чёткое визуальное разделение титанового имплантата и тканей (коллагеновые волокна, эритроциты и т.д.) может наблюдаться на обработанном изображении.

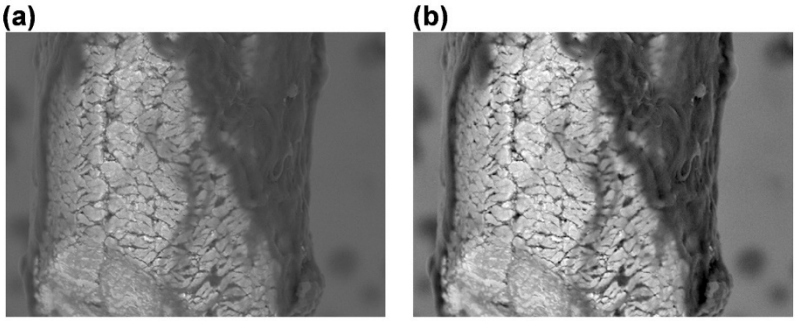


Рис. 2: Пример обработки изображения: a — исходное изображение, b — обработанное изображение

Основная часть вычислительного эксперимента — проведение расчётов с использованием алгоритмов шиарлет-преобразования. Итог расчётов — получение количественных характеристик, таких как контуры и общая площадь исследуемого имплантата, а также контуры и доля ткани, проросшей на сетке. В работе представлены оценки данных морфометрических показателей по погрешности расчёта по основным исследуемым параметрам. Оценка проводится для следующих вычислительных методов: базовая методика расчёта; расчёт с цветовым кодированием; расчёт с цветовым кодированием и коррекцией яркости.

Пример визуальных данных, сгенерированных во время применения предлагаемой методики с использованием цветового кодирования, показан на рисунке 3.

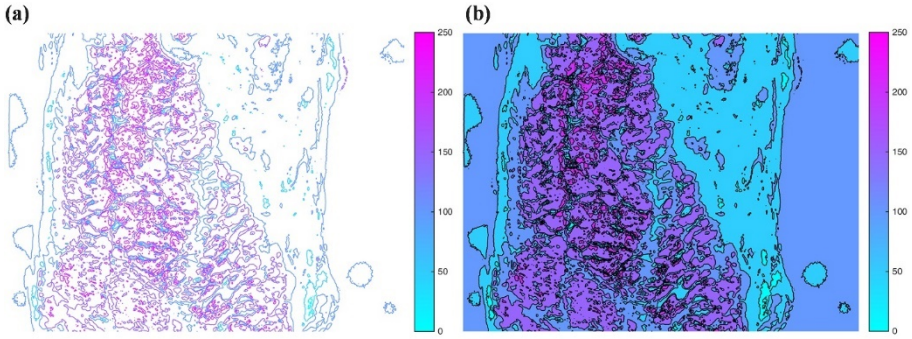


Рис. 3: Пример цветовой кодировки изображения: a — карта краёв (контуров), b — карта областей

Приведены оценки влияния коррекции яркости на рассчитанные морфометрические показатели. В случае обработки без коррекции приводятся средние значения индикаторов, а для результатов с коррекцией также предоставляется информация о минимальных и максимальных полученных значениях.

Оценка прорастания тканей проводилась на основе изображений сканирующей электронной микроскопии. Поэтому для получения более объективных данных были использованы изображения, полученные с разных ракурсов. Объекты интереса были размечены медицинским экспертом на несколько групп: ткани, волокна, эритроциты и т.д., а области со структурой имплантата были определены отдельно. Для указанных эталонных образцов параметры были рассчитаны с учётом показателей текстурных характеристик и цветового кода.

Доля проросшей ткани определялась как отношение площади проросшей ткани к площади соответствующей части сетки имплантата. Сравнение показателей прорастания тканей, полученных медицинским экспертом на основе классического оценочного подхода, проводилось с помощью предложенного метода в двух режимах — с коррекцией яркости и без этой коррекции.

В таблице 1 приведены данные о прорастании тканей, полученные при измерении изучаемых геометрических показателей текстуры для нескольких этапов медицинского эксперимента (7, 14, 21 и 45 дней).

Таблица 1: Оценка площади прорастания тканей (в %)

Период	Ракурс 1	Ракурс 2	Ракурс 3
Базовая методика расчёта			
7 дней	24,4	24,7	24,9
14 дней	47,9	48,1	48,3
21 день	59,4	59,6	59,8
45 дней	68,3	68,5	68,7
Расчёт с цветовым кодированием			
7 дней	27,1	27,4	27,6
14 дней	53,1	53,3	53,5
21 день	66,0	66,2	66,4
45 дней	76,0	76,1	76,3
Расчёт с цветовым кодированием и коррекцией яркости			
7 дней	28,1	28,3	28,5
14 дней	55,1	55,3	55,4
21 день	68,3	68,5	68,7
45 дней	78,6	78,8	78,9

Таким образом, построена зависимость площади проросшей ткани от времени. Предварительная коррекция яркости позволяет более чётко отделить ткани и имплантаты. По результатам анализа оценок гистологических параметров можно сделать вывод, что предварительная обработка изображений с учётом коррекции

яркости повышает точность оценок. Коррекция яркости на основе retinex также позволяет снизить разброс значений (погрешность) с 8–12% до 2–3%.

Заключение

Анализ результатов обработки данных сканирующего электронного микроскопа показывает эффективность разработанной вычислительной методологии для оценки основных показателей количественной гистологии. Это значительно увеличивает скорость обработки данных и получения обновлённых оценок.

Использование механизма коррекции яркости, основанного на технологии retinex, позволяет медицинским специалистам более чётко видеть детали изучаемых объектов. При оценке прорастания тканей это выражается в более чётком определении областей, в среднем на 1–3%, что важно с точки зрения оценки медицинских показателей. Разработанные алгоритмические инструменты полезны для опытных гистологов при обработке больших объёмов визуальных данных.

Список литературы

[1] Gunther V.E. Shape memory biomaterials and implants // Proceedings of International conference. Tomsk, Russia. 2001. p. 28–30.

[2] Радкевич А.А., Винник Ю.С. Каспарова И.Э. Репаративный десмогенез после замещения соединительно-тканных дефектов тканевыми имплантатами из никелида титана // Сибирский медицинский журнал. 2009. Т. 91. № 8. с. 151–153.

[3] Hauser S., Steidl G. Fast finite shearlet transform // arXiv preprint arXiv:1202.1773. 2012.

[4] Kutyniok G., Labate D. Introduction to Shearlets. In: Kutyniok G., Labate D. (eds) Shearlets. Applied and Numerical Harmonic Analysis. Birkhauser Boston. 2012.

[5] Zotin A., Simonov K., Kapsargin F., Cherepanova T., Kruglyakov A., Cadena L. Techniques for medical images processing using shearlet transform and color coding // Intelligent Systems Reference Library. 2018. V. 136. p. 223–259.

[6] Автандилов Г.Г. Медицинская морфометрия: руководство. М.: Медицина, 1990. 198 с.

[7] Zotin A. Fast Algorithm of Image Enhancement based on Multi-Scale Retinex // Procedia Computer Science. 2018. Т. 131. p. 6–14.

[8] Davies E.R. Computer and machine vision: theory, algorithms, practicalities. Academic Press. 2012.

[9] Szeliski R. Computer vision. Algorithms and applications. Springer-Verlag London Limited. 2011.

[10] Canny J. A computational approach to edge detection // Readings in computer vision. Morgan Kaufmann. 1987. p. 184–203.

[11] Горбань А.Н., Зиновьев А.Ю., Питенко А.А. Визуализация данных. Метод упругих карт. // Нейрокомпьютеры. 2002. № 4. с. 19–30.

Симонов Константин Васильевич — д.т.н., в.н.с. Института вычислительного моделирования СО РАН;
e-mail: simonovkv@icm.krasn.ru;
Зотин Александр Геннадьевич — к.т.н., доцент Сибирского государственного университета науки и технологий;
e-mail: zotinkrs@gmail.com;
Юсиф Хамад — аспирант Института космических и информационных технологий Сибирского федерального университета;
e-mail: y.albayati8@gmail.com;
Курако Михаил Александрович — ст. преп. Института космических и информационных технологий Сибирского федерального университета;
e-mail: mkurako@sfu-kras.ru;
Черепанова Татьяна Валентиновна — врач-хирург Красноярского государственного медицинского университета;
e-mail: grakova@list.ru.

Дата поступления — 30 апреля 2019 г.

ОБУСЛОВЛЕННОСТЬ ОБРАТНЫХ ЗАДАЧ И ИТЕРАЦИОННЫЕ МЕТОДЫ ИХ РЕШЕНИЯ

С. Б. Сорокин

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск
Новосибирский государственный университет, 630090, Новосибирск

УДК 519.632
DOI: 10.24411/9999-016A-2019-10073

В работе строятся и исследуются неявные градиентные двухслойные итерационные методы численного решения задачи Коши для эллиптических уравнений. Приводятся результаты вычислительных экспериментов.
Ключевые слова: задача Коши, эллиптические уравнения, неявный двухслойный итерационный метод, разделение переменных.

Введение

При численном решении задачи Коши для эллиптических уравнений одним из широко используемых подходов является сведение ее к обратной задаче (см., например, [1], [2]). Некорректность исходной задачи отражается на свойствах оператора обратной задачи в виде очень большой обусловленности его дискретного аналога. Как следствие, применение явных стационарных итерационных методов для нахождения решения соответствующего операторного уравнения с заданной точностью требует выполнения очень большого числа итераций (пропорционального значению обусловленности) [3]. С другой стороны, как показывают проведенные в [4] исследования, в случае стандартных расчетных областей (прямоугольник, круг) при решении задачи Коши для уравнения Лапласа метод разделения переменных позволяет строить экономичные прямые численные алгоритмы. Применение такого сорта алгоритмов в качестве предобусловливателей позволяет строить неявные двухслойные итерационные методы нахождения решения дискретного аналога обратной задачи. В работе строятся и исследуются неявные градиентные двухслойные итерационные методы численного решения задачи Коши для эллиптических уравнений. Приводятся результаты вычислительных экспериментов.

1 Постановка задачи

В области

$$\Omega = \{(x_1, x_2) \in R^2 : \ x_1 \in (a, b), \ x_2 \in (c, d)\}$$

рассмотрим начально-краевую задачу (задача продолжения)

$$\begin{aligned} &Lu \equiv -\frac{\partial}{\partial x_1}(k_1(x_1, x_2)\frac{\partial u}{\partial x_1}) - \frac{\partial}{\partial x_2}(k_2(x_1, x_2)\frac{\partial u}{\partial x_2}) = 0, \quad (x_1, x_2) \in \Omega, \\ &-k_1(a, x_2)\frac{\partial u}{\partial x_1}(a, x_2) = 0, \quad x_2 \in [c, d], \\ &u(a, x_2) = f(x_2), \quad x_2 \in [c, d], \\ &-k_2(x_1, c)\frac{\partial u}{\partial x_2}(x_1, c) = 0, \quad k_2(x_1, d)\frac{\partial u}{\partial x_2}(x_1, d) = 0, \quad x_1 \in [a, b]. \end{aligned} \tag{1}$$

Необходимо найти функцию $u(x_1, x_2)$ в области Ω по данным $f(x_2)$.

ISBN 978-5-901548-42-4

Сведем задачу продолжения (1) к обратной задаче [1], [2]:
Определить функцию $q(x_2)$ из соотношений

$$\begin{aligned} &Lu \equiv -\frac{\partial}{\partial x_1}(k_1(x_1,x_2)\frac{\partial u}{\partial x_1}) - \frac{\partial}{\partial x_2}(k_2(x_1,x_2)\frac{\partial u}{\partial x_2}) = 0, \quad (x_1,x_2) \in \Omega, \\ &-k_1(a,x_2)\frac{\partial u}{\partial x_1}(a,x_2) = 0, \quad x_2 \in [c,d], \\ &u(b,x_2) = q(x_2), \quad x_2 \in [c,d], \\ &-k_2(x_1,c)\frac{\partial u}{\partial x_2}(x_1,c) = 0, \quad k_2(x_1,d)\frac{\partial u}{\partial x_2}(x_1,d) = 0, \quad x_1 \in [a,b] \end{aligned} \tag{2}$$

по дополнительной информации

$$u(a,x_2) = f(x_2), \quad x_2 \in (c,d). \tag{3}$$

Обратную задачу (2)-(3) будем решать на дискретном уровне.

Построим равномерные сетки:

$$\bar{\omega}_1 = \{x_{1,i} = x_{1,i-1} + h_1,\ 1 \le i \le N_1+1,\ x_{1,0} = a,\ x_{1,N_1+1} = b,\ h_1 = \frac{b-a}{N_1+1}\},$$
$$\bar{\omega}_2 = \{x_{2,j} = x_{2,j-1} + h_2,\ 1 \le j \le N_2+1,\ x_{2,0} = c,\ x_{2,N_2+1} = d,\ h_2 = \frac{d-c}{N_2+1}\},$$
$$\bar{\omega} = \{\ x_{i,j} = (x_{1,i},x_{2,j}) \in \bar{\Omega}, \quad x_{1,i} \in \bar{\omega}_1, \quad x_{2,j} \in \bar{\omega}_2\}.$$

Скалярные произведения в пространствах сеточных функций определенных на этих сетках зададим следующим образом (на правом конце сетки $\bar{\omega}_1$ функции принимают нулевое значение)

$$(u^h,v^h)_{\bar{\omega}_1} = 0.5h_1u^h(x_{1,0})v^h(x_{1,0}) + \sum_{i=1}^{N_1} u^h(x_{1,i})v^h(x_{1,i})h_1,$$

$$(u^h,v^h)_{\bar{\omega}_2} = 0.5h_2u^h(x_{2,0})v^h(x_{2,0}) + \sum_{j=1}^{N_2} u^h(x_{2,j})v^h(x_{2,j})h_2 + 0.5h_2u^h(x_{2,N_2+1})v^h(x_{2,N_2+1}),$$

$$(u^h,v^h)_{\bar{\omega}} = (1,(u^h,v^h)_{\bar{\omega}_2})_{\bar{\omega}_1}.$$

Дискретный аналог обратной задачи (2),(3) сформулируем следующим образом: среди всех задач

$$L_hu^h(x_{1,i},x_{2,j}) \equiv \Lambda_1u^h(x_{1,i},x_{2,j}) + \Lambda_2u^h(x_{1,i},x_{2,j}) = 0, \quad i = \overline{0,N_1},\ j = \overline{0,N_2+1} \tag{4}$$

$$u^h(x_{1,N_1+1},x_{2,j}) = q^h(x_{2,j}), \quad j = \overline{0,N_2+1} \tag{5}$$

надо указать такую задачу (такую сеточную функцию $q^h(x_{2,j})$) чтобы ее решение $u^h(x_{1,i},x_{2,j})$ удовлетворяло равенству

$$u^h(x_{1,0},x_{2,j}) = f(x_{2,j}), \quad x_{2,j} \in [c,d]. \tag{6}$$

Действие операторов Λ_1 и Λ_2 определяются следующим образом [3], [5]:

$$\Lambda_1y(x_{1,i},x_{2,j}) = \begin{cases} -\dfrac{2}{h_1}k_1(x_{1,0}+0.5h_1,x_{2,j})y_{x_1}(x_{1,0},x_{2,j}), & i = 0, \\ -(k_1(x_{1,i}-0.5h_1,x_{2,j})y_{\overline{x}_1})_{x_1}(x_{1,i},x_{2,j}), & i = \overline{1,N_1}, \end{cases}$$

$$\Lambda_2y(x_{1,i},x_{2,j}) = \begin{cases} -\dfrac{2}{h_2}k_2(x_{1,i},x_{2,0}-0.5h_2)y_{x_2}(x_{1,i},x_{2,0}), & j = 0, \\ -(k_2(x_{1,i},x_{2,j}-0.5h_2)y_{\overline{x}_2})_{x_2}(x_{1,i},x_{2,j}), & j = \overline{1,N_2}, \\ \dfrac{2}{h_2}k_2(x_{1,i},x_{2,N_2+1}-0.5h_2)y_{\overline{x}_2}(x_{1,i},x_{2,N_2+1}), & j = N_2+1. \end{cases}$$

Также, как это описано, например, в [2], можно, используя задачу (4)-(5), ввести оператор

$$A_h\ :\ q^h(x_{2,j}) \to u^h(x_{1,0},x_{2,j}),$$

где $u^h(x_{1,i}, x_{2,j})$ решение прямой задачи (4)-(5).

Тогда обратная задача (4)-(6) может быть записана в следующем виде:

$$A_h\, q^h(x_{2,j}) \;=\; f(x_{2,j}). \tag{7}$$

Одним из возможных способов решения (7) является минимизация целевого функционала

$$J(q) = ||A_h\, q - f||^2_{\bar{\omega}} = (A_h\, q - f, A_h\, q - f)_{\bar{\omega}},$$

что эквивалентно решению уравнения (трансформация Гаусса)

$$A_h^* A_h\, q(x_{2,j}) \;=\; A_h^* f(x_{2,j}). \tag{8}$$

Постановка задачи: Исследовать итерационные методы решения уравнения (8).

2 Обусловленность и итерационные методы

Некорректность исходной задачи (1) отражается на свойствах оператора $A_h^* A_h$ в виде очень большой его обусловленности. Результаты полученные в [4], позволяют вычислить собственные числа оператора $A_h^* A_h$ для случая постоянных коэффициентов $k_1(x_1, x_2), k_2(x_1, x_2)$ в операторе L. Для $k_1(x_1, x_2) \equiv 1, k_2(x_1, x_2) \equiv 1$ в Таблице 1 приведены максимальное, минимальное собственные числа и обусловленность $A_h^* A_h$ при параметрах сетки $\bar{\omega}$, указанных в первом столбце.

Таблица 1

	$\lambda_{max}(A_h^* A_h)$	$\lambda_{min}(A_h^* A_h)$	$cond_2(A_h^* A_h)$
$h_1 = h_2 = \frac{1}{10}$	1	10^{-5}	10^{5}
$h_1 = h_2 = \frac{1}{20}$	1	10^{-10}	10^{10}
$h_1 = h_2 = \frac{1}{40}$	1	10^{-21}	10^{21}

Следствием наличия у оператора $A_h^* A_h$ такой большой обусловленности, является то, что применение явных стационарных итерационных методов

$$\frac{q^{k+1} - q^k}{\tau} + A_h^* A_h q^k \;=\; A_h^* f \tag{9}$$

для нахождения с заданной точностью решения операторного уравнения (8), требует выполнения неприемлемо большого числа итераций (пропорционального значению обусловленности).

Одним из широко известных подходов, позволяющих уменьшить обусловленность оператора в (8) является метод М.М.Лаврентьева. В этом методе вместо уравнения (8) предлагается решать задачу

$$(A_h^* A_h + \alpha E)\, q_\alpha(x_{2,j}) \;=\; A_h^* f(x_{2,j}). \tag{10}$$

Выбор параметра α подчинен (в частности) условию близости решений q и q_α уравнений (8) и (10), соответственно. В изучаемом нами случае параметр должен удовлетворять неравенству

$$\lambda_{\min}(A_h^* A_h) \ll \alpha \ll 1 = \lambda_{\max}(A_h^* A_h).$$

С точки зрения применения итерационных процессов это существенно не улучшает ситуацию. Действительно, обусловленность оператора $A_h^* A_h + \alpha E$ из (10) выражается в терминах собственных чисел оператора $A_h^* A_h$ в форме

$$cond_2(A_h^* A_h + \alpha E) = \frac{\lambda_{\max}(A_h^* A_h) + \alpha}{\lambda_{\min}(A_h^* A_h) + \alpha}.$$

Учитывая диапазон изменения α, имеем

$$cond_2(A_h^* A_h + \alpha E) \approx \frac{1}{\alpha}.$$

Следовательно, даже если выбрать не слишком маленький параметр α, например $\alpha = 10^{-4}$, то для нахождения решения с заданной точностью потребуется выполнить пропорционально $1/\alpha = 10000$ итераций по явному **стационарному** итерационному методу

$$\frac{q^{k+1} - q^k}{\tau} + (A_h^* A_h + \alpha E) q^k = A_h^* f \tag{11}$$

с оптимальным параметром τ.

Что касается **нестационарных** градиентных итерационных процессов, то известно, что асимптотически метод наискорейшего спуска не лучше, чем метод Ричардсона с оптимальным параметром. Однако, можно надеяться на эффективность их применения для рассматриваемого класса задач в связи со следующим.

Во-первых, метод наискорейшего спуска обладает свойством подавлять преобладающие гармоники.

Во-вторых. Как известно, коэффициенты z_i^k в разложении погрешности $z^k = q^k - q_{exact}$ по собственным функциям оператора задачи убывают по закону

$$z_i^{k+1} = (1 - \tau\lambda_i) z_i^k.$$

Если выбрать начальное приближение $q^0 = 0$, то $z^0 = -q_{exact}$ и в погрешности z^k будут присутствовать только гармоники содержащиеся в точном решении q_{exact}.

В тоже время, если известно, что q_{exact} «гладкая», то ее коэффициенты Фурье быстро убывают с номером гармоники.

Т.о. можно надеяться, что малые коэффициенты в разложении погрешности будут и дальше убывать, а метод наискорейшего спуска будет подавлять большие коэффициенты.

Другой, более эффективный, способ ускорить сходимость итерационной процедуры для решения задачи (10) – применение неявных процессов.

Как показывают проведенные в [4] исследования, использование метода разделения переменных при решении задачи Коши для уравнения Лапласа позволяет строить (в случае стандартных расчетных областей (прямоугольник, круг)) экономичные прямые численные алгоритмы, с числом операций пропорциональным числу узлов сетки $\bar{\omega}$ равным $N_1 N_2$. Применение такого сорта алгоритмов в качестве предобусловливателей приводит к неявным двухслойным градиентным итерационным методам нахождения решения дискретного аналога обратной задачи:

$$(B_h^* B_h + \varepsilon E) \frac{q^{k+1} - q^k}{\tau_{k+1}} + (A_h^* A_h + \alpha E) q^k = A_h^* f. \tag{12}$$

Здесь оператор B_h соответствует оператору Лапласа и определяется также как и A_h :

$$B_h : q^h(x_{2,j}) \to u^h(x_{1,0}, x_{2,j}),$$

где $u^h(x_{1,i}, x_{2,j})$ решение прямой задачи (4)-(5) для оператора Лапласа - $L_h = -\Delta_h$.

Очевидно, что стоимость обращения в (12) переобусловливателя $B_h^* B_h + \varepsilon E$ равная $O(N_1 N_2)$ [4] не является обременительной на фоне затрат, необходимых для вычисления $A_h^* A_h\, q^k$ (требуется два раза решить прямую задачу вида (4)-(5)).

3 Результаты вычислительных экспериментов

Задача (8) решалась методом М.М.Лаврентьева (10) для оператора L заданного в области

$$\Omega = \{(x_1, x_2) \in R^2 : \; x_1 \in (0,1), \; x_2 \in (0, \pi)\}.$$

Коэффициенты выбирались следующим образом:

$$k_1(x_1, x_2) = \begin{cases} 1, & x_1 \in [0, 0.5), x_2 \in [0, \pi] \\ 10, & x_1 \in [0.5, 1], x_2 \in [0, \pi] \end{cases}, \quad k_2(x_1, x_2) = 1, \; x_1 \in [0,1], x_2 \in [0, \pi].$$

Тест строился так.

В качестве точного решения бралась функция

$$q(x_2) = \frac{1}{3}\ \frac{x^2(\pi - x)^2}{exp(x - \pi/10)^2}.$$

Для нее решалась разностная прямая задача (4)-(5). Далее вычислялся след полученного решения $u^h(x_{1,0}, x_{2,j})$ на левой границе Ω, который и принимался за дополнительную информацию:

$$f(x_{2,j}) = u^h(x_{1,0}, x_{2,j}).$$

После этого восстанавливалось выбранное точное решение путем решения (10) итерационными методами - явным наискорейшего спуска и невным наискорейшего спуска (12).

В Таблице 2 представлены результаты расчетов для невозмущенной дополнительной информации f. В первом столбце указаны параметры сеток в прямой задаче (4)-(5), при которых выполнялись вычисления. Во втором столбце приведена точность получаемого итерационным методом решения: $||q^k - q||_2$. В последних двух столбцах помещены номера итераций, на которых была достигнута соответствующая точность по явному ($\alpha = 10^{-6}$) и неявному методам наискорейшего спуска (12) ($\alpha = 10^{-6}$, $\varepsilon = 10^{-2}$), соответственно. Так для сетки $N_1 = 9, N_2 = 9$ точность $||q^k - q||_2 = 0.1$ была достигнута на 8-ой итерации по явному методу наискорейшего спуска и на 6-ой итерации по неявному методу наискорейшего спуска.

Таблица 2

	Точность	Явный, наискорейшего спуска $\alpha = 10^{-6}$	Неявный, наискорейшего спуска $\alpha = 10^{-6}$, $\varepsilon = 10^{-2}$
$N_1 = 9, N_2 = 9$	0.1	8	6
	0.01	37	12
	0.001	98	18
	0.0001	203	24
$N_1 = 19, N_2 = 19$	0.1	10	6
	0.01	55	12
	0.001	178	18
	0.0001	463	28
$N_1 = 39, N_2 = 39$	0.1	10	6
	0.01	62	12
	0.001	212	18
	0.0001	579	35

Анализ представленных в Таблице 2 результатов показывает:

1. Явный метод наискорейшего спуска в реализованном тесте демонстрирует достаточно оптимистичные результаты, существенно лучше теоретических. Теоретическая оценка, для используемого параметра $\alpha = 10^{-6}$, дает асимптотическую оценку числа итераций для метода (11) порядка $1/\alpha = 1000000$ итераций.
2. Для явного метода наискорейшего спуска имеется зависимость числа итераций, необходимого для достижения заданной точности, от размерности задачи.
3. Неявный метод наискорейшего спуска в реализованном тесте показывает практически на порядок лучшие результаты по числу итераций по сравнению с явным методом.
4. Для неявного метода наискорейшего спуска число итераций, необходимое для достижения заданной точности, не зависит от размерности задачи.

В Таблице 3 представлены результаты расчетов решения задачи продолжения для возмущенной дополнительной информации f_δ. В отличие от предыдущего вычислительного эксперимента, в дополнительную информацию f после ее вычисления вносилось возмущение по правилу

$$f_\delta(x_{2,j}) = f((x_{2,j}))(1 + \frac{Pr}{100} \times rand(j)),$$

где Pr процент возмущения, $rand(j)$ равномерное случайное распределение не отрезке [-1,1]. В результате, вместо задачи (10), необходимо решить задачу

$$(A_h^* A_h + \alpha E)\, q_{\delta\alpha}(x_{2,j}) \;=\; A_h^* f_\delta(x_{2,j}). \tag{13}$$

Разность между точным решением q исходной задачи (7) и решением $q_{\delta\alpha}$ возмущенной задачи (13) оценивается (см., например, [2]) следующим образом

$$||q - q_{\delta\alpha}||_2 \le C_1\alpha + C_2\frac{\delta}{\alpha^{\frac{1}{3}}}.$$

Поэтому, при заданной погрешности $\delta = ||f - f_\delta||_2$, параметр α следует выбирать в виде $\alpha = \delta^{\frac{2}{3}}$.

В расчетах, поскольку константы C_1 и C_2 неизвестны, параметр α полагался $\alpha = \delta^{\frac{2}{3}}/10$. Параметр ε в неявном методе наискорейшего спуска (12) выбирался равным 10^{-2}, что в реализованных тестах соответствовало примерно удвоенному α.

В Таблице 3 в первом столбце указаны параметры сеток в прямой задаче (4)-(5), при которых выполнялись вычисления. Во втором столбце приведена вносимая в дополнительную информацию погрешность. На каждой сетке расчеты проводились для погрешностей в 1% и 3%.

В последних двух столбцах помещены номера итераций, на которых необходимая точность была достигнута по явному и неявному методам наискорейшего спуска (12), соответственно. Необходимая точность считалась достигнутой на итерации с номером k если для этой итерации выполнилось неравенство: $||q^{k+1} - q||_2 \ge ||q^k - q||_2$, ($q-$ точное решение исходной задачи (7)). Выполнение этого неравенства означает, что погрешность $||q^k - q||_2 = ||(q^k - q_{\delta\alpha}) + (q_{\delta\alpha} - q)||_2$ достигла уровня $||q_{\delta\alpha} - q||_2 = o(\delta) + o(\alpha)$ и дальнейшее уменьшение первого слагаемого $||q^k - q_{\delta\alpha}||$ не может улучшить качество полученного приближения q^k.

Так для сетки $N_1 = 9, N_2 = 9$ при вносимой в дополнительную информацию погрешности в 1% точность была достигнута на 43-ой итерации по явному методу наискорейшего спуска и на 12-ой итерации по неявному методу наискорейшего спуска.

Таблица 3

	Возмущение	Явный, наискорейшего спуска	Неявный, наискорейшего спуска
$N_1 = 9, N_2 = 9$	$\delta \to 1\%$	43	12
	$\delta \to 3\%$	34	10
$N_1 = 19, N_2 = 19$	$\delta \to 1\%$	97	12
	$\delta \to 3\%$	60	8
$N_1 = 39, N_2 = 39$	$\delta \to 1\%$	100	12
	$\delta \to 3\%$	90	10

Из результатов представленных в Таблице 3 следует:

1. В явном методе наискорейшего спуска для реализованного теста с возмущенной дополнительной информацией f_δ для выполнения критерия останова требуется выполнить по порядку такое же количество итераций как и для невозмущенной дополнительной информации.
2. Для явного метода наискорейшего спуска зависимость числа итераций от размерности задачи сохраняется.
3. Неявный метод наискорейшего спуска в реализованном тесте показывает почти на порядок лучшие результаты по числу итераций по сравнению с явным методом.
4. Для неявного метода наискорейшего спуска число итераций, необходимое для достижения заданной точности, не зависит от размерности задачи.
5. Число итераций, необходимое для достижения заданной точности полностью согласуется с распределением собственных чисел соответствующих спектральных задач:

$$(A_h^* A_h \; + \alpha E)\varphi \;= \lambda\phi, \quad \text{- явный метод} \tag{14}$$

$$(A_h^* A_h \; + \alpha E)\psi \;= \mu(B_h^* B_h + \varepsilon E)\psi \quad \text{- неявный метод.} \tag{15}$$

Результаты работы [4] позволяют вычислить за $O(N_1 N_2)$ операций собственные числа этих задач на собственные значения. На рис. 1 изображено распределение собственных чисел задач (14),(15). Нижняя кривая отвечает задаче (14), верхняя - (15). Вычисления спектров проведены для операторов, соответствующих последнему из расчетов Таблицы 3, в котором использовались параметры: $\alpha = 0.0057$, $\varepsilon = 10^{-2}$. Собственные числа, начиная с номера 10, выходят на аксимптотику равную $\lambda = 0.0057$ для задачи (14) и $\mu = 0.57$ для задачи (15). Т.о. $\lambda_{max}/\lambda_{min} \approx 175$, $\mu_{max}/\mu_{min} \approx 10$, что и определяет необходимое число итераций.

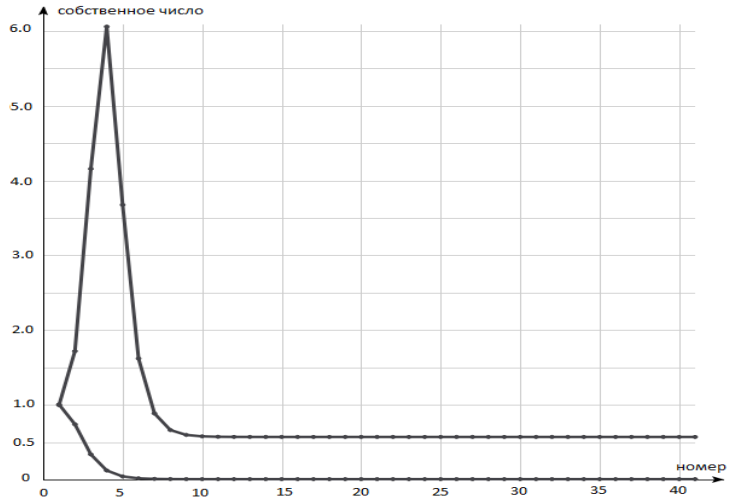


Рис. 1: Собственные числа задач (14),(15). Нижняя кривая соответствует задаче (14), верхняя - (15)

Заключение

В работе исследованы явный и неявный градиентные двухслойные итерационные методы численного решения задачи Коши для эллиптических уравнений. В качестве переобусловливателя в неявном методе использовался экономичный прямой алгоритм решения дискретной задачи Коши для уравнения Лапласа, основанный на методе разделения переменных [4]. Представленные результаты вычислительных экспериментов убедительно свидетельствуют об эффективности применения неявных методов для численного решения задачи продолжения.

Список литературы

[1] S.I. Kabanikhin, A.L. Karchevsky, Method for solving the Cauchy Problem for an Elliptic Equation, Journal of Inverse and Ill-posed Problems, 3. 1 (1995), 21-46.

[2] Кабанихин С.И Обратные и некорректные задачи, Новосибирск: Сибирское научное издательство, 2009. 457с.

[3] Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. М.: Наука. 1978. 529с.

[4] Сорокин С.Б. Экономичный прямой метод численного решения задачи Коши для уравнения Лапласа // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. – Новосибирск, 2019.– Т. 22, № 1, – С. 99–117

[5] Самарский А.А., Андреев В.Б. Разностные методы для эллиптических уравнений, М.: Наука. 1979. 152с.

Сергей Борисович Сорокин — д.ф.-м.н., вед. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: sorokin@sscc.ru;
Дата поступления — 30 апреля 2019 г.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ВОЛНОВЫХ ПРОЦЕССОВ В ГЕОЛОГИЧЕСКИХ СРЕДАХ С ГАЗОВЫМИ КАРМАНАМИ В ЗОНЕ АРКТИЧЕСКОГО ШЕЛЬФА С ПОМОЩЬЮ СЕТОЧНО-ХАРАКТЕРИСТИЧЕСКОГО МЕТОДА

П. В. Стогний[1], Н. И. Хохлов[1,2], И. Б. Петров[1,2]

[1] *Московский физико-технический институт (национальный исследовательский университет), 141701, Долгопрудный*

[2] *Научно-исследовательский институт системных исследований РАН, 117218, Москва*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10074

В работе представлены результаты численного моделирования распространения сейсмических волн в моделях с газонасыщенными средами в течение четырехлетнего периода. Моделирование проводилось с помощью сеточно-характеристического метода на основе схемы Русанова для трехмерного случая. Представлены волновые картины сейсмических откликов от различных геологических слоев с течением времени.

Ключевые слова: Арктика, сеточно-характеристический метод, газовые карманы.

Введение

Газовые карманы — залежи свободного газа с аномально высоким пластовым давлением [1]. Залежи газа представляют опасность для буровых установок, расположенных в водном пространстве, в частности, в Арктической зоне России. Сами по себе газовые залежи не представляют опасности, так как самопроизвольные выбросы газовых залежей в атмосферу случаются крайне редко. Но при бурении скважины может произойти вскрытие газовых карманов. В результате, происходит дальнейший прорыв газа к поверхности воды. Выбросы газа влекут за собой аварийные и катастрофические ситуации с гибелью буровых установок и судов, нередко вместе с членами экипажа. С целью понимания данных процессов, чтобы в дальнейшем уметь предотвращать подобные выбросы, проводится моделирование геологических сред, содержащих подобные залежи, на акватории Мирового океана. В данной работе представлены результаты трехмерного моделирования распространения упругих волн для геологических газонасыщенных сред в Северных морях глубиной до 200-300 м. Получены графические результаты распространения упругих волн сквозь толщу многослойной геологической среды с газонасыщенными породами к поверхности водного пространства для случаев расположения газовых карманов в 1-й, 2-й, 3-й и 4-й года. Представлены волновые отклики непосредственно от геологических сред с газовыми карманами для 2-го, 3-го и 4-го года расчета, соответственно. Все расчеты проводились с помощью сеточно-характеристического метода [2].

1 Математическая модель

Для описания динамических процессов в геологических средах использовались уравнения линейной упругости (1), (2) и акустики (3), (4):

$$\rho \frac{\partial}{\partial t} \vec{v} = (\nabla \cdot \sigma)^T \tag{1}$$

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00366.

ISBN 978-5-901548-42-4

$$\frac{\partial}{\partial t}\sigma = \lambda(\nabla \cdot \vec{v})I + \mu(\nabla \otimes \vec{v} + (\nabla \otimes \vec{v})^T) \tag{2}$$

$$\rho\frac{\partial}{\partial t}\vec{v} = -\nabla\rho \tag{3}$$

$$\frac{\partial}{\partial t}\rho = -c^2\rho(\nabla \cdot \vec{v}) \tag{4}$$

В (1), (2), (3), (4) υ — это скорость движения, σ — тензор напряжения Коши, ρ — плотность материала, I — единичный тензор, t — время, p — давление, c — скорость звука в идеальной жидкости. Во всех расчетах использовались следующие граничные условия. На верхней границе, границе вода-воздух, было установлено условие свободной границы. На боковых границах моделей и нижней границе использовалось не отражающее граничное условие (условие поглощения).

2 Численный метод

Для численного решения систем уравнений (1), (2) и (3), (4) использовался сеточно-характеристический метод [2] на структурированных прямоугольных сетках. Его описание можной найти в работах [3], [4]. Сеточно-характеристический метод позволяет ставить корректные контактные и граничные условия.

3 Результаты

Моделировалось распространение сейсмических волн в слоистой структуре грунта для трехмерного случая [5]. Основная модель для 1-го года расчета состояла из 9 геологических слоев: воды, ила, трех слоев влагонасыщенного песка, трех слоев глины, нефтесодержащего слоя. Модели для 2-го, 3-го и 4-го года расчета дополнительно включали в себя газовые карманы. Схематичное изображение модели для 2-го года расчета с 1 газовым карманом представлено на рис. 1.

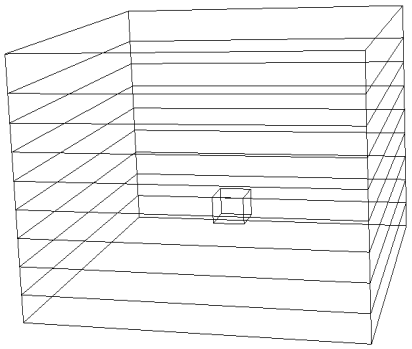


Рис. 1: Схематичное изображение модели для 2-го года расчета с 1 газовым карманом

Расчет всех четырех моделей проводился со следующими параметрами. Шаг по пространству равнялся 1 м. Площадь расчетной области составляла 16 млн. м^2. Шаг по времени равнялся 10^{-4} с. Всего было сделано 20 тыс. шагов по времени. Глубина слоя воды составляла 200 м. Глубина каждого из остальных восьми слоев равнялась 150 м. Воздействие осуществлялось с помощью одиночного источника импульса Рикера, который располагался на поверхности воды в центре расчетной области. 200 приемников сигнала было установлено на поверхности воды в центре расчетной области. На рис. 2 представлены волновые картины сейсмических откликов для моделей с газовыми карманами. На рис. 2 а, 2 в и 2 д изображены волновые поля сейсмических откликов для 2-го, 3-го и 4-го года расчета, соответственно. На рис. 2 б, 2 г и 2 е представлены сейсмические отклики только от газовых карманов, за вычетом поля слоистой среды. Из рис. 2 следует, что волновые отклики от газовых карманов поднимаются со временем, и на 4-й год расчета приближаются к поверхности

воды. Таким образом, можно оценить примерную скорость распространения газа в данной многослойной среде, для которой были выполнены расчеты, и определить, как скоро произойдет выброс газа в атмосферу. В данной постановке задачи газ приближается к поверхности воды уже на 4-й год расчета, что грозит скорым выбросом газа в атмосферу.

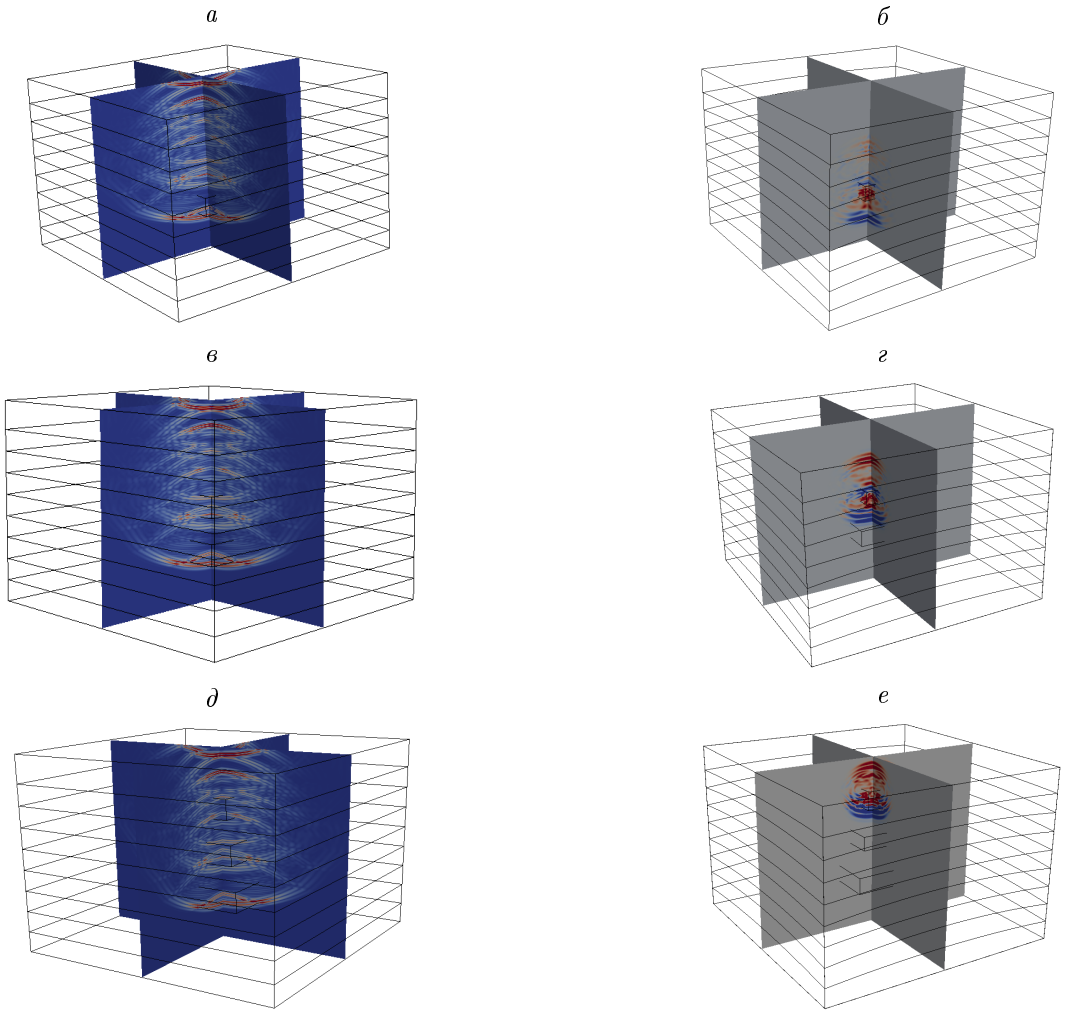


Рис. 2: Волновые поля сейсмических откликов для моделей 2-го, 3-го и 4-го года расчета:
а – волновое поле для модели 2-го года расчета с 1 газовым карманом; *б* – сейсмические отклики только от газового кармана для 2-го года расчета; *в* – волновое поле для модели 3-го года расчета с 2 газовыми карманами; *г* – сейсмические отклики только от газовых карманов для 3-го года расчета; *д* – волновое поле для модели 4-го года расчета с 3 газовыми карманами; *е* – сейсмические отклики только от газовых карманов для 4-го года расчета

Заключение

В работе представлены результаты численного моделирования распространения сейсмических волн сквозь слоистую структуру грунта в трехмерных моделях с газовыми карманами. Получены волновые отражения откликов от геологических сред с газовыми карманами в зависимости от их местоположения относительно поверхности воды. Волновые картины на 4-й год расчета показали существенное отличие сейсмических откликов от газонасыщенных сред по сравнению с аналогичными волновыми полями для моделей 1-го, 2-го и 3-го года расчета. В представленной постановке задачи газ приближается к поверхности воды уже на 4-й год расчета, что грозит скорым выбросом газа в атмосферу. В дальнейшем, планируется провести сравнение

полученных результатов моделирования с реальными исследованиями на акватории Мирового океана для корректировки и улучшения расчетов.

Список литературы

[1] Judd A., Hovland M. Seabed Fluid Flow // The Impact on Geology, Biology, and the Marine Environment. 2007. V. 61, iss. 1. P. 235–248.

[2] Магомедов К. М., Холодов А. С. Сеточно-характеристические численные методы. М.: Наука. 1988.

[3] Favorskaya A.V., Breus A.V., Galitskii B.V. Application of the Grid-Characteristic Method to the Seismic Isolation Model: Proceedings of the Conference 50 Years of the Development of Grid-Characteristic Method // Smart Modeling for Engineering Systems. 2019. P. 167–181. DOI: 10.1007/978-3-030-06228-6_15

[4] Khokhlov N.I., Golubev V.I. On the Class of Compact Grid-Characteristic Schemes: Proceedings of the Conference 50 Years of the Development of Grid-Characteristic Method // Smart Modeling for Engineering Systems. 2019. P. 64–77. DOI: 10.1007/978-3-030-06228-6_7

[5] Фаворская А. В., Петров И. Б. Исследование особенностей трещиноватых зон путем полноволнового численного моделирования // Матем. моделирование. 2018. Т. 30, № 11. С. 105–126.

Стогний Полина Владимировна — аспирант
Московского физико-технического института
e-mail: stognii@phystech.edu;
Хохлов Николай Игоревич — к.ф.-м.н., зам.зав. лабораторией
прикладной вычислительной геофизики МФТИ (НИУ)
e-mail: k_h@inbox.ru;
Петров Игорь Борисович — д.ф.-м.н., заведующий кафедрой
информатики и вычислительной математики
Московского физико-технического института (НИУ);
e-mail: petrov@mipt.ru.
Дата поступления — 30 апреля 2019 г.

ТРЁХМЕРНОЕ МОДЕЛИРОВАНИЕ СИГНАЛОВ ЭЛЕКТРОМАГНИТНОГО ЗОНДА С ТОРОИДАЛЬНЫМИ КАТУШКАМИ

И. В. Суродина

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск
Институт нефтегазовой геологии и геофизики СО РАН,630090, Новосибирск

УДК 519.632.4+550.832.7
DOI: 10.24411/9999-016A-2019-10075

В работе рассматривается математическое моделирование электромагнитных откликов нового каротажного прибора в несимметричных изотропных средах, выполненное с помощью конечно-разностного метода. Описывается постановка задачи, методы решения. Приводятся численные эксперименты, показывающие необходимость центратора прибора для его использования на практике. Обсуждаются результаты моделирования.

Ключевые слова: электромагнитный каротаж, математическое моделирование, зонд с тороидальными катушками, конечные разности, система линейных алгебраических уравнений, итерационные методы, прямые методы

Введение

В последнее десятилетие произошли большие инженерные достижения в каротажном приборостроении. Появилась возможность создавать приборы, основанные на нетрадиционных для каротажа источниках возбуждения электромагнитного поля. В ИНГиГ и НПО «Луч» создан каротажный зонд нового поколения — устройство для регистрации характеристик электромагнитного поля с использованием тороидальных катушек [1]– [7]. В процессе создания прибора применялось математическое моделирование. В работе [8] проведено математическое обоснование нового зонда, описано решение прямой задачи электромагнитного каротажа для тороидального источника в цилиндрически-слоистой геоэлектрической модели. В работе [9] на основе двумерной программы для симметричных изотропных сред оценено влияние конечного металлического корпуса прибора на регистрируемые сигналы. Результаты моделирования были учтены при выборе оптимальной конфигурации зонда. Цель настоящего исследования состоит в том, чтобы правильно сформулировать задачу для несимметричных моделей, найти метод решения и оценить влияние эксцентриситета прибора на измеряемые характеристики.

1 Постановка задачи

Рассмотрим случай изотропной проводящей немагнитной среды. Для амплитуд электрического и магнитного поля справедливы уравнения Максвелла

$$\begin{cases} rot\vec{H} = \sigma\vec{E} \\ rot\vec{E} = i\omega\mu\vec{H} - \vec{j} \end{cases} \tag{1}$$

где $\sigma = 1/\rho$ — электропроводность, $\mu = 4\pi * 10^{-7}$ Гн/м — магнитная проницаемость, ω — циклическая частота, $\vec{j}$ — источник стороннего тока (плотность тока), $\vec{E}$ — электрическое поле, $\vec{H}$ — магнитное поле. Зависимость тока от времени гармоническая.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 19-05-00595).

ISBN 978-5-901548-42-4

Введём цилиндрическую систему координат $\{r, \phi, z\}$.

Источник поля в виде тороидальной катушки вдали от нее формально можно описать круговым магнитным током [10]. Поэтому считаем, что источник может быть задан в виде

$\vec{j} = \{0, j_\phi^\mu, 0\}$, где $j_\phi^\mu = -i\omega\mu M_\phi \delta(z - z_0)\delta(r - r_0)$, M_ϕ — магнитный момент, δ — дельта функция Дирака, r_0 — радиус кольца с током. Преобразуем уравнение (1) к виду

$$rot\rho\left(rot\vec{H}\right) - i\omega\mu\vec{H} = -j_\phi^\mu \tag{2}$$

Принимая во внимание уравнение Максвелла $div\vec{H} = 0$ и выделяя явно оператор Лапласа, приходим к эквивалентной зписи

$$\rho\Delta\vec{H} + grad\rho \times rot\vec{H} - i\omega\mu\vec{H} = -j_\phi^\mu \tag{3}$$

Запишем уравнение (3) в матрично-векторном виде, объединяя первые два оператора:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \\ H_z \end{pmatrix} = -j_\phi^\mu \tag{4}$$

Элементы матрицы определяются следующим образом:

$$a_{11} = -\frac{\partial}{\partial z}\left(\rho\frac{\partial}{\partial z}\right) - \frac{1}{r^2}\frac{\partial}{\partial\phi}\left(\rho\frac{\partial}{\partial\phi}\right) - \rho\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial(r)}{\partial r}\right),$$
$$a_{12} = \frac{1}{r^2}\frac{\partial\rho}{\partial\phi}\frac{\partial(r)}{\partial r} + \frac{2\rho}{r^2}\frac{\partial}{\partial\phi},$$
$$a_{13} = \frac{\partial\rho}{\partial z}\frac{\partial}{\partial r},$$
$$a_{21} = \frac{1}{r^2}\frac{\partial\rho}{\partial\phi}\frac{\partial(r)}{\partial r} - \frac{2\rho}{r^2}\frac{\partial}{\partial\phi},$$
$$a_{22} = -\frac{\partial}{\partial z}\left(\rho\frac{\partial}{\partial z}\right) - \frac{\rho}{r^2}\frac{\partial^2}{\partial\phi^2} - \frac{\partial}{\partial r}\left(\frac{\rho}{r}\frac{\partial(r)}{\partial r}\right),$$
$$a_{23} = \frac{1}{r}\frac{\partial\rho}{\partial z}\frac{\partial}{\partial\phi},$$
$$a_{31} = \frac{\partial\rho}{\partial r}\frac{\partial}{\partial r},$$
$$a_{32} = \frac{1}{r}\frac{\partial\rho}{\partial\phi}\frac{\partial}{\partial z},$$
$$a_{33} = -\rho\frac{\partial^2}{\partial z^2} - \frac{1}{r^2}\frac{\partial}{\partial\phi}\left(\rho\frac{\partial}{\partial\phi}\right) - \frac{1}{r}\frac{\partial(r)}{\partial r}\left(\rho\frac{\partial}{\partial r}\right).$$

В соответствии с условием излучения на бесконечности магнитное поле затухает с удалением от источника $|H| \to 0$ при $R, |Z| \to \infty$ Это позволяет приближенно поставить нулевые граничные условия для компонент H_r,H_ϕ,H_z, на большом расстоянии от источника. Таким образом, имеем задачу Дирихле для уравнения (4). Для построения конечно-разностной схемы для уравнения(4) источник расположим по кольцу $(z_0 = 0, r_0)$. Введем прямоугольную неравномерную координатную сетку по переменным r,z и равномерную по переменной ϕ:

$$\omega_h = \{(r_i, \phi_k, z_j), i = 0, ...N_r, k = 0, ..., N_\phi, z = -N_z, ..., N_z\}, r_{N_r} = R, z_{N_z} = Z \tag{5}$$

Нас интересует ограниченное решение при $r = 0$, которое удовлетворяет условию $\lim\limits_{r\to\,0}\dfrac{d\vec{H}}{dr} = 0$. Это условие наиболее удобно учесть путём специального выбора сетки – сдвигаем сетку по переменной r на полшага [11].

Рассмотрим на сетке (5) линейное конечномерное пространство сеточных функций H^0, равных нулю на границе и скалярным произведением

$$(U, V) = \sum_{i=0}^{N_r-1}\sum_{k=0}^{N_\phi-1}\sum_{j=1-N_z}^{N_z-1} U_{ikj}V_{ikj}r_i\hbar_i^{(r)}\hbar_k^{(\phi)}\hbar_j^{(z)}, \tag{6}$$

где $\hbar^{(r)}$ — средний шаг на неравномерной сетке по переменной r: $\hbar^{(r)} = \frac{h_i^{(r)} + h_{i+1}^{(r)}}{2}$. Средний шаг по переменной z определяется аналогично. U и V — сеточные функции. С помощью однородной консервативной конечно-разностной схемы [11], [12] в пространстве H^0 аппроксимируем уравнение (4), принимая во внимание условие периодичности по переменной ϕ и условие ограниченности решения. После дискретизации получим разностное уравнение вида $AX = F$. Матрица A является комплексной, не эрмитовой, не симметричной в пространстве H^0.

2 Тестирование и численные эксперименты.

Для решения полученной системы применялся итерационный метод бисопряжённых градиентов и прямой метод. Реализация прямого метода осуществлена с помощью программы PARDISO из библиотеки INTEL MKL. Первые эксперименты были проведены для цилиндрически-слоистых сред с беконечным корпусом прибора,а затем с учётом конкретной конфигурации прибора (рис. 1) для рабочих частот 50, 100, 250 кГц. Радиус медного корпуса прибора полагался равным 0.038м, радиус генераторных и приёмных катушек такой же. Расчёты проведены для одной генераторной катушки, расположенной условно в плоскости $z = 0$. Приёмные катушки располагались в плоскостях $z = 0.25, 0.5, 0.75$ м.



Рис. 1: Модель прибора

Протестированы и выбраны сетки для частот 50, 100, 250 кГц, дающие наименьшую погрешность на достаточно большом диапазоне моделей (от 0.1 Ом м до 100 Ом м). При расчётах контролировались по точности вычислений 2 измеряемые величины H_ϕ и E_z (J_z). Результаты расчётов сравнивались с аналогичными расчётами по двумерной программе [9] для реального зонда и с результатами расчётов по программе для цилиндрически-слоистых сред [8] для бесконечно-длинного зонда.

Компонента E_z рассчитывалась при помощи первого уравнения (1). Особенности вычисления данной компоненты состоят в том, что необходима очень мелкая сетка по переменной r (для достижения необходимой точности) из-за большого градиента в значениях в приповерхностной части прибора. Скин-слой для медного проводника для частоты 100 кГц составляет всего 0.21мм. На рисунках 2, 3 представлены графики функций действительной и мнимой частей E_z, вычисленные для цилиндрически-слоистой среды [8] для источника с частотой 250 кГц на малом отрезке длиной в 1 мм. Как видно из рисунка 3, на отрезке от 0.038 м до 0.03799 м вещественная часть компоненты меняется -2132 до -3065. Если вычислить $Re(Ez)$ в точке 0.03799, то будем иметь погрешность в 40%. Очевидно, что шаг сетки здесь должен быть менее чем 0.00001м. Следует отметить, что компонента H_ϕ (в двумерной и трёхмерной программах) вычисляется достаточно хорошо и на грубой сетке, но с обязательным присутствием точки в сетке, где производятся измерения.

Как показали расчёты, для хорошего вычисления компоненты E_z необходим шаг 1.d-6 и меньше по переменной r вблизи поверхности корпуса прибора. На рисунке 4 приведены графики действительных частей компоненты E_z, рассчитанной для зонда с частотой 250 кГц (рис. 1), располагающегося в пласте с сопротивлением 100 Ом м. По оси ординат отложена длина корпуса прибора (условно от -1 м до 1 м), по оси абсцисс — рассчитанные значения компоненты по двумерной и трёхмерной программам. Источник расположен в точке — 0.5 м по оси X. Из рисунка 4 видно, что компонента рассчитывается достаточно точно в измеряемом прибором диапазоне. Рассмотрим следующую модель. Пусть вышеописанный зонд располагается в скважине радиусом 0.108 метра, имеющей электрическое сопротивление 1 Ом м. Сопротивление пласта, вмещающего скважину – 100 Ом м. На рисунках 5, 6 приведены графики рассчитанных (веществен-

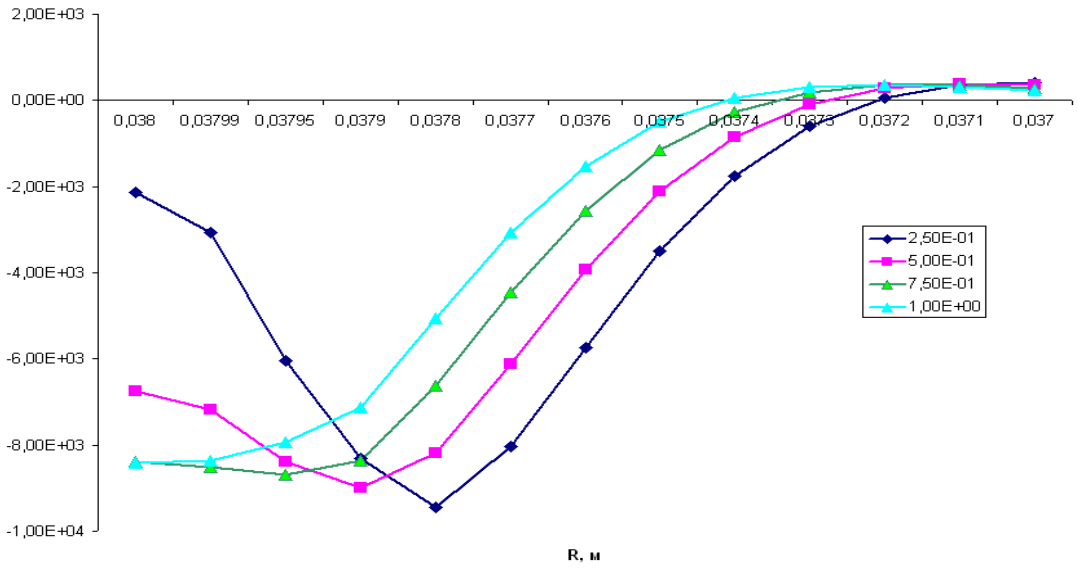


Рис. 2: Вещественная часть E_z, рассчитанная по программе для цилиндрически-слоистой среды

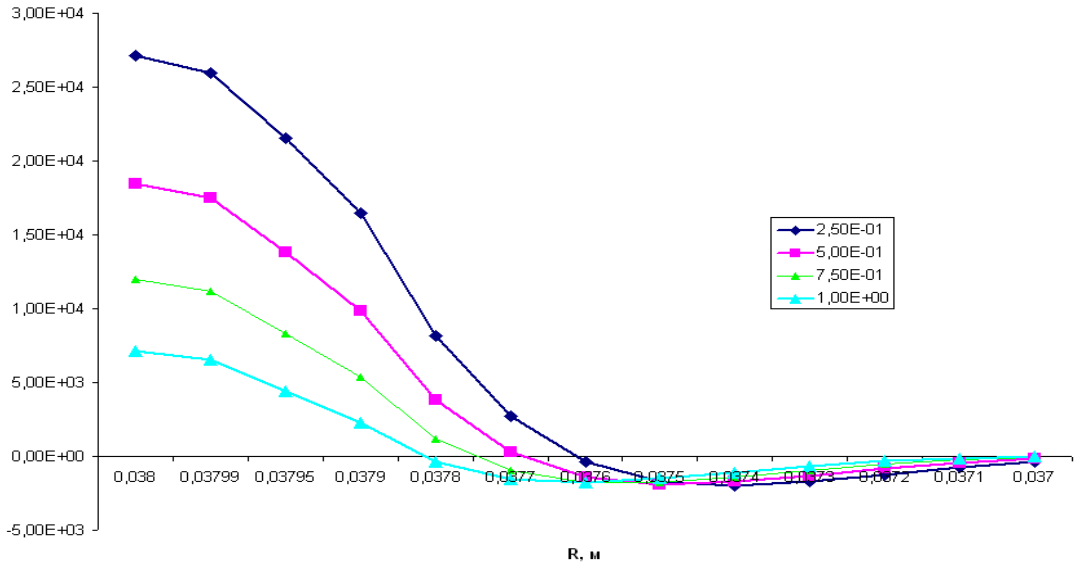


Рис. 3: Мнимая часть E_z, рассчитанная по программе для цилиндрически-слоистой среды

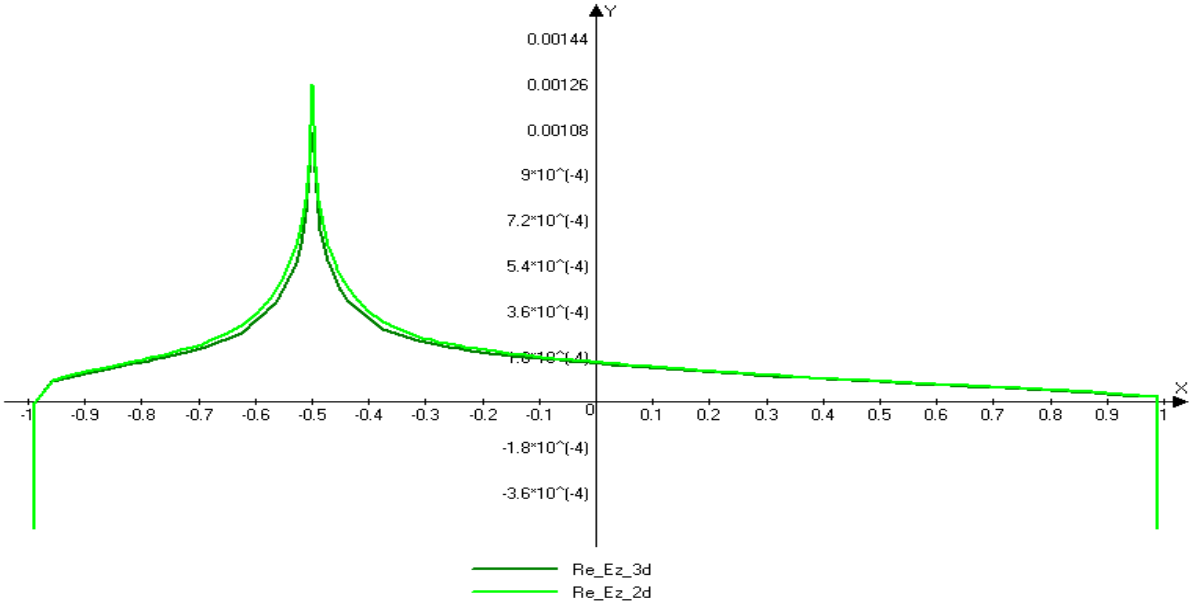


Рис. 4: Вещественная часть E_z, рассчитанная по двумерной и трёхмерной программам

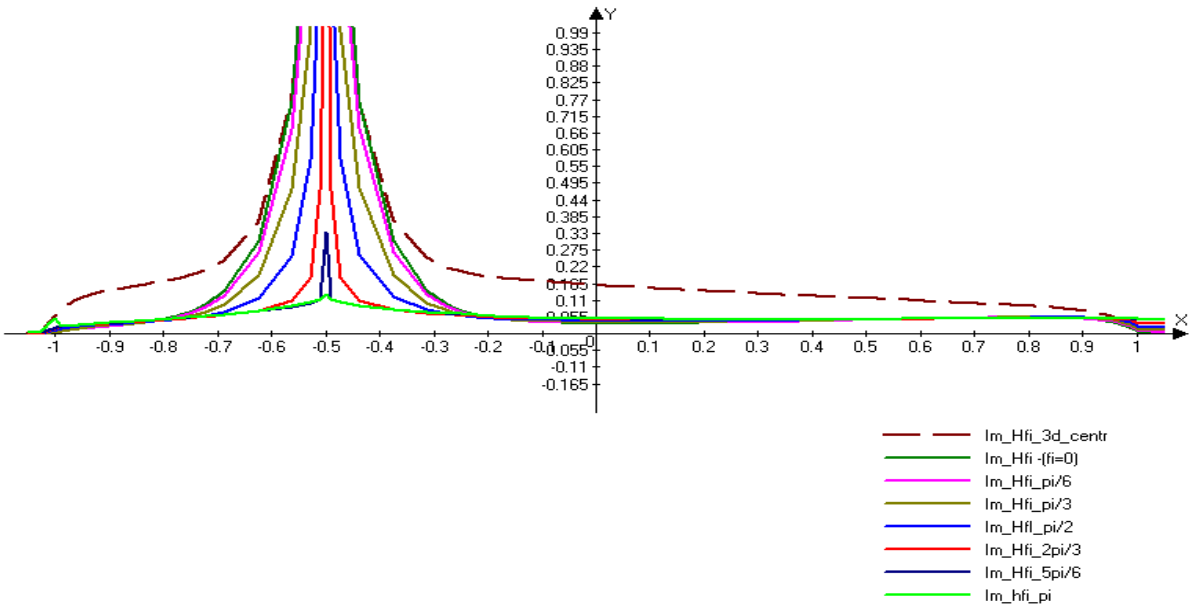


Рис. 5: Мнимая часть H_ϕ, рассчитанная для зонда, находящегося в центре скважины и сдвинутого на её стенку

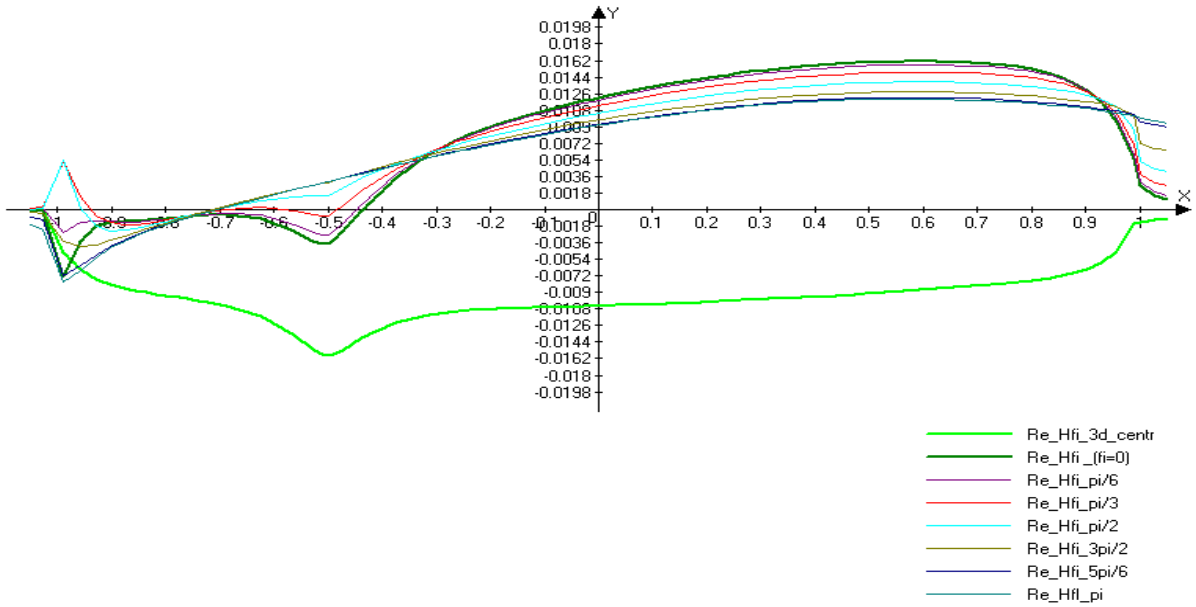


Рис. 6: Вещественная часть H_ϕ, рассчитанная для зонда, находящегося в центре скважины и сдвинутого на её стенку

ной и мнимой частей) H_ϕ для зонда, находящегося в центре скважины и для зонда, сдвинутого на её стенку. Значения H_ϕ приведены для различных углов $\phi = 0, \pi/6, \pi/3, 2\pi/3, 5\pi/6, \pi$.

Во всём измеряемом диапазоне значения компонент у сдвинутого зонда очень сильно отличаются от рассчитанных компонент для зонда, находящегося в центре скважины. Вещественная часть H_ϕ даже имеет другой знак. Заметным при сдвиге становится и влияние изолирующей проставки (не в измеряемом диапазоне), что выражается на графиках явным скачком в точке — 1 м, особенно это чётко видно для вещественной части H_ϕ (6). Аналогичная картина наблюдается и для компоненты электрического поля E_z. Проведённые расчёты показывают, что необходимо в конструкции прибора предусмотреть центратор, который будет удерживать зонд в центре скважины, иначе интерпретация сигналов будет весьма затруднительной.

Оба метода (и итерационный, и прямой) при расчёте модели с центральным положением зонда относительно центра скважины привели к практически одинаковым результатам. При этом итерационный метод существенно выигрывал по времени. Но как только перешли к трёхмерным моделям (сдвиг зонда относительно центра скважины) — итерационный метод перестал сходиться для большинства моделей. Прямой метод, хотя и требует больших ресурсов по памяти, позволяет получить решение с хорошей точностью.

Расчёты проведены на кластере НКС-1П ЦКП ССКЦ СО РАН.

Заключение

Создана программа прямого трёхмерного моделирования для зонда с тороидальными катушками. Проведены тестовые расчёты на правильность работы программы. Показана зависимость показаний зонда от эксцентриситета прибора.

Благодарности

Автор выражает признательность коллегам по лаборатории скважинной геофизики Института нефтегазовой геологии и геофизики им. А.А. Трофимука СО РАН за содержательное обсуждение постановки задачи и благодарит В.Н. Глинских за постоянные консультации, М.Н. Никитенко за предоставленную программу расчёта тороидального зонда в слоистой среде, М.И. Эпова за полезные советы и замечания.

Список литературы

[1] Эпов М.И., Глинских В.Н., Никитенко М.Н. Способ измерения удельной электропроводности и электрической макроанизотропии горных пород. // Патент на изобретение № 2525149. Опубл. 10.08.2014. Бюл. № 22.

[2] Эпов М.И., Еремин В.Н., Манштейн А.К., Петров А.Н., Глинских В.Н. Устройство для измерения удельной электропроводности и электрической макроанизотропии горных пород. // Патент на изобретение № 2528276. Опубл. 10.09.2014. Бюл. № 25

[3] Эпов М.И., Еремин В.Н., Петров А.Н., Глинских В.Н., Суродина И.В., Киселев В.В.,Никитенко М.Н. Устройство для регистрации характеристик электромагнитного поля с использованием тороидальных катушек // Патент на изобретение № 2578774. Заявл. 14.01.2015; опубл.27.03.2016. Бюл. № 9.

[4] Эпов М.И., Еремин В.Н., Петров А.Н., Глинских В.Н., Суродина И.В., Киселев В.В. Устройство для генерации электромагнитного поля тороидальной катушкой в геологической среде // Патент на изобретение № 2579177. Заявл. 14.01.2015 г. Опубл. 10.04.2016. Бюл. № 10

[5] Эпов М.И., Еремин В.Н., Петров А.Н., Глинских В.Н. Электромагнитный зонд для каротажа в нефтегазовых скважинах. // Патент на изобретение № 2583867. Опубл. 10.05.2016. Бюл. № 13

[6] Эпов М.И., Глинских В.Н., Еремин В.Н., Никитенко М.Н., Петров А.Н., Суродина И.В., Михайлов И.В. Математическое и физическое моделирование сигналов электромагнитного зонда для изучения макроанизотропии осадочных отложений [Электронный ресурс] // Геомодель 2017: 19-я конференция по вопросам геологоразведки и разработки месторождений нефти и газа – Геленджик, 11-14 сентября 2017 г. Тезисы докладов. (43809)

[7] Epov M.I., Glinskikh V.N., Eremin V.N., Nikitenko M.N., Petrov A.N., Mikhaylov I.V. Electromagnetic tool for high-resolution logging: theoretical and experimental studies [Электронный ресурс] // SPE Russian Petroleum Technology Conference. Moscow, October 16-18, 2017. Paper SPE-187904-MS

[8] Эпов М.И., Никитенко М. Н., Глинских В. Н. Математическое обоснование нового электромагнитного зонда с тороидальными катушками для высокоразрешающего каротажа нефтегазовых скважин // Вестн. НГУ. Серия: Информационные технологии. 2018. Т. 16, № 1. С. 113–129.

[9] Суродина И.В. Математическое моделирование сигналов электромагнитного зонда с тороидальными катушками в двумерных изотропных моделях геологических сред. Интерэкспо Геосибирь-2018. Т. 2. № 4. С.162–170

[10] Светов Б. С. Электродинамические основы квазистационарной геоэлектрики. М.: Изд-во ИЗМИРАН, 1984. 183 с.

[11] Самарский А.А., Андреев В.Б. Разностные методы для эллиптических уравнений – М.: Наука, 1976. — 352 с.

[12] Самарский А.А, Николаев Е.С. Методы решения сеточных уравнений, М.: Наука,1978, 592с.

Суродина Ирина Владимировна — к.ф.-м.н., ст. науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: sur@ommfao1.sscc.ru.
Дата поступления — 5 июня 2019 г.

“БУДУЩЕЕ ЗЕМЛИ”: ГЛОБАЛЬНЫЕ ПРОБЛЕМЫ И РАДИАЦИОННОЕ ПОЛЕ, РАДИАЦИОННЫЙ ФОРСИНГ И ДИСТАНЦИОННОЕ ЗОНДИРОВАНИЕ ЗЕМЛИ

Т. А. Сушкевич, С. А. Стрелков, С. В. Максакова

ФИЦ "Институт прикладной математики им. М.В. Келдыша Российской академии наук", 125047, Москва

УДК 52-6; 52-64; 52-65
DOI: 10.24411/9999-016A-2019-10076

Фундаментальные основы для реализации Всемирной научной ПРОГРАММЫ "Будущее Земли" (World Global Research Projects "Future Earth") и исследования Климатической системы Земли под влиянием радиационного форсинга заложены в XX-м веке благодаря изобретению компьютера и выхода человека в космос и, конечно, достижений в математике. Сложнейшие задачи эволюции, климата, экологии, глобального мониторинга и дистанционного зондирования Земли (ДЗЗ) с гиперспектральными подходами и нанодиагностикой природной среды и объектов предлагается рассматривать как сопряженные. Электромагнитное излучение — единое физическое поле, объединяющее радиационное поле Земли с радиационно–активными компонентами. Гиперспектральное дистанционное зондирование Земли — важнейшая технология оценки воздействия естественно-природных и антропогенных факторов на планету Земля.

Ключевые слова: Будущее Земли, радиационное поле, световые технологии.

Введение. Математика, компьютер, космос и "Будущее Земли"

В 2014 г. началось формирование Всемирной научной ПРОГРАММЫ "Будущее Земли" (World Global Research Projects "Future Earth") [1] для координации международных исследований по устойчивому развитию окружающей среды и общества по совместной инициативе Международного совета по науке (ICSU) и Международного научного совета по общественным наукам (ISSC) при поддержке крупнейших международных правительственных организаций в сфере науки, образования и культуры — ЮНЕСКО, Программы Объединенных Наций по окружающей среде (ЮНЕП), Международного университета ООН, Международной метеорологической организации (WM0), а также Бельмонтским форумом и Сетью организаций по устойчивому развитию (SDSP). В ГЛОБАЛЬНУЮ ПРОГРАММУ "Будущее Земли" вошло большинство всех ранее существовавших международных программ по изучению глобальных изменений — Международная геосферно–биосферная программа (IGBP), Международная программа по общественному измерению глобальных изменений (IHDP), Всемирная программа по изучению климата (WCRP) и Международная программа по изучению биоразнообразия (Diversitas) и др., которые ориентированы на реализацию "Парижского соглашения по климату" [2].

В 2018 г. к ПРОГРАММЕ присоединилась Россия в лице Российской академии наук: Постановлением Президиума РАН 27 июня 2018 г. создан Российский национальный Комитет РАН по международной программе "Будущее Земли" [3]. Эта ПРОГРАММА — глобальный вызов "Повестки XXI–го века" [4] — не имеет аналогов по масштабам и значимости для человечества в мировой истории.

В мире нарастают процессы глобализации и человечество все серьезнее сталкивается с проблемами изменения климата, истощения ресурсов, перенаселения и последствий технологической революции. Фундаментальные основы для реализации ПРОГРАММЫ заложены в XX–м веке благодаря изобретению компьютера и выхода человека в космос и, конечно, достижений в математике и науках о Земле.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проектов 18-01-00609, 17-01-00220).

ISBN 978-5-901548-42-4

ФУНДАМЕНТАЛЬНАЯ НАУКА и "РАКЕТНО-ЯДЕРНЫЙ ЩИТ" — двигатели научно-технического прогресса в XX–м веке! Двадцатый век в истории земной цивилизации — это век научно–технической революции (НТР), связанной с ТРЕМЯ ВЕЛИКИМИ ОТКРЫТИЯМИ:
— проникновение в тайны и овладение ядерной энергией,
— покорение космического пространства и выход человека в космос,
— изобретение электронно–вычислительных машин (ЭВМ) и создание информационных технологий.

Компьютер явился главным действующим лицом, основным двигателем НТР: использование ядерной энергии, полет в космос, информационные технологии были бы невозможны без ЭВМ.

Эти великие открытия и международное сотрудничество в космосе [5], [6] связаны с именем математика–легенды Мстислава Всеволодовича Келдыша (10.02.1911 – 24.06.1968) [7], масштабы интеллектуальной, научной и организационной деятельности которого потрясают воображение. Это русский гений — "Великий ум России" — "Ломоносов XX–го века", заложивший фундаментальные основы цивилизации и постиндустриального информационного общества XXI–го века, а также "цифровой экономики" и современных глобальных проблем по спасению планеты Земля [8]. Это был лучший Президент за всю историю Академии Наук, который в 1964 г. спас Академию Наук СССР от волюнтаризма Н. С. Хрущева, угрожавшего на июльском Пленуме ЦК КПСС "разогнать академию наук....", и при котором Академия Наук стала форпостом СССР в мире, существенно расширилось международное сотрудничество в науке, сложились ведущие в мире научные школы, советские ученые получали Нобелевские и Ленинские премии.

АКАДЕМИЧЕСКИЙ Институт прикладной математики — "Институт Келдыша", основанный в 1953 г., где 65 лет назад в 1954 г. ввели в строй ПЕРВУЮ ламповую серийную ЭВМ "Стрела", — активный участник научно-технической революции и реализации "Ракетно–ядерного щита".

МАТЕМАТИКА в современном мире была, есть и будет всегда ЦАРИЦЕЙ ВСЕХ НАУК! Галилео Галилей (15.02.1564 – 08.01.1642): "Природа формулирует свои законы языком математики". Гаусс Карл Фридрих (30.04.1777 – 23.02.1855): "Математика — царица наук, а арифметика — царица математики". ПЕРВЫЕ расчеты для атомного проекта проводили на арифмометрах. Без МАТЕМАТИКИ и высочайшего уровня МАТЕМАТИКОВ не было бы успехов ни в космосе ни в атомном проекте!

Не случайно МАТЕМАТИК М. В. Келдыш — Главный Теоретик Космонавтики! Благодаря М. В. Келдышу мы стали ПЕРВЫМИ в космосе: ПЕРВЫЙ СПУТНИК в 1957 г., ПЕРВЫЙ полет космонавта в 1961 г., три ПЕРВЫХ корабля к Луне запущены 60 лет назад в 1959 г., ПЕРВЫЕ Долгосрочные орбитальные станции в 1971 г., ПЕРВАЯ МКС в 1989 г., ПЕРВАЯ стыковка кораблей и ПЕРВАЯ стыковка с ДОС и МКС!

Талантливые люди талантливы во всем... Такими были покорители космоса. Подтверждение этих слов — те же Ломоносов, Пушкин, Менделеев, Лермонтов, Келдыш... М. Ю. Лермонтов (15.10.1814 – 27.07.1841), поэт и математик (юный поэт владел математикой значительно лучше большинства своих знакомых): "Математику уже затем учить надо, что она ум в порядок приводит". А. С. Пушкин (06.06.1799 – 10.02.1837): "Вдохновение нужно в геометрии, как и в поэзии".

Настоящая публикация посвящена великим советским и российским ученым — основоположникам и создателям фундаментальных основ математики и наук о Земле, научное наследие которых должно развиваться в ПРОГРАММЕ "Будущее Земли", — это академики:

- Владимир Иванович Вернадский (12.03.1863 – 06.01.1945) [9],
- Отто Юльевич Шмидт (30.01.1891 – 07.09.1956),
- Андрей Николаевич Тихонов (30.10.1906 – 08.10.1993),
- Гурий Иванович Марчук (08.06.1925 – 24.03.2013) [10],
- Виктор Амазаспович Амбарцумян (18.09.1908 – 12.08.1996),
- Виктор Викторович Соболев (02.09.1915 – 07.01.1999),
- Кирилл Яковлевич Кондратьев (14.06.1920 – 01.05.2006),
- Александр Михайлович Обухов (05.05.1918 – 03.12.1989),
- Никита Николаевич Моисеев (23.08.1917 – 29.02.2000) [11],
- Михаил Иванович Будыко (20.01.1920 – 10.12.2001) [12],
- Владимир Евсеевич Зуев (29.01.1925 – 06.06.2003),
- Андрей Сергеевич Монин (02.07.1921 – 22.09.2007),
- Николай Павлович Лаверов (12.01.1930 – 27.11.2016),

- Юрий Антониевич Израэль (15.05.1930 – 23.01.2014) [13] — в 2007 г. Нобелевский лауреат премии мира вместе с Альбертом Гором и др.,

  а также
- член-корреспондент Илья Афанасьевич Кибель (19.10.1904 – 05.09.1970),
- член-корреспондент Геннадий Павлович Курбаткин (23.07.1930 – 09.09.2011),

  профессора
- Александр Александрович Фридман (16.06.1988 – 16.09.1925) [14],
- Кусиэль Соломонович Шифрин (26.07.1918 – 02.06.2011),
- Михаил Арамаисович Петросянц (04.12.1919 – 18.07.2005),
- Ева Михайловна Фейгельсон (07.09.1915 – 17.01.2010)

  и профессора
- Георгий Владимирович Розенберг (29.04.1914 –09.12.1982),
- Матвей Семенович Малкевич (1923-2010),
- Неон Александрович Арманд (01.01.1932 – 09.10.2009)

  — ПИОНЕРЫ покорения космического пространства и создания основ космических исследований, которые стояли у истоков создания современной космической оптики для аэрокосмического ДЗЗ.

Значимый вклад в теорию климата внесли академики Г. С. Голицын, И. И. Мохов, В. П. Дымников, а также профессора А. Н. Филатов, И. Л. Кароль, В. М. Катцов, В. П. Мелешко, А. В. Фролов и другие.

Особого внимания заслуживает профессор Евграф Сергеевич Кузнецов (13.03.1901 – 17.02.1966) — основатель отечественной научной школы по теории переноса излучения, нейтронов, заряженных частиц и создатель уникального и единственного в мире отдела "Кинетические уравнения" в Институте Келдыша. В США был Субрахманьян Чандрасекар (Subrahmanyan Chandrasekhar; 19.10.1910 – 21.08.1995), получивший Нобелевскую премию 10 декабря 1983 г. за "черные дыры" [15]. А в СССР жил и работал Е. С. Кузнецов — оба участники "атомных проектов" и во время войны помогали... В связи со 100–летием Е. С. Кузнецова при поддержке РФФИ в 2003 г. изданы научные труды [16], в которые вошли публикации с 1925 по 1966 гг.

"Будущее Земли" зависит и от космоса, тем важнее работы физика–теоретика и астрофизика С. Чандрасекара, который работал с 1937 по 1995 гг. в Чикагском университете, т.е. 58 лет в одном месте (Т. А. Сушкевич в Институте Келдыша тоже 58 лет!), и при этом сменил СЕМЬ направлений исследований, результаты которых подытоживал и систематизировал в монографиях для потомков исследователей. В 2018 г. исполнилось 65 лет изданию Е. С. Кузнецовым в 1953 г. монографии С. Чандрасекара "Перенос лучистой энергии" — перевода на русский язык монографии "Radiative Transfer", в США изданной в 1950 г. [17]. Это ПЕРВАЯ в мировой науке монография по теории переноса излучения и лучистой энергии, в которой системно и четко сформулированы наиболее важные задачи и математические модели (одномерные плоские и сферические) и изложены методы их решения, полезные для тех, кто моделированием и расчетами занимается, и потому на этой монографии учились несколько поколений молодых исследователей не только в мире, но и в СССР и в России. Важно знать и помнить... Однако, советские и российские ученые в те времена мало издавали монографий по радиационным полям и первыми были монографии К. Я. Кондратьева [18], [19], но внесли огромный научный фундаментальный вклад в эту область знаний, имеющей широкую сферу приложений и вечной, пока светит Солнце и мерцают звезды...

В 2002 г. за достижения в ДЗЗ была присуждена премия Правительства Российской Федерации в области науки и техники за работу "Разработка и внедрение методов и технологий аэрокосмического мониторинга природной среды" коллективу: Макриденко Леонид Алексеевич, Полищук Георгий Максимович, Волков Алексей Михайлович (посмертно), Исаев Александр Сергеевич, Коровин Георгий Николаевич, Сухих Василий Иванович, Козодеров Владимир Васильевич, Ушаков Сергей Александрович, Бондур Валерий Григорьевич, Малинников Василий Александрович, Савиных Виктор Петрович, Викторов Алексей Сергеевич, Смоктий Олег Иванович, Сушкевич Тамара Алексеевна [20], [21].

К проектам покорения космоса были привлечены лучшие советские специалисты в разных областях. Настоящая публикация — это признание заслуг первопроходцев в освоении космического пространства и космических технологий ДЗЗ, кто своими трудами развивал фундаментальные основы ПРОГРАММЫ:
— это космонавты-исследователи на ПЕРВЫХ пилотируемых космических кораблях и долгосрочных орбитальных станциях [22], [23]: Ю. А. Гагарин, Г. С. Титов, А. Г. Николаев, П. Р. Попович, В. В. Терешкова, В. Ф. Быковский, А. В. Филипченко, Н. Н. Рукавишников, А. А. Леонов, В. Н. Кубасов, В. А. Шаталов, А. С.

Елисеев, В. И. Севастьянов, В. Г. Лазарев, О. Г. Макаров, П. И. Климук, Б. В. Волынов, Е. Б. Хрунов, В. Н. Волков, В. С. Комаров, В. В. Горбатко, В. В. Аксенов, В. В. Ковалёнок, В. В. Рюмин, Г. Т. Береговой..., — это космонавты-исследователи, которые защитили докторские диссертации по материалам ДЗЗ из космоса — Г. М. Гречко, член-корреспонденты РАН В. П. Савиных и В. В. Лебедев;
— это советские ученые, внесшие значимый вклад в становление космических исследований и ДЗЗ: Г. А. Михайлов, О. И. Смоктий, В. В. Козодеров, А. И. Лазарев, М. М. Мирошников, Е. О. Федорова, В. П. Козлов, А. С. Селиванов, В. Н. Сергеевич, И. И. Кокшаров, Л. И. Чапурский, Т. А. Гермогенова, М. В. Масленников, М. С. Малкевич, Г. В. Розенберг, А. Б. Сандомирский, Г. И. Горчаков, А. Х. Шукуров, И. Н. Минин, А. А. Бузников, А. П. Гальцев, О. Б. Васильев, Ю. М. Тимофеев, О. М. Покровский, Л. С. Ивлев, Б. С. Непорент, М. С. Киселева, Э. Г. Яновицкий, В. М. Орлов, В. Г. Бондур, Н. И. Арженеко, А. П. Тищенко, Ч. Й. Виллман, О. А. Авасте, В. Н. Досов, В. В. Филюшкин, М. А. Назаралиев, М. В. Кабанов, С. Д. Творогов, Г. Г. Матвиенко, Ю. С. Макушкин, Г. М. Креков, В. М. Фомин, Ю. Н. Пономарев, В. П. Лукин, В. В. Белов, В. А. Крутиков, И. В. Самохвалов, М. В. Панченко, А. Г. Боровой, А. М. Волков, Л. А. Пахомов, А. А. Феоктистов, Д. А. Усиков, В. Г. Золотухин, А. К. Городецкий, В. В. Бадаев, Я. Л. Зиман, Г. А. Аванесов, У. М. Султангазин, Г. Ш. Лившиц, В. Е. Павлов, В. Л. Филиппов. Н. И. Москаленко, В. Н. Арефьев, А. М. Броунштейн и др.

В 1965 г. был организован Институт космических исследований АН СССР, в который перешли подразделения из Института Келдыша и других организаций. Организатором и первым директором (1965-1973) ИКИ АН СССР по рекомендации Президента АН СССР М. В. Келдыша являлся его соратник академик Георгий Иванович Петров (31.05.1912 – 13.05.1987). Особо следует отметить ученых из Белоруссии: основатель Института физики Борис Иванович Степанов (28.04.1913 – 07.12.1987), экс–ректор Минского госуниверситета Леонид Иванович Киселевский (12.04.1927 – 08.10.1991), Ф. И. Федоров, А. П. Иванов, К. С. Адзерихо, В. Е. Плюта, К. Г. Предко, Э. П. Зеге, И. Л. Кацев, А. П. Пришивалко, П. Я. Ганич, Л. И. Чайковская, А. Б. Гаврилович, П. Я. Бойко, С. А. Макаревич, Б. И. Беляев, С. Б. Костюкевич, А. А. Ковалев, С. И. Кононович, Е. К. Науменко, А .Н. Валентюк, Ю. А. Лебединский и др.

Радиационное поле Земли и радиационный форсинг

Изменения климата, капризы погоды воспринимались человечеством как данность, вынуждающая жителей планеты к этому приспосабливаться. И только сравнительно недавно, во многом благодаря прорывам в космических системах ДЗЗ и "computer sciences" появились возможности всесторонне изучать процессы формирования погоды и климата, причины их изменений и перспективы влияния на них антропогенной деятельности и естественно-природных факторов. Электромагнитное излучение — единое физическое поле, объединяющее экологию, климат, эволюцию, мониторинг и дистанционное зондирование Земли — радиационное поле Земли с радиационно–активными компонентами [24], [25].

Климатическая система Земли (КСЗ) — это природная среда, включающая атмосферу, гидросферу (океаны, моря, озера, реки), криосферу (поверхность суши, снег, морской и горный лед и т.д.), биосферу, объединяющую всё живое. Для количественных оценок значимости разных климатообразующих факторов, зависящих от солнечного и собственного излучения, ввели специальную характеристику КСЗ — радиационное воздействие (форсинг) [26]. По экспертным оценкам последнего времени от 40% до 60% приходится на радиационный форсинг на эволюцию климата. Удалите Солнце и будущее Земли очевидно...

Радиационный форсинг — это изменение притока радиации (солнечной коротковолновой и длинноволновой) в глобальной системе "атмосфера – земная поверхность – океан" под влиянием радиационно–активных факторов: альбедо земной поверхности, облачность, океаны и моря, снежный и ледовый покров, газовый состав атмосферы, аэрозольный состав атмосферы, солнечная постоянная, спектральные характеристики рассеяния и поглощения, изотропная и анизотропная (при осадках и низких температурах) среда, оптическая и метеорологическая "погода" (температура, давление, влажность), биофизические, биогеофизические и биогеохимические процессы, круговорот веществ в биосфере и экосистеме, нефтегазовый комплекс и т.д.

Гиперспектральное дистанционное зондирование [27] радиационно–активных компонент и метеорологических параметров атмосферы, поверхности и океана — актуальная перспективная сложнейшая задача будущего, без решения которой все модели климата будут недостоверными. Объективно оценивать и контролировать выбросы газов и загрязнения окружающей среды сложно, поскольку КСЗ — это нелинейная динамическая система и локальные выбросы тут же распространяются в воздушной и водной средах в зависимости от метеорологической ситуации и взаимодействуют с биосферой. Например, извержение вулканов и трансграничный перенос загрязнений, которые влияют на экологию и состояние окружающей среды, могут

быть обнаружены методами ДЗЗ, а далее через перенос лучистой энергии, зависящей от загрязнителей природных сред, может влиять на климат и в конечном итоге на тренд эволюции Земли как планеты. Наглядной иллюстрацией радиационного форсинга является "ядерная зима" после ядерного удара [28].

Непреодолимая сложность проблемы состоит в том, что для исследований планеты не допустимы натуральные эксперименты и возможны только мониторинг и наблюдения разными средствами, с одной стороны, а с другой стороны, на момент измерений радиации невозможно восстановить весь набор оптико–геофизических и оптико–метеорологических параметров системы "атмосфера – суша – океан", от которых зависит радиация, и невозможно повторить условия наблюдений, так как среда непрерывно изменяется и никогда не повторяется. И только междисциплинарные исследования и математическое моделирование "больших" прямых и обратных задач теории переноса излучения с параллельным суперкомпьютингом позволят провести теоретико–расчетные исследования столь сложных проблем и получить качественные и количественные оценки для анализа и прогнозов, а также для разных тематических приложений на основе "сценарного" подхода и широкого использования компьютерного моделирования при контролируемых входных данных модели через коэффициенты уравнений, граничные условия, функции источников излучения.

Заключение

65 лет назад, 14 февраля 1954 г., в кабинете М. В. Келдыша (с 1981 г. Мемориальный музей–кабинет академика М. В. Келдыша РАН) состоялось ПЕРВОЕ совещание, на котором ВПЕРВЫЕ обсуждался вопрос о возможности создания и запуска в космическое пространство ПЕРВОГО искусственного спутника Земли. В совещании участвовали ученики М. В. Келдыша кандидаты физ.-мат. наук Т. М. Энеев и Д. Е. Охоцимский, ставшие академиками за достижения в космосе. Присутствовали С. П. Королев, П. Л. Капица, И. А. Кибель, М. К. Тихонравов, А. Ю. Ишлинский, С. Н. Вернов и целый ряд других людей [8], [29]. Это были те ученые, конструкторы и инженеры — ведущие профессионалы и лучшие специалисты, кто был непосредственно связан с созданием космической техники, и те, кто мог высказать предложения по научным исследованиям, которые нужно было бы проводить со спутников.

О программе космических исследований заговорили в 1955 г. По указанию М. В. Келдыша в 1955 г. из Академии наук СССР с помощью референта Г. А. Скуридина разослали письма в разные организации и ученым разных специальностей с одним вопросом "Как можно использовать космос?" Для убеждения руководителей СССР в необходимости освоения космического пространства и запусков космических спутников и кораблей М. В. Келдыш как уже признанный государственный деятель выделил две главные задачи: разведка и наблюдения Земли, вокруг которых сформировались многие научно-исследовательские проекты, определилась новая отрасль человеческой деятельности, в 1955 г. началось строительство космодрома "Байконур" и были созданы Министерство общего машиностроения и другие ведомства.

Подтверждается стратегический выбор, сделанный Главным Теоретиком Космонавтики, единственным математиком трижды Героем Социалистического Труда М. В. Келдышем в 1955 г. и актуальный в XX и XXI веках. Юрий Гагарин (09.03.1934 – 27.03.1968) — ПЕРВЫЙ в истории человечества увидел планету Земля из космоса и воскликнул "Земля голубая!", а позже добавил "Земля такая маленькая...". По инициативе академика В. А. Садовничего, отметившего 3 апреля 2019 г. свое 80–летие и который много сделал для создания и поддержки космических исследований в МГУ имени М. В. Ломоносова, начало формироваться новое научное и научно-практическое направление "Космическое землеведение" [30], о котором Г .И. Марчук и К. Я. Кондратьев писали в 1985 г. [31], а также о "Спутниковой климатологии" [32].

Т. А. Сушкевич, редактор трудов [16], — последняя ученица Е. С. Кузнецова, который в 1959 г. предложил изучить монографию С. Чандрасекара [17] и тем самым определил направление научной деятельности на всю жизнь. Настоящая публикация подготовлена с позиции 60–летнего опыта работы [33], [34] (в монографии содержится более 400 ссылок на публикации автора) и 55–летия первой статьи [35] в области теории переноса излучения, 56–летия участия в космических проектах [20], [21] и 50–летия публикации [36], которая потрясла весь космический мир: никто в мире до сих пор не превзошел "сферическую модель" радиационного поля Земли в масштабах всей планеты [34]. Писать было легко, поскольку Т. А. Сушкевич знала не только работы всех выше упомянутых, но и была лично знакома со всеми, кроме Вернадского и Фридмана.

В истории и памяти потомков останутся имена наших великих УЧЕНЫХ, которые в XX–м веке на государственном уровне возглавили проекты и своими трудами внесли огромный вклад в становление наук о Земле и исследования Земли из космоса — это В. И. Вернадский, его продолжатель Н. Н. Моисеев, Г. И. Марчук, К. Я. Кондратьев, Ю. А. Израэль. Кто возглавит ГЛОБАЛЬНУЮ ПРОГРАММУ "Будущее Земли", тот возьмет на себя ответственность за "сбережение планеты", как Н. Н. Моисеев [37].

Список литературы

[1] Future Earth. Global Research Projects. http://futureearth.org

[2] Парижское соглашение согласно Рамочной конвенции Организации Объединенных Наций об изменении климата ("Парижское соглашение"). ООН. 2016. 19 с.

[3] Постановление Президиума РАН № 103 от 29.05.2018 О создании Комитета РАН по международной программе "Будущее Земли".

[4] Коптюг В. А. Конференция ООН по окружающей среде и развитию (Рио-де-Жанейро, июнь 1992 года). Информационный отчет. Новосибирск: Сибирское отделение Российской академии наук , 1992. 79 с.

[5] Сушкевич Т. А. Главный Теоретик М. В. Келдыш и Главный Конструктор космонавтики С. П. Королев — покорители космоса // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8, № 1. С. 9–25.

[6] Сушкевич Т. А. Главный Теоретик М. В. Келдыш — организатор международного сотрудничества в космосе и первой советско–американской Программы "Союз-Аполлон" (ЭПАС) // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8, № 4. С. 9–22.

[7] Келдыш М. В. Творческий портрет по воспоминаниям современников. М.: Наука, 2001. 398 с.

[8] Прикладная небесная механика и управление движением. Сборник статей, посвященный 90–летию со дня рождения Д. Е. Охоцимского / Сост.: Т. М. Энеев, М. Ю. Овчинников, А. Р. Голиков // Труды ИПМ им. М. В. Келдыша РАН. М.: ИПМ им. М. В. Келдыша, 2010. 368 с.

[9] Вернадский В. И. Биосфера и ноосфера. М.: Наука, 1989. 261 с.

[10] Марчук Г. И. Математическое моделирование в проблеме окружающей среды. М.: Наука, Главная редакция физико-математической литературы, 1982. 320 с.

[11] Моисеев Н. Н. Как далеко до завтрашнего дня. Свободные размышления. 1917–1993. М.: "Аспект пресс", 1994. 304 с. (Глава X. Эпопея ядерной зимы... Карл Саган и первые сценарии ядерной войны)

[12] Будыко М. И., Голицын Г. С., Израэль Ю. А. Глобальные климатические катастрофы. Л: Гидрометеоиздат. 1986. 159 с.

[13] Антропогенные изменения климата / Под ред. Будыко М. И. и Израэля Ю. А. Л: Гидрометеоиздат, 1987. 406 с.

[14] Фридман А. А. Мир как пространство и время. Серия "Современная культура". Петербург: Издатель "Academia", 1923. 132 с.

[15] Сушкевич Т. А. Нобелевский лауреат С. Чандрасекар: к 65–летию первой монографии по переносу лучистой энергии (Посвящается 65–летию ИПМ имени М. В. Келдыша и памяти профессора Е. С. Кузнецова, основателя отечественной научной школы и создателя отдела "Кинетические уравнения")// Препринты ИПМ им. М. В. Келдыша. 2018. № 228. 32 с.

[16] Кузнецов Е. С. Избранные научные труды (в связи со 100–летием со дня рождения) / Отв. ред. и сост. Сушкевич Т. А. М.: ФИЗМАТЛИТ, 2003. 784 с.

[17] Chandrasekhar S. Radiative transfer. London, Oxford, Clarendon Press, 1950. 405 p.(Чандрасекар С. Перенос лучистой энергии / Пер. с англ. издания под ред. Е. С. Кузнецова. М.: Изд-во иностранной литературы, 1953. 432 с.).

[18] Кондратьев К.Я. Перенос длинноволнового излучения в атмосфере. М.-Л.: Гос. изд-во техн.-теорет. лит., 1950. 288 с.

[19] Кондратьев К.Я. Лучистая энергия солнца. Л.: Гидрометеоиздат, 1954. 600 с.

[20] Сушкевич Т. А. О пионерских работах по математическому моделированию радиационного поля Земли при освоении космоса // Современные проблемы дистанционного зондирования Земли из космоса. 2008. Вып. 5. Т. 1. С. 165–180.

[21] Сушкевич Т. А. К истории первого научного эксперимента по дистанционному зондированию Земли на пилотируемом космическом корабле // Современные проблемы дистанционного зондирования Земли из космоса. 2008. Вып. 5. Т. 1. С. 315–322.

[22] Лазарев А. И., Николаев А. Г., Хрунов Е.В. Оптические исследования в космосе. Л.: Гидрометеоиздат, 1979. 256 с.

[23] Лазарев А. И., Коваленок В. В., Авакян С. В. Исследование Земли с пилотируемых космических кораблей. Л.: Гидрометеоиздат, 1987. 400 с.

[24] Сушкевич Т. А., Стрелков С. А., Максакова С. В. Земля, космос и суперкомпьютинг: сопряженные радиационные задачи (Посвящается 65–летию ИПМ имени М. В. Келдыша и памяти математика–легенды М. В. Келдыша)

[25] Сушкевич Т. А., Стрелков С. А., Максакова С. В. Космос, Земля и суперкомпьютинг: сопряженные задачи экологии, климата, мониторинга и дистанционного зондирования Земли, гиперспектральный подход и нанодиагностика природных сред (посвящается 65–летию ИПМ им. М. В. Келдыша РАН) // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2018. Т. 7, № 4. С. 5–29.

[26] Кароль И. Л., Катцов В. М., Киселев А. А., Кобышева Н. В. О климате по существу и всерьез. Санкт-Петербург: ГГО им. А. И. Воейкова, 2008. 55 с.

[27] Фомин Б. А., Колокутин Г. Э. Новая спектроскопическая база HITRAN–2016 в полинейных моделях, применяемых в дистанционном зондировании Земли методами инфракрасной спектрометрии // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16, № 1. С. 17–24.

[28] Моисеев Н. Н., Александров В. В. Ядерный конфликт глазами климатологов и математиков // Вестник АН СССР. 1984. № 11. С. 65–76.

[29] Сушкевич Т. А., Стрелков С. А., Максакова С. В. 60 лет от первого совещания по ИСЗ до современных систем дистанционного зондирования и мониторинга Земли из космоса: информационно–математический аспект (история и перспективы) // Оптика атмосферы и океана. 2014. Т. 27, № 7. С. 21–27.

[30] Козодеров В. В., Косолапов В. С., Садовничий В. А., Тимошин О. А., Тищенко А. П., Ушакова Л. А., Ушаков С. А. Космическое землеведение: информационно–математические основы / Под ред. Садовничего В. А. М.: Издательство МГУ, 1998. 571 с.

[31] Кондратьев К. Я., Григорьев А. А. Космическое землеведение / Отв. ред. Г. И. Марчук. Серия "Планета Земля и Вселенная". М.: Наука, 1985. 160 с.

[32] Кондратьев К. Я. Спутниковая климатология. Л.: Гидрометеоиздат, 1983. 263 с.

[33] Сушкевич Т. А., Стрелков С. А., Иолтуховский А. А. Метод характеристик в задачах атмосферной оптики. М.: Наука, 1990. 296 с. (При поддержке РИСО АН СССР)

[34] Сушкевич Т. А. Математические модели переноса излучения. М.: БИНОМ. Лаборатория знаний, 2005. 661 с. (При поддержке РФФИ)

[35] Масленников М. В., Сушкевич Т. А. Асимптотические свойства решения характеристического уравнения теории переноса излучения в сильно поглощающих средах // Ж. вычисл. матем. и матем. физ. 1964. Т. 4, № 1. С. 23–34.

[36] Гермогенова Т.А., Копрова Л.И., Сушкевич Т.А. Исследование угловой, пространственной и спектральной структуры поля яркости Земли для характерной модели сферической атмосферы // Изв. АН СССР. Серия Физика атмосферы и океана. 1969. Т. 5, № 12. С. 1266–1277.

[37] Ларин И. И. Он учил беречь Землю. Моисеев Н. Н. Серия: "Жизнь замечательных ученых". М.: Росэкопресс, 2002. 236 с.

Сушкевич Тамара Алексеевна — д.ф.-м.н., гл. науч.сотр. ФИЦ "Институт прикладной математики им. М.В.Келдыша РАН"; e-mail: tamaras@keldysh.ru;
Стрелков Сергей Александрович — к.ф.-м.н., ст. науч.сотр. ФИЦ "Институт прикладной математики им. М.В.Келдыша РАН"; e-mail: strelka50@mail.ru;
Максакова Светлана Викторовна — науч.сотр. ФИЦ "Институт прикладной математики им. М.В.Келдыша РАН"; e-mail: svmmr@yandex.ru.
Дата поступления — 30 апреля 2019 г.

О РЕШЕНИИ ОБРАТНОЙ ГРАНИЧНОЙ ЗАДАЧИ ТЕПЛОПРОВОДНОСТИ ДЛЯ ШАРА, СОСТОЯЩЕГО ИЗ КОМПОЗИТНЫХ МАТЕРИАЛОВ

В. П. Танана[1], Б. А. Марков[2]

[1] *Южно-Уральский государственный университет, 454080, Челябинск*

[2] *Челябинское высшее военное авиационное краснознамённое училище штурманов, 454015, Челябинск,*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10077

В работе предлагается постановка обратной граничной задачи теплопроводности композитных материалов для шара. Впервые обратная граничная задача для уравнения теплопроводности была рассмотрена в [1].
Сложность решения задачи состоит в том, что не удаётся использовать классическое решение, так как, в силу композитности материала, производная решения терпит разрыв на границе раздела сред.
Для обратной задачи приведена оценка погрешности [4, 8] приближённого решения.

Ключевые слова: обратные задачи, композитные материалы, погрешность приближённого решения.

Введение

Термоизмерительные устройства в современной технике нередко подвергаются б**о**льшим тепловым нагрузкам, чем могут выдержать, не выходя из строя. Поэтому их помещают в различные теплозащитные оболочки, имеющие разнообразные геометрические формы — одной из наиболее распространённых является шаровой слой, изготовленный из разных материалов.

Теплозащитный материал вносит весьма значительные искажения в показания приборов. Поэтому необходимо решать обратную задачу для уравнения теплопроводности в случае, когда свойство температуропроводности среды меняется скачком.

Классического решения такой задачи не существует, так как на границе раздела сред производная решения по пространственной координате имеет точки разрыва. Мы предлагаем рассмотреть решение такой задачи для шара, которое, с одной стороны, не является обобщённым, а с другой — допускает точки разрыва производной по координате.

1 Постановка и исследование прямой задачи шаровой термопары

Пусть тепловой процесс задан системой уравнений:

$$\frac{\partial u_1(r,t)}{\partial t} = \frac{a_1^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_1(r,t)}{\partial r}\right), \quad r \in (0;r_0),\ t \in (0;+\infty), \tag{1}$$

$$\frac{\partial u_2(r,t)}{\partial t} = \frac{a_2^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_2(r,t)}{\partial r}\right), \quad r \in (r_0;1),\ t \in (0;+\infty), \tag{2}$$

$$u_1(r,0) = 0,\ r \in [0;r_0], \quad u_2(r,0) = 0,\ r \in [r_0;1], \tag{3}$$

$$\exists A_1 > 0,\ |u_1(0,t)| < A_1, \quad t \in [0;+\infty), \tag{4}$$

$$u_2(1,t) = q(t), \quad t \in [0;+\infty), \tag{5}$$

$$u_1(r_0,t) = u_2(r_0,t), \quad t \in [0;+\infty), \tag{6}$$

$$a_1\frac{\partial u_1(r_0,t)}{\partial r} = a_2\frac{\partial u_2(r_0,t)}{\partial r}, \quad t \in [0;+\infty), \tag{7}$$

ISBN 978-5-901548-42-4

где используются обозначения: r — сферический радиус, t — координата по времени.

Пусть функция $q(t)$ финитна, дважды дифференцируема и равна нулю вместе с производной в нуле:

$$q(t) \in C^2[0;+\infty),\ q(0) = q'(0) = 0, \exists t_0 > 0,\ \forall t > t_0 : q(t) = 0. \tag{8}$$

В прямой задаче требуется найти функции $u_1(r,t)$ и $u_2(r,t)$.

Будем искать формальное решение задачи (1–7) в виде ряда по собственным функциям $\{S_n(r)\}_{n=1}^{\infty}$ соответствующей задачи Штурма-Лиувилля:

$$u(r,t) = q(t) - \sum_{n=1}^{\infty} u_n(t) S_n(r), \tag{9}$$

где функции $u(r,t)$ и $S_n(r)$ определены формулами:

$$u(r,t) = \begin{cases} u_1(r,t), r \in [0;r_0],\ t \in (0;+\infty), \\ u_2(r,t), r \in [r_0;1],\ t \in (0;+\infty), \end{cases} \quad S_n(r) = \begin{cases} S_n^{(1)}(r), r \in [0;r_0], \\ S_n^{(2)}(r), r \in [r_0;1]. \end{cases}$$

Задача Штурма-Лиувилля для системы (1–7) приведена в [2], равно как и система собственных функций:

$$S_n(r) = \frac{\beta_n}{r} \begin{cases} \sin\left(\dfrac{\lambda_n r}{a_1}\right) \sin\left(\dfrac{\lambda_n (1-r_0)}{a_2}\right), & r \in [0;r_0], \\ \sin\left(\dfrac{\lambda_n r_0}{a_1}\right) \sin\left(\dfrac{\lambda_n (1-r)}{a_2}\right), & r \in [r_0;1], \end{cases} \tag{10}$$

где норма функции β_n имеет вид:

$$\beta_n^2 = \frac{2}{\sin^2\left(\dfrac{\lambda_n(1-r_0)}{a_2}\right)\left(\dfrac{r_0}{a_1} + \dfrac{1-r_0}{a_2}\right)} + \frac{A_2}{n} + \overline{\overline{o}}\left(\frac{1}{n}\right),\ A_2 = const, \tag{11}$$

$$\lambda_n = \frac{a_1 a_2}{a_2 r_0 + (1-r_0) a_1} (\pi n + \gamma_n), n = 1,2,3,..., \gamma_n = \frac{A_3}{n} + \overline{\overline{o}}\left(\frac{1}{n}\right),\ A_3 = const, \tag{12}$$

а выражения $u_n(t)$ имеют вид:

$$u_n(t) = -\frac{\beta_n}{\lambda_n} \sin\left(\frac{\lambda_n r_0}{a_1}\right) \int_0^t q'(\tau) \exp(-\lambda_n^2 (t-\tau)) d\tau. \tag{13}$$

Утверждение 1. Система функций $\{S_n(r)\}_{n=1}^{\infty}$ полна и ортогональна в пространстве $L_2[0;1]$ с соответствующим весом.

Доказательство. Ортогональность системы функций $\{S_n(r)\}_{n=1}^{\infty}$ следует из вида соответствующей задачи Штурма-Лиувилля, а полнота системы — из положительности всех собственных значений на основании теоремы Гильберта-Шмидта [3].

Определим решение задачи (1–7) следующим образом:

Определение 1. Решение задачи (1–7) заключается в определении функций $u_1(r,t)$ и $u_2(r,t)$, таких, что:

1. $u_1(r,t)$ удолетворяет уравнению (1) на множестве $(0;r_0) \times (0;+\infty)$, а $u_2(r,t)$ удовлетворяет уравнению (2) на множестве $(r_0;1) \times (0;+\infty)$;
2. $u(r,t) \in C([0,1] \times [0,\infty)) \cap C^{2,1}((0,r_0) \times (0;\infty)) \cap C^{2,1}((r_0,1) \times (0,\infty))$;
3. $u(r,t)$ удовлетворяет начальному условию (3), граничным условиям (4), (5), а также условиям согласования (6) и (7).

Справедливо также следующее утверждение:

Утверждение 2. Пусть дана задача (1-7). Тогда её решение в смысле определения 1 задано формулой (9).

2 Применимость преобразования Фурье для решения обратной задачи.

Для аналитической оценки погрешности приближённого решения нам потребуется преобразование Фурье по времени t, для которого справедлива теорема Планшереля [3]. Для этого продолжим функцию $u(r,t)$ на область отрицательных t нулём и покажем, что финитность функции $q(t)$ будет достаточна для применимости преобразования Фурье.

Теорема 1. Пусть решение задачи (1–7) в смысле определения 1 рассматривается для финитной функции $q(t)$, удовлетворяющей условию (8). Тогда существует Фурье-преобразование по t функций $u_1(r,t)$ и $u_2(r,t)$.

Доказательство. Доказательство аналогично соответствующему доказательству, приведённому в [12]. Так как функция $q(t)$ финитна, то существует такое $t_0 > 0$, что $\forall t > t_0\ q(t) = 0$. Представим задачу (1-7) как две, используя обозначения:

$$u_1(r,t) = \begin{cases} v_1(r,t),\ r \in [0,r_0],\ t \in [0,t_0], \\ v_2(r,t),\ r \in [r_0,1],\ t \in [0,t_0], \end{cases} \quad u_2(r,t) = \begin{cases} w_1(r,t),\ r \in [0,r_0],\ t \in [t_0,+\infty), \\ w_2(r,t),\ r \in [r_0,1],\ t \in [t_0;+\infty). \end{cases}$$

Тогда решение задачи (1-7) будет иметь вид (9) для функции $v(r,t)$, а для функции $w(r,t)$ примет вид:

$$w(r,t) = q(t) - \sum_{n=1}^{\infty} v_n(t_0) S_n(r) \exp(-\lambda_n^2 t). \tag{14}$$

Тогда преобразование Фурье $\hat{u}(r,\tau) = \int\limits_0^{+\infty} \exp(-i\tau t) u(r,t)dt$, $\tau \in (-\infty;+\infty)$, от функции $u(r,t)$ примет вид:

$$\hat{u}(r,\tau) = \int\limits_0^{t_0} \exp(-i\tau t) v(r,t)dt + \int\limits_{t_0}^{+\infty} exp(-i\tau t) w(r,t)dt. \tag{15}$$

Несложно видеть, что первый интеграл существует в силу финитности по времени функции $v(r,t)$.

Второй интеграл рассмотрим отдельно. Ряд (14) сходится равномерно, и формально может быть проинтегрирован почленно. В результате получается ряд:

$$\hat{w}(r,\tau) = \hat{q}(\tau) - \sum_{n=1}^{\infty} v_n(t_0) S_n(r) \frac{\exp(-i\tau t_0 - \lambda_n^2 t_0)}{\lambda_n^2 + i\tau}.$$

Несложно видеть, что этот ряд может быть оценен сверху сходящимся числовым рядом, а, следовательно, сходится равномерно как по τ, так и по r.

Следствие Из применимости преобразования Фурье следует, что решение задачи (1–7) с условиями (8) единственно.

Теперь выпишем преобразование Фурье всей задачи (1–7) по t. Получим:

$$i\tau \hat{u}_1(r,\tau) = \frac{a_1^2}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \hat{u}_1(r,\tau) \right),\ r \in (0;r_0),\ \tau \in (-\infty;+\infty), \tag{16}$$

$$i\tau \hat{u}_2(r,\tau) = \frac{a_2^2}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \hat{u}_2(r,\tau) \right),\ r \in (r_0;1),\ \tau \in (-\infty;+\infty), \tag{17}$$

$$|\hat{u}_1(0,\tau)| \le C_1,\ \tau \in (-\infty;+\infty), \tag{18}$$

$$\hat{u}_1(r_0,\tau) = \hat{u}_2(r_0,\tau) = \hat{f}(\tau),\ \tau \in (-\infty;+\infty). \tag{19}$$

Здесь использованы обозначения: $\hat{p}(r,\tau) = \int\limits_0^{+\infty} p(r,t) \exp(i\tau t)dt$.

Решив обыкновенные дифференциальные уравнения (16–19), получим решение. Вид решения приведём только для случая $r = r_0$:

$$\hat{u}_1(r_0,\tau) = \hat{u}_2(r_0,\tau) = T(\tau)\hat{q}(\tau), \quad T(\tau) = -\frac{1}{r_0} \frac{\mathrm{sh}(\alpha)}{\mathrm{sh}(\alpha+\beta)}, \tau \in (-\infty;+\infty), \tag{20}$$

где использованы обозначения: $\alpha = \mu\frac{r_0}{a_1}\sqrt{\tau}$, $\beta = \mu\frac{1-r_0}{a_2}\sqrt{\tau}$, $\mu = \frac{1+i}{\sqrt{2}}$.

Так как $\hat{u}_1(r_0,\tau) = \hat{u}_2(r_0,\tau) = \hat{f}(\tau)$, то задача (20) принимает вид:

$$T(\tau)\hat{q}(\tau) = \hat{f}(\tau). \tag{21}$$

3 Постановка обратной граничной задачи

Пусть в условии (5) функция $q(t)$ неизвестна, а вместо неё дана функция $f(t)$, определяемая формулой

$$u(r_0,t) = f(t). \tag{22}$$

Требуется, используя $f(t)$, определить функцию $q(t)$ так, что при подстановке её в условие (5) решение задачи $u(r,t)$ задачи (1–7) удовлетворяло бы соотношению (21).

Так как обратная задача некорректно поставлена, то дополнительно предположим, что при $f(t) = f_0(t) \in C[0;+\infty)$ существует решение $q_0(t)$ обратной задачи (1–4) — (21), которое удовлетворяет условиям (8) и (9). При этом оно при подстановке функции $q_0(t)$ в условие (5) прямой задачи (1–7) будет удовлетворять условию (21).

Пусть теперь нам вместо функции $f_0(t)$ даны некоторое приближение $f_\delta(t) \in C[0;+\infty)$ и уровень погрешности $\delta > 0$ такие, что справедливо неравенство:

$$\sup_{t\geq 0}|f_\delta(t) - f_0(t)| \leq \delta \tag{23}$$

Введём класс корректности: $q_0(t) \in M_{r_1}$, где

$$M_{r_1} = \left\{ g(t) : g(t) \in H_0^1[0;+\infty), ||g(t)||^2_{H_0^1} \leq r_1^2 \right\}.$$

При этих условиях приступим к решению задачи (21). Уравнение (21) является некорректной задачей, так как требует обращения преобразования Фурье неограниченного оператора. Поэтому для решения этого уравнения необходима дополнительная информация.

Пусть при $\hat{f}(\tau) = \hat{f}_0(\tau) \in L_2(-\infty;+\infty)$ существует точное решение $\hat{q}_0(\tau)$, то есть справедливо точное равенство

$$T(\tau)\hat{q}_0(\tau) = \hat{f}_0(\tau), \tag{24}$$

но $\hat{f}_0(\tau)$ неизвестна, а вместо неё дана функция $\hat{f}_\delta(\tau)$ такая что

$$||\hat{f}_\delta(\tau) - \hat{f}_0(\tau)||^2_{L_2} \leq \delta^2. \tag{25}$$

Кроме того, из Фурье-преобразования и свойств, заданных формулой (21) следует, что

$$\hat{q}_0(\tau) \in \hat{H}^1(-\infty;+\infty), \tag{26}$$

где

$$||\hat{q}_0(\tau)||^2_{\hat{H}^2} = \int_{-\infty}^{+\infty}(1+\tau^2)|\hat{q}_0(\tau)|^2 d\tau, \tag{27}$$

а из финитности функции $q(t)$ и (8) — что

$$||\hat{q}_0(\tau)||^2_{\hat{H}^1} \leq r_1^2. \tag{28}$$

Для решения задачи (24–28, 21) используем известный метод невязки [5]– [8]. Этот метод заключается в сведении данной задачи к вариационной:

$$\inf\left\{ ||\hat{q}(\tau)||^2_{\hat{H}^1} : ||T(\tau)\hat{q}(\tau) - \hat{f}_\delta(\tau)||^2_{L_2} \leq \delta^2 \right\}. \tag{29}$$

Решение (29) может быть построено методом регуляризации по Тихонову:

$$\hat{q}_\delta^\alpha(\tau) = \left\{|T(\tau)|^2 + \alpha(1+\tau^2)\right\}^{-1} \times \left\{\overline{T}(\tau)\hat{f}_\delta(\tau)\right\},\ \tau \in (-\infty;+\infty). \tag{30}$$

В методе регуляризации необходимо выбрать параметр регуляризации α.

Для дальнейшего удобно разбить оператор $T(\tau)$ следующим образом: $T(\tau) = T_1(\tau) + T_2(\tau)$, где

$$T_1(\tau) = \begin{cases} 0, & \tau \in (-\infty;-\alpha)\cup(\alpha;+\infty), \\ T(\tau), & -\alpha \le \tau \le \alpha, \end{cases} \quad T_2(\tau) = \begin{cases} T(\tau), & \tau \in (-\infty;-\alpha)\cup(\alpha;+\infty), \\ 0, & -\alpha \le \tau \le \alpha. \end{cases} \tag{31}$$

В этом случае справедлива следующая теорема:

Теорема Пусть $T_2(\tau)$ определяется формулой (31). Тогда

$$T_2^{-1}(\tau) \sim \exp(k\sqrt{\tau}),\ \tau \to \infty, \tag{32}$$

где $k = \dfrac{1-r_0}{a_2\sqrt{2}}$, и верна оценка

$$\sqrt{\frac{1-2\epsilon}{1+2\epsilon}}\exp(k\sqrt{\tau}) \le |T_2^{-1}(\tau)| \le \sqrt{\frac{1+2\epsilon}{1-2\epsilon}}\exp(k\sqrt{\tau}),\ \epsilon = \frac{1}{\mathrm{ch}(k\sqrt{\alpha})}. \tag{33}$$

Вводя $\hat{q}_1^\alpha(\tau) = T^{-1}(\tau)\hat{f}(\tau)$, $\hat{q}_2^\alpha(\tau) = T_2^{-1}(\tau)\hat{f}(\tau)$, получаем оценку:

$$||\hat{q}(\tau) - \hat{q}_\delta(\tau)||_{L_2} \le ||\hat{q}(\tau) - \hat{q}_1^\alpha(\tau)||_{L_2} + ||\hat{q}_1^\alpha(\tau) - \hat{q}_\delta(\tau)||_{L_2}. \tag{34}$$

Так как известны оценки $||\hat{q}_1^\alpha(\tau) - \hat{q}_\delta(\tau)||_{L_2} \le C_T\delta$, где C_T — некоторая постоянная, оценивающая $||T_1^{-1}||$ (этот оператор определён на конечном промежутке и конечен), и $||\hat{q}_1^\alpha(\tau) - \hat{q}(\tau)||_{L_2} = ||\hat{f}(\tau)||_{L_2}||T_2^{-1}(\tau)||_{L_2}$. Т.к. $||\hat{f}(\tau)||_{L_2} \le \dfrac{r_1}{\sqrt{1+\alpha^2}}$ в силу принадлежности $q(t) \in M_{r_1}$ и (21), то из (34) получаем, выбрав число $C_2 > 0$, что $C_2 \ge \dfrac{1}{\sqrt{1+\alpha^2}}$:

$$||\hat{q}(\tau) - \hat{q}_\delta(\tau)||_{L_2} \le C_T\delta + C_2 r_1 \exp(k\sqrt{\tau}). \tag{35}$$

Выбрав $\alpha = \tau$ и считая, что $C_T\delta = C_2\exp(k\sqrt{\tau})$, получаем соотношение для нахождения парамера регуляризации α :

$$\alpha = \frac{1}{k^2}\ln^2\left(\frac{C_T\delta\sqrt{1-2\epsilon}}{r_1 C_2\sqrt{1+2\epsilon}}\right), \tag{36}$$

где ϵ определяется формулой (33).

Рассмотрим теперь оценку по порядку погрешности [6]. Из [8] следуют существование и единственность решения задачи (29). Обозначим это решение через $\hat{q}_\delta(\tau)$ и назовём его приближённым решением (21).

Из [9] следует, что для этого решения справедлива оценка

$$||\hat{q}_\delta(\tau) - \hat{q}_0(\tau)||_{L_2} \le 2\omega(\delta, r_1), \tag{37}$$

где, следуя [10],

$$\omega(\delta, r_1) = \sup_{\hat{q}(\tau)}\left\{||\hat{q}(\tau)||_{L_2} :\ ||\hat{q}(\tau)||_{\dot{H}^1}^2 \le r_1^2,\ ||T(\tau)\hat{q}(\tau) - \hat{f}_\delta(\tau)||_{L_2}^2 \le \delta^2\right\}. \tag{38}$$

Из [11], [12], (37), (38) следует существование числа A, такого, что при достаточно малых значениях $\delta > 0$ справедливо

$$\omega(\delta, r_1) \le A\ln^{-2}(\delta). \tag{39}$$

Теперь, используя функцию $\hat{q}_\delta(\tau)$, определим приближённое решение $q_\delta(t)$ обратной задачи (21):

$$q_\delta(t) = \begin{cases} \Re\left(\int\limits_{-\infty}^{+\infty}\exp(i\tau t)\hat{q}_\delta(\tau)d\tau\right), & t \ge 0, \\ 0, & t < 0. \end{cases} \tag{40}$$

Так как преобразование Фурье изометрично в пространстве $L_2(-\infty;+\infty)$, то из (40) следует, что

$$||q_\delta(t) - q_0(t)||_{L_2} \le \frac{A}{\sqrt{2}}\ln^{-2}\left(\frac{1}{\delta}\right). \tag{41}$$

Заключение

В данной работе дана постановка обратной граничной задачи теплопроводности для шара, состоящего из композитных материалов, и кратко рассмотрено решение обратной задачи. Для получения оценки погрешности приближённого решения обратной задачи использовано преобразование Фурье по временной переменной t. В результате получена оценка погрешности приближённого решения исходной задачи.

Список литературы

[1] Тихонов А.Н., Гласко В.Б. К вопросу о методах определения температуры поверхности тела // 1967 ЖВМиМФ. Т. 7. № 4. С. 267-273.

[2] Будак Б.М., Самарский А.А., Тихонов А.Н. Сборник задач по математической физике. М.: Физматлит. 2004.

[3] Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Физматлит. 2004.

[4] Кабанихин С.И. Обратные и некорректные задачи. Новосибирск. 2009.

[5] Phillips D.L. A technique for the numerical solution of certain integral equations of the first kind // Journal of ACM. 1962. V. 9, iss. 1. P. 84–97.

[6] Танана В.П. Об оптимальной по порядку метода проекционной регуляризации при решении обратных задач // Сибирский журнал индустриальной математики. 2004. Т. 7. № 2. С. 117–132.

[7] Иванов В.К. О приближённом решении операторных уравнений // Журнал вычислительной математики и математической физики. 1966. Т. 6. № 6. С. 1089–1094.

[8] Иванов В.К., Васин В.В., Танана В.П. Теория линейных некорректных задач и её приложения. М.: Наука. 1978. 208 с.

[9] Танана В.П. Об оптимальности методов решения нелинейных неустойчивых задач // ДАН СССР. Т. 220. № 5. С. 1035–1037.

[10] Иванов В.К., Королюк Т.И. Об оценке погрешности при решении нелинейных некорректно поставленных задач // Журнал вычислительной математики и математической физики. 1969. Т. 9. № 1. С. 30–41.

[11] Танана В.П., Колесникова Н.Ю. Об оценке погрешности при решении линейных некорректно поставленных задач // Известия Уральского государственного университета. Сер. Математика. Механика. Информатика. 2008. № 58. С. 155–162.

[12] Танана В.П., Сидикова А.И. О гарантированной оценке точности приближённого решения одной обратной задачи тепловой диагностики // Труды Института математики и механики УрО РАН. 2010. Т. 16. № 2. С. 238–252.

Танана Виталий Павлович — д.ф.-м.н., главный научный сотрудник, кафедра системного программирования, Южно-Уральский государственный университет;
e-mail: tvpa@susu.ac.ru;
Марков Борис Анатольевич — к.ф.-м.н., доцент кафедры математики и естественно-научных дисциплин ВУНЦ ВВС ВВА ЧВВАИШ;
e-mail: smpx1969@mail.ru.
Дата поступления — 30 апреля 2019 г.

ВОССТАНОВЛЕНИЕ КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ ЗЕМНОЙ ПОВЕРХНОСТИ ПО ДАННЫМ ПРИБОРОВ MODIS, САНГУР, ГСА И МСУ-100

М. В. Тарасенков, В. В. Белов, М. В. Энгель, А. В. Зимовая, Ю. В. Гриднев

Институт оптики атмосферы им. В.Е. Зуева СО РАН, 634021, Томск

УДК 528.85:535.51, 535.361
DOI: 10.24411/9999-016A-2019-10078

Рассматриваются результаты восстановления коэффициентов отражения земной поверхности по спутниковым измерениям приборов MODIS, Сангур, ГСА и МСУ-100. Результаты получены с помощью разработанного программного обеспечения. Решение задачи состоит в формировании, исходя из спутниковых или наземных измерений, моделей атмосферы; моделировании методом Монте-Карло потоков, формирующих принимаемое излучение; и решении обратной задачи восстановления коэффициентов отражения земной поверхности по измерениям интенсивности принимаемого излучения.

Ключевые слова: атмосферная коррекция, коэффициент отражения земной поверхности, метод Монте-Карло, MODIS, САНГУР, ГСА, МСУ-100.

Введение

Спутниковые данные о коэффициенте отражения земной поверхности имеют в настоящий момент самый широкий круг применений. Вместе с тем задача получения коэффициентов отражения земной поверхности из спутниковых измерений является сложной обратной задачей. Одно из направлений связано с решением уравнения переноса излучения (УПИ). В настоящее время существует достаточно много алгоритмов решения УПИ. Подробный обзор методов можно найти, например, в [1]. Существенный вклад в разработку алгоритмов решения УПИ был внесен в таких отечественных работах как [2–6]. Вместе с тем наиболее универсальным алгоритмом решения УПИ является метод Монте-Карло, основы которого были заложены в таких работах как [2]. Он почти не имеет ограничений на геометрию задачи и оптические параметры атмосферы. Кроме того, достаточно широкое распространение имеют методы дискретных ординат и сферических гармоник.

Разработка RTM-алгоритмов ведется на протяжении десятков лет (начиная с таких работ как [7]). В общем случае в видимом и ближнем ИК-диапазонах принимаемое спутниковой системой излучение формируется следующими потоками:
а) нерассеянное излучение от наблюдаемого участка земной поверхности I_0;
б) излучение, рассеянное в атмосфере и не взаимодействовавшее с земной поверхностью I_{sun};
в) излучение, отраженное соседними участками земной поверхности и далее рассеянное в атмосфере — боковой подсвет I_{surf}.

В свою очередь, отраженное излучение формируется как однократно отраженным излучением Солнца E_0, так и многократно отраженным излучением E_1, E_2 и т.д.

В общем случае земная поверхность неоднородна. Решение обратной задачи в полной постановке имеет существенные трудности. Для ускорения получения результатов в первых работах (например, [7]) использовалось приближение однородной поверхности и однократного рассеяния излучения в среде. В более поздних работах (например, [8]) использовалось приближение однородной поверхности. В работе [9] обратная задача решалась для упрощенной модели формирования бокового подсвета и дополнительной освещенности. В работе [10] автор точно учитывает боковой подсвет, а дополнительную освещенность земной поверхности

ISBN 978-5-901548-42-4

учитывает в приближении однородной земной поверхности. В разрабатываемом нами алгоритме [11,12] влияние бокового подсвета и дополнительной освещенности описывается с учетом того, что земная поверхность неоднородна.

Причина, почему требуется более точный (и более сложный) алгоритм, состоит в том, что оценки, обосновывающие использование более простых алгоритмов, например, [13], выполнялись для длин волн $\lambda \geqslant 0.5$ мкм, оптических толщин $\tau_{atm} < 1$ и коэффициентов отражения $r_{surf} \leqslant 0.4$. Однако, как показывают наши оценки, решение рассматриваемой задачи возможно и для значительно больших оптических толщин (вплоть до $\tau_{atm} \backsim 4$). Это, в свою очередь, требует более точного учета потоков, формирующих принимаемое излучение. На длинах волн $\lambda < 0.5$ мкм влияние неоднородности земной поверхности повышается, так как увеличиваются размеры областей, формирующих боковой подсвет и дополнительную освещенность. Помимо этого на земной поверхности встречаются области с большим отражением, например, покрытые снегом территории. И для участков с резким переходом от сильно отражающих поверхностей к слабоотражающим требуется учет влияния неоднородности земной поверхности. Например, в работе [14] указывается, что в ситуации высокой мутности атмосферы при наблюдении маленьких озер, окруженных растительностью, во втором канале прибора AVHRR (центр канала $\lambda = 0.85$ мкм) нельзя пренебрегать ошибками, связанными с учетом бокового подсвета.

Помимо рассмотренных факторов на восстановление коэффициента отражения земной поверхности могут влиять рельеф поверхности, неламбертовость отражения, поляризация излучения, облачные поля. Учет влияния рельефа можно выполнять путем нормализации поверхности (например, [15]). Неламбертовость поверхности можно учитывать после решения задачи для ламбертовской поверхности подходом, описанным в [16]. Влияние поляризации излучения рассматривалось в наших работах [17,18] и работах других авторов (например, [4, 5, 19]). Работы по оцениванию влияния облачности на безоблачные участки начаты нами в таких работах как [20] и рассматриваются, например в [21]. Используемый в настоящей работе алгоритм строился в приближении отражения Ламберта. Учет рельефа, облачности и неламбертовости отражения можно выполнять независимо от рассматриваемого алгоритма.

В настоящий момент созданы и развиваются штатные алгоритмы для восстановления коэффициентов отражения по измерениям приборов MODIS [16], POLDER [22], MISR [23] и др. Однако аналогичных средств для современных отечественных спутниковых систем нет.

Для оценки возможностей предложенного алгоритма рассматриваются спутниковые данные MODIS. Для данных отечественных спутниковых приборов приводятся примеры восстановления коэффициентов отражения.

1 Математическая модель

1.1 Постановка задачи

Задача рассматривается в следующей постановке. На верхнюю границу атмосферы в направлении $\overrightarrow{\omega_{sun}}$ падает параллельный поток солнечного излучения. На некоторой высоте от поверхности Земли h_d расположена спутниковая сканирующая оптико-электронная система, которая формирует изображение участка земной поверхности на длине волны λ в направлении $\overrightarrow{\omega_d}$. Интенсивность принимаемого спутниковой системой излучения считается известной. Атмосфера разделена на сферические однородные слои, в каждом из которых заданы коэффициенты аэрозольного и молекулярного рассеяния и ослабления $\sigma_{s,a}$, $\sigma_{s,m}$, $\sigma_{t,a}$, $\sigma_{t,m}$, а также индикатриса аэрозольного рассеяния g_a. Поверхность Земли неоднородна и отражает излучение по закону Ламберта. Требуется, зная интенсивности принимаемого излучения для всех наблюдаемых пикселей земной поверхности (т.е. ее яркостное изображение) и оптические параметры атмосферы, восстановить коэффициенты отражения земной поверхности на длине волны λ.

1.2 Основные уравнения

В приближении однородной поверхности суммарная принимаемая интенсивность излучения определяется по формуле:

$$I_{sum,i} = I_{sun,i} + \frac{r_{surf1,i} E_0}{(1 - r_{surf1,i}\gamma_1)}(\frac{1}{\pi} exp(-\tau_{d,i}) + I_{surf,i}) \tag{1}$$

где $I_{sun,i}$ — интенсивность излучения, не взаимодействовавшего с земной поверхностью при наблюдении i-го пикселя; $r_{surf1,i}$ — коэффициент отражения наблюдаемого пикселя; γ_1 — вклад однократно отраженного

излучения в освещенность земной поверхности или сферическое альбедо атмосферы; $\tau_{d,i}$ — оптическая толщина трассы от наблюдаемого пикселя до приемной системы; $I_{surf,i}$ — интенсивность рассеянного излучения от поверхности при единичной светимости поверхности.

Исходя из (1), получаем соотношение для определения коэффициента отражения земной поверхности в приближении однородной поверхности:

$$r_{surf1,i} = \frac{\frac{Q_i}{E_0}}{1 + \gamma_1 \frac{Q_i}{E_0}}, i = \overline{1,N} \tag{2}$$

$$Q_{1i} = \frac{I_{sum,i} - I_{sun,i}}{\frac{1}{\pi} exp(-\tau_{d,i}) + I_{surf,i}} \tag{3}$$

где Q_i — суммарная светимость земной поверхности в приближении однородной поверхности. Величины I_{sun}, E_0, γ_1, $I_{surf,i}$ определяются методом Монте-Карло.

В разработанном алгоритме, учитывающим влияние неоднородности земной поверхности на боковой подсвет и дополнительную освещенность земной поверхности отраженным излучением, коэффициент отражения земной поверхности $r_{surf2,i}$ определяется из систем уравнений:

$$I_{sum,i} = I_{sun,i} + \frac{Q_{2i}}{\pi} exp(-\tau_{d,i}) + \sum_{j=1}^{N} Q_{2j} A_{ij} + A_{out,i} Q_{aver}, i = \overline{1,N} \tag{4}$$

$$A_{ij} \approx h(\theta_{d,i}, x_{w,j}, y_{w,j}) S_j \tag{5}$$

$$A_{out,i} = \int\limits_S h(\theta_{d,i}, x_w, y_w) dx_w dy_w - \sum_{j=1}^{N} A_{ij} \tag{6}$$

$$Q_{2i} \equiv r_{surf2,i} E_{sum,i} = r_{surf2,i} E_0 (1 + \sum_{j=1}^{N} r_{surf2,j} C_{ij} + C_{out,i} \overline{r_{surf}} + \frac{(\overline{r_{surf}} \gamma_1)^2}{1 - \overline{r_{surf}} \gamma_1}) \tag{7}$$

$$C_{ij} \approx h_1(r_{w,ij}) S_j \tag{8}$$

$$C_{out,i} = \gamma_1 - \sum_{j=1}^{N} C_{ij} \tag{9}$$

$$\overline{r_{surf}} = \frac{1}{N} \sum_{j=1}^{N} r_{surf1,i} \tag{10}$$

где N — количество пикселей, для которых выполняется восстановление коэффициентов отражения; $r_{surf2,i}$ — коэффициент отражения земной поверхности, при учете влияния неоднородности отражения земной поверхности; $E_{sum,i}$ — суммарная освещенность земной поверхности с учетом влияния неоднородности земной поверхности и вклада отраженного излучения; A_{ij} — величина, характеризующая интенсивность бокового подсвета, приходящего от j-го пикселя с площадью S_j при единичной светимости земной поверхности; $A_{out,i}$ — интенсивность бокового подсвета при единичной светимости земной поверхности, приходящего из областей вне области восстановления; Q_{aver} — средняя светимость земной поверхности в приближении однородной поверхности по участку, для которого выполняется восстановление; h — ФРТ канала формирования бокового подсвета; θ_d — угол отклонения линии визирования от направления в надир; C_{ij} — дополнительная освещенность i-го пикселя однократно отраженным излучением j-го пикселя при единичной светимости земной поверхности; h_1 — ФРТ канала формирования дополнительной совещенности; $r_{w,ij}$ — поверхностное расстояние от центра i-го пикселя до центра j-го пикселя; $C_{out,i}$ — дополнительная освещенность i-го пикселя участками вне области восстановления однократно отраженным излучением при единичной светимости земной поверхности; $\overline{r_{surf}}$ — средний по рассматриваемой области коэффициент отражения в приближении однородной поверхности; S — вся площадь земной поверхности.

Величины I_{sun}, h, E_0, γ_1, h_1 определяются методом Монте-Карло. В [24] показано, что (4) сходится при $A_{ii} \geq \sum\limits_{j=1, j\neq i}^{N} A_{ij}$, что выполняется при оптической толщине атмосферы $\tau_{atm} \leqslant 1$ при малых размерах пикселей S_j и для $\tau_{atm} \leqslant 4$ при размерах пикселей $S_j \backsim 0.78$ км^2. Система уравнений (7) сходится для всех рассмотренных оптических толщин атмосферы плоть до $\tau_{atm} < 7$.

2 Методы решения уравнений

Для ускорения получения решений применялось ряд приемов:

1. величина I_{sun} определялась с использованием расчетов методом Монте-Карло для узловых направлений и аппроксимационной формулы из [25] для всех остальных направлений;

2. по критерию изопланарности [26] определялись диапазоны углов, в пределах которых с точностью δ_1 можно считать функцию h независящей от угла θ_d;

3. вводя радиус бокового подсвета и радиус области формирования дополнительной освещенности по критериям [26], число членов в суммах по j в системах (4), (7) можно существенно уменьшить с контролируемыми погрешностями δ_2, δ_3.

Система уравнений (4) решалась методом Зейделя [27] с начальным приближением, определяемым в приближении однородной поверхности (3). Система нелинейных уравнений (7) решалась методом Ньютона [27] с итерационным решением систем линейных уравнений методом Зейделя.

Для расчета величин I_{sun}, h, E_0, γ_1, h_1, I_{surf} разрабатывались программы метода Монте-Карло. Методическую основу для программ составляла работа [2]. Созданные программы тестировались в работах [11,17,25] путем сравнения с результатами J. Lenoble, Y.J. Kaufman и K. Coulson. Сравнение показывает правильность работы разработанных программ. При расчетах контролировалось, чтобы при заданном числе траекторий относительная погрешность определения величин h и h_1 не превышала 0.01, а для остальных величин не превышала 0.005. Как показала работа [17], можно существенно снизить погрешность восстановления коэффициентов отражения земной поверхности для слабоотражающих поверхностей, если величину I_{sun} определять с учетом поляризации излучения. Модифицированная программа для I_{sun} также основана на работе [2]. Выполненное в [17] тестирование программы показало правильность работы разработанной программы.

3 Результаты

Разработанный комплекс программ апробировался на серии снимков прибора MODIS для 1-4 и 8 каналов. Были рассмотрены 3 тестовых участка: 1) юг Томской области $55.95^0 - 56.85^0$ с.ш. и $84.05^0 - 84.95^0$ в.д., 2) участок г. Москвы $55.72^0 - 55.95^0$ с.ш. и $37.56^0 - 38.10^0$ в.д., 3) участок Иркутской области $51.42^0 - 52.67^0$ с.ш. и $103.64^0 - 105.47^0$ в.д. Количество пикселей в рассматриваемых участках лежало в интервале от 2000 до 15000. Для построения используемой в алгоритме модели атмосферы использовались наземные измерения со станций Аэронет [28] аэрозольной оптической толщины, распределения частиц по размерам, комплексного показателя преломления, спутниковые измерения MODIS профиля температур и модели атмосферы LOWTRAN-7 [29]. Полученные результаты с учетом и без учета поляризации излучения сопоставлялись для участков, покрытых хвойным лесом, с результатами алгоритма MOD09 [30], результатами без атмосферной коррекции и с результатами наземных измерений из [31], принятых как эталонные. На рис. 1 приведены средние по рассмотренной серии снимков отличия результатов от эталонных значений для трех участков. Анализ результатов показывает, что модернизированный алгоритм дает погрешности существенно ниже, чем алгоритм MOD09 NASA на длинах волн 0.469, 0.555 и 0.649 мкм, погрешность одного порядка на длине волны 0.860 мкм и большую погрешность на длине волны 0.420 мкм. Последнее обстоятельство, по видимому, связано с необходимостью совершенствовать используемую в расчетах оптическую модель атмосферы.

Аналогичная работа проделывалась для снимков приборов Сангур, ГСА (спутник Ресурс-П) и МСУ-100 (спутник Метеор-М). Количество пикселей в рассматриваемых снимках этих приборов составлял порядка 3.5 млн. пикселей на 1 канал для прибора Сангур, 8.3 млн. пикселей на канал для прибора ГСА и 95 млн. пикселей на канал для прибора МСУ-100. В силу гораздо большего числа пикселей используемая ЭВМ для алгоритма в нынешнем виде оказалась не способна обработать такие крупные участки за разумное время, поэтому для атмосферной коррекции на данном этапе использвались формулы в приближении однородной поверхности (2) и (3). Каналы приборов МСУ-100 и Сангур гораздо шире каналов MODIS, а количество каналов ГСА превышает 100 в диапазоне 0.4-1.0 мкм. По этой причине используемая при восстановлении модель атмосферы была существенно модифицирована. Оптические коэффициенты определялись с учетом аппаратных функций приборов. Для построения более качественной модели молекулярного поглощения дополнительно использовались спутниковые измерения MODIS общего содержания озона и водяного пара [30], а также сечения поглощения кислорода, озона и водяного пара из базы HITRAN [32]. Пример полученных результатов приведен на рис. 2.

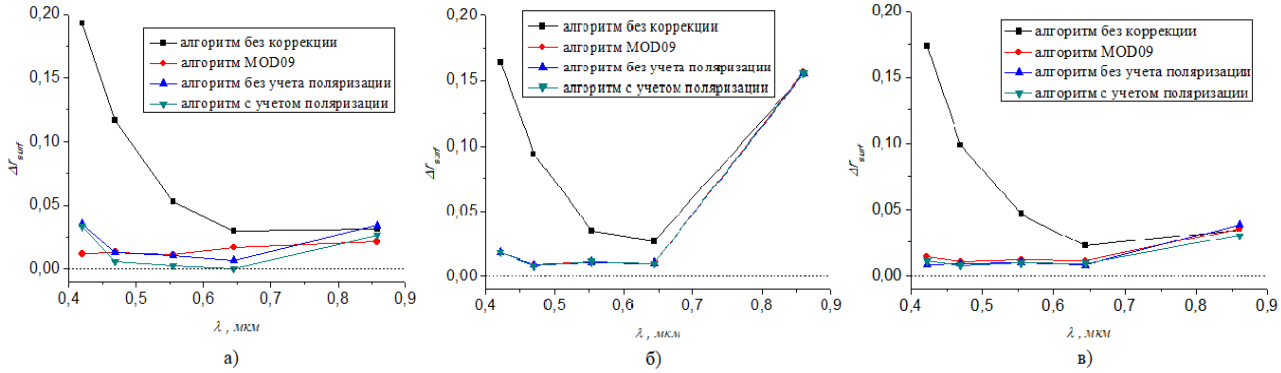


Рис. 1: Модули средних по рассмотренным снимкам отличий коэффициентов отражения земной поверхности, полученных разными подходами, от эталонных значений [31]; (а) — тестовая точка на юге Томской области, (б) — тестовая точка г. Москвы, (в) — тестовая точка Иркутской области

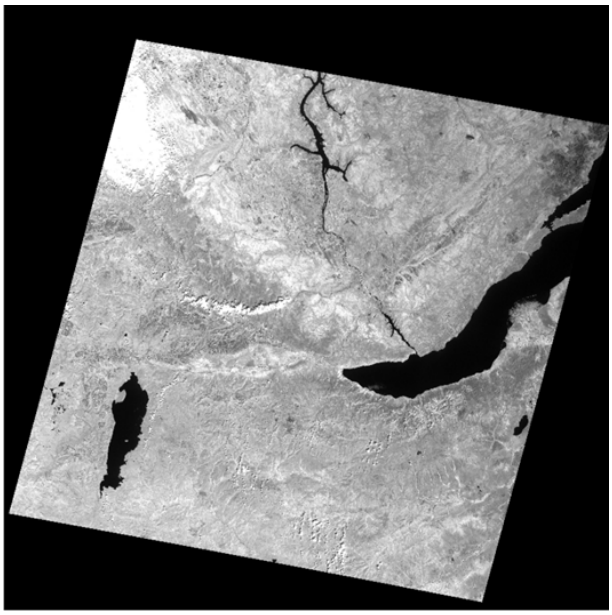


Рис. 2: Пример восстановленных коэффициентов отражения земной поверхности по данным МСУ-100

Анализ результатов показывает, что погрешность восстановления коэффициентов отражения по данным отечественных приборов получается несколько выше, чем для данных прибора MODIS, но, тем не менее, эти данные пригодны для решения широкого круга задач.

Заключение

В работе выполнена аппробация алгоритма восстановления коэффициентов отражения земной поверхности по спутниковым измерениям приборов MODIS, ГСА, Сангур и МСУ-100. Сопоставление с результатами MOD09 показывает, что разработанный алгоритм дает меньшие погрешности для рассмотренных ситуаций на длинах волн 0.469, 0.555 и 0.649 мкм, погрешность одного порядка на длине волны 0.860 мкм и большую погрешность на длине волны 0.420 мкм.

Список литературы

[1] Перенос радиации в рассеивающих и поглощающих атмосферах. Стандартные методы счета / ред. Ж. Ленобль, пер. Ж.К. Золотовой. Л: Гидрометеоиздат. 1990. 263с.

[2] Марчук Г.И., Михайлов Г.А., Назаралиев М.А., Дарбинян Р.А., Каргин Б.А., Елепов Б.С. Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука. Сибирское отд-ние, 1976. 284 с.

[3] Каргин Б.А. Статистическое моделирование поля солнечной радиации в атмосфере / под ред. Г.А. Михайлова. Новосибирск: Ротапринт ВЦ СО АН СССР. 1984. 206 с.

[4] Назаралиев М.А. Статистическое моделирование радиационных процессов в атмосфере. Новосибирск: Наука. Сибирское отделение, 1990. 227 с.

[5] Сушкевич Т.А. Математические модели переноса излучения. М: Бином. Лаборатория знаний, 2005. 661с.

[6] Будак В.П., Желтов В.С., Лубенченко А.В., Фрейдлин К.С., Шагалов О.В. Быстрый и точный алгоритм численного моделирования переноса излучения в мутной среде на основе метода синтетических итераций // Оптика атмосферы и океана. 2016. Т. 29, № 09. С. 739–746.

[7] Otterman J., Fraser R.S. Adjacency effects on imaging by surface reflection and atmospheric scattering: cross radiance to zenith // Applied optics. 1979. V. 18, iss. 16. P. 2852–2860.

[8] Vermote E.F., El Saleous N.Z., Justice C.O. Atmospheric correction of MODIS data in the visible to middle infrared: first results // Remote Sensing of Environment. 2002. V. 83, iss. 1–2. P. 97–111.

[9] Vermote E.F., Vermeulen A. Atmospheric correction algorithm: spectral reflectances (MOD09). Algorithm Theoretical Background document, version 4.0. 1999. [Электронный ресурс]: http:modis.gsfc.nasa.gov.atbd.atbd mod08.pdf

[10] Lyapustin Al., Wang Y. Parameterized code SHARM-3D for radiative transfer over inhomogeneous surfaces // Applied Optics. 2005. V. 44, iss. 35. P. 7602–7610.

[11] Белов В.В., Тарасенков М.В. О точности и быстродействии RTM-алгоритмов атмосферной коррекции спутниковых изображений в видимом и УФ-диапазонах // Оптика атмосферы и океана. 2013. Т. 26, № 07. С. 564–571.

[12] Тарасенков М.В., Белов В.В. Комплекс программ восстановления отражательных свойств земной поверхности в видимом и УФ-диапазонах // Оптика атмосферы и океана. 2014. Т. 27, № 07. С. 622–627.

[13] Dave J.V. Effect of Atmospheric Conditions on Remote Sensing of a Surface Nonhomogeneity // Photogrammetric Engineering and remote Sensing. 1980. V. 46, iss. 9. P. 1173–1180.

[14] Tanre D., Holben B.N., Kaufman Y.J. Atmospheric correction algorithm for NOAA-AVHRR products: theory and application // IEEE Transactions on Geoscience and Remote Sensing. 1992. V.30, iss. 2. P. 231–248.

[15] Егоров В.А., Барталев С.А. Метод радиометрической коррекции искажений отражательных характеристик земного покрова в данных спутниковых измерений, вызванных влиянием рельефа местности // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13, № 5. С. 192–201.

[16] Breon, F.-M., Vermote E. Correction of MODIS surface reflectance time series for BRDF effects // Remote Sensing of Environment. 2012. V. 125. P. 1–9.

[17] Зимовая А.В., Тарасенков М.В., Белов В.В. Влияние поляризации излучения на восстановление коэффициента отражения земной поверхности по спутниковым данным в видимом диапазоне длин волн // Оптика атмосферы и океана. 2017. Т. 30, № 11. С. 927–932.

[18] Зимовая А.В., Тарасенков М.В., Белов В.В. Учет поляризации при пассивном зондировании из космоса отражательных свойств земной поверхности // Оптика атмосферы и океана. 2015. Т. 28, № 12. С. 1085–1089.

[19] Гермогенова Т.А. О влиянии поляризации на распределение интенсивности рассеянного излучения // Изв. АН СССР. Сер. геофиз. 1962. №6. С. 854–856.

[20] Тарасенков М.В., Кирнос И.В., Белов В.В. Наблюдение земной поверхности из космоса через просвет в облачном поле // Оптика атмосферы и океана. 2016. Т. 29, № 09. С. 767–771.

[21] Marshak A., Wen G., Coakley J.A., Remer L.A., Loeb N.G., Cahalan R.F. A simple model for the cloud adjacency effect and the apparent bluing of aerosols near clouds // J. Geophys. Res. 2008. V. 113. D14S17.

[22] Leroy M., Deuze J., Breon F., Hautecoeur O., Herman M., Buriez J., Tanre D., Bouffies S., Chazette P., Roujean J. Retrieval of atmospheric properties and surface bidirectional reflectances over land from POLDER/ADEOS // J. Geophys. Res. 1997. V. 102. P. 17023–17037.

[23] Diner D.J., Martonchik J.V., Borel C., Gerstl S.A.W., Gordon H.R., Knyazikhin Y., Myneni R., Pinty B., Michel V.M. Multi-Angle Imaging Spectro-Radiometer Level 2 Surface Retrieval Algorithm Theoretical Basis; Jet Propulsion Laboratory: La Canada Flintridge, CA, USA, 2008.

[24] Tarasenkov M.V., Belov V.V., Engel M.V. Algorithm for reconstruction of the Earth surface reflectance from MODIS satellite measurements in a turbid atmosphere // Proceedings of SPIE. 2018. V. 10833. CID: 10833 16 [10833-58].

[25] Белов В.В., Тарасенков М.В., Пискунов К.П. Параметрическая модель солнечной дымки в видимой и УФ-области спектра // Оптика атмосферы и океана. 2010. Т. 23, № 04. С. 294–297.

[26] Белов В.В., Тарасенков М.В. Статистическое моделирование функции размытия точки в сферической атмосфере и критерий выделения зон изо-планарности изображений // Оптика атмосферы и океана. 2010. Т. 23, № 05. С. 371–377.

[27] Березин И.С., Жидков Н.П. Методы вычислений. М: Гос. из-во физ-мат литературы. 1959. Т. 2. 620 с.

[28] Aerosol Robotic Network (AERONET), http://aeronet.gsfc.nasa.gov

[29] Kneizys F.X., Shettle E.P., Anderson G.P., Abreu L.W., Chetwynd J.H., Selby J.E.A. Clough S.A., Gallery W.O. User guide to LOWTRAN-7. ARGL-TR-86-0177.ERP 2010 / Hansom AFB. MA 01731. 137 p.

[30] https://lpdaac.usgs.gov.data access

[31] Кринов Е.Л. Спектральная отражательная способность природных образований. Ленинград. Издательство Академии Наук СССР, 1947. 271 с.

[32] hitran.iao.ru

Тарасенков Михаил Викторович — к.ф.-м.н., ст. науч.сотр. Института оптики атмосферы им. В. Е. Зуева СО РАН;
e-mail: TMV@iao.ru;
Белов Владимир Васильевич — д.ф.-м.н., зав. лабораторией Института оптики атмосферы им. В. Е. Зуева СО РАН;
e-mail: Belov@iao.ru;
Энгель Марина Владимировна — к.т.н., науч. сотрудник Института оптики атмосферы им. В. Е. Зуева СО РАН;
e-mail: Angel@iao.ru.
Зимовая Анна Викторовна — мл. науч.сотр. Института оптики атмосферы им. В. Е. Зуева СО РАН;
e-mail: AVK@iao.ru;
Гриднев Юрий Владимирович — науч. сотрудник Института оптики атмосферы им. В. Е. Зуева СО РАН;
e-mail: Yuri@iao.ru;
Дата поступления — 30 апреля 2019 г.

ВАРИАЦИОННЫЕ МЕТОДЫ ПОСТРОЕНИЯ РАЗНОСТНЫХ УРАВНЕНИИ И ИТЕРАЦИОННЫХ АЛГОРИТМОВ ДЛЯ ПРИБЛИЖЕННОГО РЕШЕНИЯ ЗАДАЧИ ДИРИХЛЕ

Н. М. Темирбеков[1, 2], Ж. Р. Жаксылыкова[3]

[1] *Национальная Инженерная академия РК, Алматы, Казахстан*

[2] *Казахстанкий инженерно-технологический университет, Алматы, Казахстан*

[3] *Казахский Национальный педагогический университет имени Абая, Алматы, Казахстан*

УДК 519.6
DOI: 10.24411/9999-016A-2019-10079

Рассматривается задача Дирихле для уравнения Пуассона, описываемое начально-краевой задачей для эллиптического типа второго порядка. Для ее приближенного решения используются вариационные методы построения разностных уравнений и вариационные методы построения итерационных алгоритмов. Приведены результаты расчетов для модельной задачи.

Ключевые слова: эллиптическое уравнение второго порядка, задача Дирихле, разностная схема, метод Ритца, метод сопряженных градиентов.

Введение

В данной работе рассматривается эллиптическое дифференциальное уравнение. При решении:

1) для построения разностных уравнений был применен вариационный метод, предложенный в 1908 году немецким математиком В. Ритцем, который называется методом Ритца. Найденное этим методом решение $u_n(x)$ при некоторых условиях, стремится к точному решению $u(x)$, когда $n \to \infty$. Вопросы сходимости решений, получаемых методом Ритца, рассматриваются в многочисленных работах и монографиях.

2) Для построения итерационных алгоритмов был применен метод сопряженных градиентов, который среди известных итерационных методов, используемых для решения систем линейных алгебраических уравнений, выделяется своей эффективностью.

Приведены примеры расчетов для модельных задач. Результаты вычислительного эксперимента демонстрируют высокую эффективность предлагаемого итерационного метода.

1 Вариационные методы построения разностных уравнений

Рассмотрим в общем виде задачу в операторной форме $Lu = f$, $u \in \Phi(L)$ где $\Phi(L)$ — область определения оператора L. Данная задача эквивалентна соответствующей вариационной задаче $J(u) = \min\limits_{v \in \Phi(L)} J(v)$ где $J(v) = (Lv, v) - 2(f, u)$.

Рассмотрим применение метода Ритца в решении эллиптического дифференциального уравнения вида

$$-\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(A_{ij}(x)\frac{\partial u}{\partial x_j}\right) = f(x), \tag{1}$$

в D с краевым условием

$$u|_{\partial D} = 0 \tag{2}$$

ISBN 978-5-901548-42-4

где D ограниченная область с кусочно-линейной границей ∂D, $A_{ij}(x) = A_{ji}(x)$ ограниченные функции и для произвольного вектора $\xi = (\xi_1, \xi_2)^{/}$ выполняется неравенство $\mu_0 \sum_{i=1}^{2} \xi_i^2 \leq \inf\limits_{x \in D} \sum_{i,j=1}^{2} A_{ij}(x)\, \xi_i \xi_j \leq \sup\limits_{x \in D} \sum_{i,j=1}^{2} A_{ij}(x)\, \xi_i \xi_j \leq \mu_1 \sum_{i=1}^{n} \xi_i^2$ с положительными постоянными $\mu_0 \leq \mu_1$.

Можно показать, что оператор задач (1) и (2) является симметричным и положительно определенным, а сама задача сводится к нахождению функции, минимизирующей в пространстве $\mathring{W}_2^1(D)$ квадратичный функционал

$$J(u) = \int_D \left(\sum_{i,j=1}^{2} A_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \right) dD - 2 \int_D u f dD \tag{3}$$

Для нахождения приближенного решения задачи (3) применим метод Ритца с подпространствами F_h специального вида, удовлетворяющих условию (2)

Далее, построим подпространства F_h. Для упрощения изложения проиллюстрируем это на примере кусочно-линейных аппроксимаций, когда область $D = \{(x_1, x_2) : 0 < x_1, x_2 < 1\}$ является квадратом. Покроем эту область равномерной квадратной сеткой с шагом $h = \frac{1}{N+1}$ (N — положительное целое число), а затем осуществим триангуляцию области D. Пронумеруем все внутренние вершины треугольников через $\{p_{k,l}\}_{k,l=1}^{N}$, объединение всех треугольников имеющих точку своей вершиной $p_{k,l}$ обозначим через $D_{k,l}^h$. $D_{k,l}^h$ можно представить в виде объединения шести треугольников $\left\{D_{k,l,m}^h\right\}_{m=1}^{6}$, порядок нумерации которых указан на рис. 1.

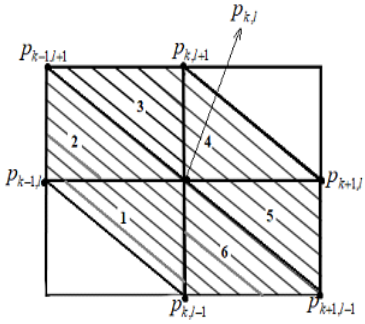


Рис. 1: Обозначения треугольников $\left\{D_{k,l,m}^h\right\}_{m=1}^{6}$

Каждому внутреннему узлу (x_k, y_l), $k, l = \overline{1, N}$ сетки поставим в соответствие кусочно-линейную базисную функцию $\varphi_{k,l}(x, y)$. Определяем каждую из этих функций $\varphi_{k,l}(x, y)$ для введенной сетки.

Чтобы задать $\varphi_{k,l}(x, y)$ аналитически, достаточно для каждого треугольника, входящего в носитель $\varphi_{k,l}(x, y)$, составить уравнение плоскости, проходящей через единицу в $@_{k,l}$, а в остальных его двух вершинах через нуль. Вычисляя базисные функции $\varphi_{k,l}(x, y)$ в каждом из треугольников $\left\{D_{k,l,m}^h\right\}_{m=1}^{6}$, построим систему базисных функций

$$\varphi_{k,l}(x, y) = \begin{cases} 1 - \frac{1}{h}(x_k - x) - \frac{1}{h}(y_l - y), & (x, y) \in D_{k,l,1}^h, \\ 1 - \frac{1}{h}(x_k - x), & (x, y) \in D_{k,l,2}^h, \\ 1 + \frac{1}{h}(y_l - y), & (x, y) \in D_{k,l,3}^h, \\ 1 + \frac{1}{h}(x_k - x) + \frac{1}{h}(y_l - y), & (x, y) \in D_{k,l,4}^h, \\ 1 + \frac{1}{h}(x_k - x), & (x, y) \in D_{k,l,5}^h, \\ 1 - \frac{1}{h}(y_l - y), & (x, y) \in D_{k,l,6}^h. \end{cases} \tag{4}$$

Для построения приближенного решения $u^h(x)$ задачи (1) и (2) воспользуемся методом Ритца с применением

базиса $\{\varphi_{k,l}(x)\}_{k,l=1}^{N}$

$$u^h(x) = \sum_{k,l=1}^{N} \alpha_{k,l}\varphi_{k,l}(x) \tag{5}$$

В результате приходим к системе линейных уравнений

$$A\alpha = g \tag{6}$$

где $\alpha = (\alpha_1, \alpha_2, ..., \alpha_{N^2})^T$ — вектор, составленный из коэффициентов $\left\{\alpha_{N(k-1)+l} = \alpha_{k,l}\right\}_{k,l=1}^{N}$ разложения, $g = (g_1, g_2, ..., g_{N^2})^T$ — вектор с компонентами

$$g_{N(k-1)+l} = g_{k,l} = \int_{D_{k,l}} f\varphi_{k,l}(x)dD,\ k,l = \overline{1,N} \tag{7}$$

и элементы матрицы A вычисляются по формулам

$$a_{k,l}^{i,j} = a_{N(k-1)+l,N(i-1)+j} = \int_D \sum_{s,t=1}^{2} A_{st}(x)\frac{\partial\varphi_{k,l}}{\partial x_s}\frac{\partial\varphi_{i,j}}{\partial x_t}dD, \quad k,l,i,j = \overline{1,N} \tag{8}$$

Матрица A является блочной трехдиагональной вида

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & 0 & ... & 0 & 0 \\ A_{2,1} & A_{2,2} & A_{2,3} & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & A_{N,N-1} & A_{N,N} \end{pmatrix} \tag{9}$$

где $_{k,k} =^*_{k,k}$, $_{k,k+1} =^*_{k+1,k}$, $k,l = \overline{1,N}$ и каждая из матриц $_{k,l}$ является трехдиагональной матрицей порядка N.

Рассмотрим задачу Дирихле в области D с границей ∂D с переменными коэффициентами $p(x,y)$ и $q(x,y)$

$$-\frac{\partial}{\partial x}p(x,y)\frac{\partial u}{\partial x} - \frac{\partial}{\partial y}q(x,y)\frac{\partial u}{\partial y} = f(x,y), \tag{10}$$

с краевым условием

$$u|_{\partial D} = 0 \tag{11}$$

Уравнение (10) умножаем на функцию $u(x,y)$и интегрируем по D учитывая граничное условие $u|_{\partial D} = 0$

$$\int_D p(x,y)\left(\frac{\partial u}{\partial x}\right)^2 dD + \int_D q(x,y)\left(\frac{\partial u}{\partial y}\right)^2 dD = \int_D f(x,y)\cdot u dD$$

Согласно (3) построим функционал

$$J(u) = \int_D p(x,y)\left(\frac{\partial u}{\partial x}\right)^2 dD + \int_D q(x,y)\left(\frac{\partial u}{\partial y}\right)^2 dD - 2\int_D f(x,y)\cdot u dD \tag{12}$$

Для применения метода Ритца $u(x,y)$ заменяем разложением $\sum_{k,l=1}^{N}\alpha_{k,l}\varphi_{k,l}(x)$

$$J\left(u^h(\alpha_{k,l})\right) = \int_D p(x,y)\left(\frac{\partial}{\partial x}\sum_{k,l=1}^{N}\alpha_{k,l}\varphi_{k,l}(x)\right)^2 dD + \int_D q(x,y)\left(\frac{\partial}{\partial y}\sum_{k,l=1}^{N}\alpha_{k,l}\varphi_{k,l}(x)\right)^2 dD -$$

$$-2\int_D f(x,y)\sum_{k,l=1}^{N}\alpha_{k,l}\varphi_{k,l}(x)\,dD, \quad k,l = 1,...,N$$

Далее, производные $\frac{\partial J\left(u^h(\alpha_{k,l})\right)}{\partial\alpha_{k,l}}$ $(k,l = 1,...,N)$ приравниваем нулю и $\alpha_{k,l}$ выносим за скобку

$$\left[\int_D p(x,y)\left(\sum_{k,l=1}^{N}\frac{\partial\varphi_{k,l}(x)}{\partial x}\right)\frac{\partial\varphi_{k,l}(x)}{\partial x}dD+\int_D q(x,y)\left(\sum_{k,l=1}^{N}\frac{\partial\varphi_{k,l}(x)}{\partial y}\right)\frac{\partial\varphi_{k,l}(x)}{\partial y}dD\right]\alpha_{k,l}- \tag{13}$$
$$-\int_D f(x,y)\,\varphi_{k,l}(x)\,dD=0,\quad k,l=\overline{1,N}.$$

Таким образом, согласно (8) и (9) вводя обозначение

$$a_{k,l}^{i,j}=a_{N(k-1)+l,N(i-1)+j}=\int_D p(x,y)\left(\sum_{k,l=1}^{N}\frac{\partial\varphi_{k,l}(x)}{\partial x}\right)\frac{\partial\varphi_{i,j}(x)}{\partial x}dD+ \tag{14}$$
$$+\int_D q(x,y)\left(\sum_{k,l=1}^{N}\frac{\partial\varphi_{k,l}(x)}{\partial y}\right)\frac{\partial\varphi_{i,j}(x)}{\partial y}dD,\ k,l,i,j=\overline{1,N},$$

получаем систему линейных алгебраических уравнений (6).

Учитывая симметричность и блочную трехдиагональность матрицы A, нам достаточно определить $a_{k,l}^{k,l},\ a_{k,l}^{k,l-1},\ a_{k,l}^{k-1,l},\ a_{k,l}^{k-1,l+1},\ k,l=\overline{1,N}$.

1. Согласно (12), где $i,j=k,l$ находим $a_{k,l}^{k,l}$:

$$a_{k,l}^{k,l}=\int_{D_{k,l}^h}\left[p(x,y)\left(\frac{\partial\varphi_{k,l}}{\partial x}\right)^2+q(x,y)\left(\frac{\partial\varphi_{k,l}}{\partial y}\right)^2\right]dD \tag{15}$$

Вычисляем интегралы в каждой из области $D_{k,l}^h$, учитывая, вид базисной функции $\varphi_{k,l}(x,y)$ в рассматриваемой области. Далее, все найденные шесть значения подставляем в (15). При этом, объединяем интегралы с одинаковыми значениями и используем обозначения

$$\tilde{P}_{k\pm\frac{1}{2},l}=\frac{1}{h^2}\int_{D_{k,l}^h\bigcap D_{k\pm1,l}^h}p(x,y)\,dxdy \tag{16}$$

$$\tilde{Q}_{k,l\pm\frac{1}{2}}=\frac{1}{h^2}\int_{D_{k,l}^h\bigcap D_{k,l\pm1}^h}q(x,y)\,dxdy \tag{17}$$

Значит,

$$a_{k,l}^{k,l}=\tilde{P}_{k-\frac{1}{2},l}+\tilde{Q}_{k,l-\frac{1}{2}}+\tilde{P}_{k+\frac{1}{2},l}+\tilde{Q}_{k,l+\frac{1}{2}}. \tag{18}$$

2. Для того чтобы найти элемент $a_{k,l}^{k,l-1}$ как и в первом случае воспользуемся формулой (14), где $i,j=k,l-1$ и для того чтобы определить место $p_{k,l-1}$ вершину $p_{k,l}$ спускаем на один шаг вниз. По рис.1 видим что, вершины $p_{k,l}$ и $p_{k,l-1}$ лежат в треугольниках D_1 и D_6. Значит, $a_{k,l}^{k,l-1}$ определяем только в этих областях

$$\begin{aligned}&a_{k,l}^{k,l-1}=\textstyle\int_{D_1\bigcup D_6}\left(p(x,y)\frac{\partial\varphi_{k,l}}{\partial x}\frac{\partial\varphi_{k,l-1}}{\partial x}+q(x,y)\frac{\partial\varphi_{k,l}}{\partial y}\frac{\partial\varphi_{k,l-1}}{\partial y}\right)dxdy=\\&=\textstyle\int_{D_1}\left(p(x,y)\frac{\partial\varphi_{k,l}}{\partial x}\frac{\partial\varphi_{k,l-1}}{\partial x}+q(x,y)\frac{\partial\varphi_{k,l}}{\partial y}\frac{\partial\varphi_{k,l-1}}{\partial y}\right)dxdy+\\&+\textstyle\int_{D_6}\left(p(x,y)\frac{\partial\varphi_{k,l}}{\partial x}\frac{\partial\varphi_{k,l-1}}{\partial x}+q(x,y)\frac{\partial\varphi_{k,l}}{\partial y}\frac{\partial\varphi_{k,l-1}}{\partial y}\right)dxdy\end{aligned} \tag{19}$$

Найденные значения интегралов, подставляя в (19) и используя обозначение (17) получаем

$$a_{k,l}^{k,l-1}=-\frac{1}{h^2}\int_{D_1}q(x,y)dxdy-\frac{1}{h^2}\int_{D_6}q(x,y)dxdy=-\frac{1}{h^2}\cdot\int_{D_1\bigcup D_6}q(x,y)dxdy=-\tilde{Q}_{k,l-\frac{1}{2}} \tag{20}$$

3. Чтобы найти $a_{k,l}^{k-1,l}$ по формуле $a_{k,l}^{i,j} = \int_D p(x,y)\frac{\partial\varphi_{k,l}(x)}{\partial x}\frac{\partial\varphi_{i,j}(x)}{\partial x}dD + \int_D q(x,y)\frac{\partial\varphi_{k,l}(x)}{\partial y}\frac{\partial\varphi_{i,j}(x)}{\partial y}dD$, где $i,j = k-1,l$ передвигаем вершину $p_{k,l}$ на один шаг влево и по рис.1 видим, что $p_{k,l}$ и $p_{k-1,l}$ являются вершинами треугольников D_1 и D_2. Значит, $a_{k,l}^{k-1,l}$ определяется в треугольниках D_1 и D_2.

$$\begin{aligned} a_{k,l}^{k-1,l} &= \int_{D_1 \bigcup D_2}\left(p(x,y)\frac{\partial\varphi_{k,l}}{\partial x}\frac{\partial\varphi_{k-1,l}}{\partial x} + q(x,y)\frac{\partial\varphi_{k,l}}{\partial y}\frac{\partial\varphi_{k-1,l}}{\partial y}\right)dxdy = \\ &= \int_{D_1}\left(p(x,y)\frac{\partial\varphi_{k,l}}{\partial x}\frac{\partial\varphi_{k-1,l}}{\partial x} + q(x,y)\frac{\partial\varphi_{k,l}}{\partial y}\frac{\partial\varphi_{k-1,l}}{\partial y}\right)dxdy + \\ &+ \int_{D_2}\left(p(x,y)\frac{\partial\varphi_{k,l}}{\partial x}\frac{\partial\varphi_{k-1,l}}{\partial x} + q(x,y)\frac{\partial\varphi_{k,l}}{\partial y}\frac{\partial\varphi_{k-1,l}}{\partial y}\right)dxdy. \end{aligned} \tag{21}$$

Найденные значения $a_{k,l}^{k-1,l}$ в треугольниках D_1 и D_2 подставляем в (21) с учетом (16) и получим

$$a_{k,l}^{k-1,l} = -\frac{1}{h^2}\cdot\int_{D_1} p(x,y)dxdy - \frac{1}{h^2}\cdot\int_{D_2} p(x,y)dxdy = -\frac{1}{h^2}\cdot\int_{D_1\bigcup D_2} p(x,y)dxdy = -\tilde{P}_{k-\frac{1}{2},l} \tag{22}$$

4. Осталось определить элемент $a_{k,l}^{k-1,l+1}$. Для этого вершину $p_{k,l}$ нужно будет с начало передвигать на один шаг влево, затем на один шаг поднять вверх. Рассуждая по рис.1 видим, что $p_{k,l}$ и $p_{k-1,l+1}$ являются вершинами треугольников D_2 и D_3. Но при выбранном сдвиге $p_{k,l}$ вершина $p_{k-1,l+1}$ не будут принадлежать ни одному из шести треугольников. Поэтому

$$a_{k,l}^{k-1,l+1} = 0 \tag{23}$$

Все найденные значения (18), (20), (22), (23) подставляя в (6) с учетом симметричности матрицы A получаем

$$\begin{aligned} &\left(\tilde{P}_{k-\frac{1}{2},l} + \tilde{Q}_{k,l-\frac{1}{2}} + \tilde{P}_{k+\frac{1}{2},l} + \tilde{Q}_{k,l+\frac{1}{2}}\right)\alpha_{k,l} - \tilde{Q}_{k,l-\frac{1}{2}}\alpha_{k,l-1} - \tilde{Q}_{k,l+\frac{1}{2}}\alpha_{k,l+1} - \tilde{P}_{k-\frac{1}{2},l}\alpha_{k-1,l} - \\ &-\tilde{P}_{k+\frac{1}{2},l}\alpha_{k+1,l};\ k,l = \overline{1,N} \end{aligned} \tag{24}$$

Пусть в (24)

$$(A_1\alpha)_{k,l} = -\tilde{P}_{k-\frac{1}{2},l}\alpha_{k-1,l} + \left(\tilde{P}_{k-\frac{1}{2},l} + \tilde{P}_{k+\frac{1}{2},l}\right)\alpha_{k,l} - \tilde{P}_{k+\frac{1}{2},l}\alpha_{k+1,l};\ k,l = \overline{1,N} \tag{25}$$

$$(A_2\alpha)_{k,l} = -\tilde{Q}_{k,l-\frac{1}{2}}\alpha_{k,l-1} + \left(\tilde{Q}_{k,l-\frac{1}{2}} + \tilde{Q}_{k,l+\frac{1}{2}}\right)\alpha_{k,l} - \tilde{Q}_{k,l+\frac{1}{2}}\alpha_{k,l+1};\ k,l = \overline{1,N} \tag{26}$$

тогда, используя $A_1 + A_2 =$ получим систему (6).

2 Вариационные методы построения итерационных алгоритмов

Для численного решения системы (6) можно применять методы вариационного типа, такие как, метод скорейшего спуска, метод минимальных поправок, метод сопряженных градиентов и т.д.

Метод сопряженных градиентов является наиболее предпочтительным для систем с самосопряженной положительной матрицей $A = A^* > 0$. При плохой обусловленности матрицы этот метод не всегда становится вычислительно устойчивым.

Оператор A является самосопряженным и положительно определенным оператором в пространстве сеточных функций H_h^0. В H_h^0 введем скалярное произведение $(u,v) = \sum_{i,j=1}^{N-1} u_{i,j}v_{i,j}h_1h_2$и норма $\|u\| = \sqrt{(u,u)}$.

Для решения уравнения (6) воспользуемся методом сопряженных градиентов. Итерационный процесс реализуется в следующем порядке:

0. Подготовка перед итерационным процессом.
   По заданному $\alpha_{i,j}^0$ вычисляется невязка $s_{i,j}^0 = r_{i,j}^0 = \left(A\alpha^0\right)_{i,j} - g_{i,j}$, $i = 1,...,n_1,\ j = 1,...,n_2$;

k-я итерация метода:

1. Вычисляется параметр: $\gamma_k = \frac{\left(r_{i,j}^{k-1}, r_{i,j}^{k-1}\right)}{\left(As_{i,j}^k, s_{i,j}^{k-1}\right)}$, $r_{i,j}^k = r_{i,j}^{k-1} - \gamma_k As_{i,j}^k$

2. Следующее приближение решения вычисляется по формуле: $\alpha_{i,j}^k = \alpha_{i,j}^{k-1} - \gamma_k s_{i,j}^k$

3. Вычисляется параметр: $\beta_k = \frac{\left(r_{i,j}^k, r_{i,j}^k\right)}{\left(r_{i,j}^{k-1}, r_{i,j}^{k-1}\right)}$

4. Вспомогательное значение вычисляется по формуле:$s_{i,j}^{k+1} = r_{i,j}^k + \beta_n s_{i,j}^k$ где

$$\left(As^k\right)_{i,j} = \frac{1}{h_1^2}\left[\tilde{P}_{i+\frac{1}{2},j}\left(s_{i+1,j} - s_{i,j}\right) - \tilde{P}_{i-\frac{1}{2},j}\left(s_{i,j} - s_{i-1,j}\right)\right] + \\ + \frac{1}{h_2^2}\left[\tilde{Q}_{i+\frac{1}{2},j}\left(s_{i,j+1} - s_{i,j}\right) - \tilde{Q}_{i,j-\frac{1}{2}}\left(s_{i,j} - s_{i,j-1}\right)\right];\ i,j = \overline{1,N}$$

Этот процесс продолжается до тех пор, пока не выполнится критерий остановки итераций $\left\|\alpha^{k+1} - \alpha^k\right\| \le \varepsilon$.

3 Примеры расчетов

Для иллюстрации предложенного метода рассмотрим пример задачи (10)–(11) в круге.

Пусть в прямоугольной области D имеется круг Ω радиуса R с центром (A_1, A_2). Требуется найти приближенное решение задачи Дирихле для уравнения Пуассона

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -x_1 \cdot x_2 \tag{27}$$

в круге, при условии $u\left(x_1, x_2\right)|_{\partial\Omega} = 0$.

Вспомогательная задача метода фиктивных областей

$$\begin{cases} \frac{\partial}{\partial x_1}\left(p\left(x_1,x_2\right)\frac{\partial u^\varepsilon}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(q\left(x_1,x_2\right)\frac{\partial u^\varepsilon}{\partial x_2}\right) = -x_1 \cdot x_2, \\ u^\varepsilon\left(x_1,x_2\right)|_{\partial D} = 0 \end{cases}$$

$$p\left(x_1,x_2\right) = q\left(x_1,x_2\right) = \begin{cases} 1, & \left(x_1,x_2\right) \in \Omega, \\ 1/\varepsilon, & \left(x_1,x_2\right) \in D/\Omega. \end{cases}, f^\varepsilon\left(x_1,x_2\right) = \begin{cases} -x_1 \cdot x_2, & \left(x_1,x_2\right) \in \Omega \\ 0, & \left(x_1,x_2\right) \in D/\Omega. \end{cases}$$

Были проведены численные эксперименты. Ниже приводится таблица 1 для числа итераций N и погрешности вычисления $\|u_h - u\|_C$ при $n_1 = n_2 = 101$ и различных значениях ε для метода Ритца.

Таблица 1: Результаты численных экспериментов

n_1	n_2	$1/\varepsilon$	N	$\|u_h - u\|_C$
101	101	10^2	1221	0.003888
101	101	10^3	3395	0.003820
101	101	10^4	7920	0.003819
101	101	10^5	12737	0.003818
101	101	10^6	20042	0.003817
101	101	10^7	26592	0.003816

Полученные результаты показывают, что при увеличении числа узлов сетки погрешность решения уменьшается.

На рис. 2 приведен результат решения задачи по явной разностной схеме при $\varepsilon_* = \varepsilon = 10^{-4}$ на равномерной сетке с размером 101×101. Значит, исходная система линейных алгебраических уравнений имеет 101×101 неизвестных.

Заключение

На сегодняшний день остается актуальной задача разработки и модификации численных методов. Однако процесс развития вычислительной техники смещает упор с создания новых численных методов, на исследование и классификацию старых, с целью выявлению лучших. Теперь для современных мощных компьютеров, такие характеристики как объём требуемой памяти, и число арифметических операции, не обязательно

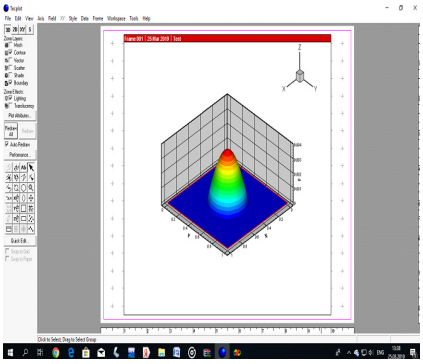


Рис. 2: Решение задачи Дирихле для уравнения Пуассона в круге

стоят на первом плане. Более предпочтительными становятся те методы, которые отличаются удобством реализации на ЭВМ, и позволяют решать более широкий класс задач.

Особыми достоинствами данного метода являются его простота и низкие затраты памяти, что делает его эффективным при решении задач большой размерности.

Приведенные результаты вычислительного эксперимента подтверждают работоспособность предлагаемого метода решения задачи Дирихле для уравнения Пуассона и ее достаточно высокую эффективность.

Список литературы

[1] Самарский А.А. Теория разностных схем — М.: Наука, 1983. — 616 с.

[2] Марчук Г.И. Методы вычислительной математики — М.: Наука, 1989. — 608 с.

[3] Мухамбетжанов А. Т., Отелбаев М. О., Смагулов Ш. С. Об одном новом приближенном методе решение нелинейных краевых задач.// Вычислительные технологии, Том 3, №4, 1998.

[4] Н.М. Темирбеков, Ж.Р. Жаксылыкова. Об итерационном методе фиктивных областей для решения эллиптического уравнения в областях со сложной геометрией // Известия МКТУ им. Х.А.Ясави, Серия математика, физика, информатика. — 2018. — Т.1. — №1, С.128–132.

[5] N.M.Temirbekov, Zh.R.Zhaksylykova. AnIterative Method for Solving Non linear Navier-Stokes Equationsin Complex Domains Taking into Account Boundary Conditions with Uniform Accuracy // AIP Conference Proceedings — 2018. —V.1997, 020036 (2018); doi: 10.1063/1.5049030.

[6] Temirbekov A.N., Wójcik W. Numerical Implementation of the Fictitious Domain Method for Elliptic Equations//International Journal of Electronics and Telecommunications.—2014. —Vol.60, Issue 3, —P.219-223.

[7] Temirbekov A.N. Numerical implementation of the method of fictitious domains for elliptic equations // 3rd International Conference on Analysis and Applied Mathematics, ICAAM 2016.—Vol. 1759, —P. 020053-1–020053-6; doi: 10.1063/1.4959667.

Темирбеков Нурлан Муханович — д.ф.-м.н., профессор
вице-президент Национальной Инженерной академии РК;
e-mail: temirbekov@rambler.ru;
Жаксылыкова Жадра Рахметовна — докторант 2-курса по специальности "6D060100–Математики"
Казахского национального педагогического университета имени Абая;
email: zhaksylykova0507@mail.ru
Дата поступления — 30 апреля 2019 г.

ДВУХСЕТОЧНЫЙ МЕТОД ДЛЯ НЕЛИНЕЙНОГО СИНГУЛЯРНО ВОЗМУЩЕННОГО УРАВНЕНИЯ ВТОРОГО ПОРЯДКА С ДВУМЯ ПАРАМЕТРАМИ

С. В. Тиховская

Институт математики им. С. Л. Соболева СО РАН, 630090, Новосибирск

УДК 519.624.2
DOI: 10.24411/9999-016A-2019-10080

Рассмотрена краевая задача для нелинейного сингулярно возмущенного обыкновенного дифференциального уравнения второго порядка с двумя малыми параметрами. Для линеаризации используются итерации Ньютона и Пикара. Для решения задачи на каждой итерации применяется разностная схема второго порядка на сетке Шишкина, которая сходится равномерно по обоим малым параметрам. Для уменьшения требуемого количества арифметических операций для решения разностной схемы предлагается каскадный двухсеточный метод. Для повышения точности разностной схемы, применяется экстраполяция Ричардсона. Получена оценка первой производной решения дифференциальной задачи в случае различных малых параметров. Обсуждаются результаты некоторых численных экспериментов.

Ключевые слова: сингулярно возмущенная задача, два малых параметра, сетка Шишкина, двухсеточный метод.

Введение

Известно, что применение классических разностных схем для сингулярно возмущенных задач приводит к большим погрешностям при малых значениях параметров возмущения. Равномерная сходимость разностной схемы для таких задач может быть достигнута подгонкой схемы к погранслойной составляющей [1, 2] или сгущением сетки в пограничных слоях [3, 4]. Рассматривается сингулярно возмущенная краевая задача для нелинейного обыкновенного дифференциального уравнения второго порядка с двумя малыми параметрами, влияющих на конвекционный и диффузионный члены, на кусочно-равномерной сетке. Разностные схемы высокого порядка точности для сингулярно возмущенных краевых задач очень важны, например [5–8].

Двухсеточный метод исследован в [8–14] и других работах, в том числе для решения нелинейных сингулярно возмущенных краевых задач в [8–12]. Для повышения ε-равномерной точности разностной схемы на сетке Шишкина в двухсеточном методе практически без дополнительных вычислений применяется экстраполяция Ричардсона [6, 15, 16].

В [17] построена монотонная разностная схема второго порядка точности для линейного сингулярно возмущенного обыкновенного дифференциального уравнения второго порядка с двумя малыми параметрами, влияющими на конвекционный и диффузионный члены. Показано, что эта схема на сетке Шишкина сходится равномерно по обоим малым параметрам.

В [18] рассмотрено нелинейное сингулярно возмущенное обыкновенное дифференциальное уравнения второго порядка с малыми параметрами при второй производной ε^2 и первой производной ε.

В [19] предложен двухсеточный метода с экстраполяцией Ричардсона для нелинейного сингулярно возмущенного обыкновенного дифференциального уравнения второго порядка с двумя малыми параметрами, основанного на разностной схеме второго порядка точности на сетке Шишкина.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (код проекта 18-31-00487).

ISBN 978-5-901548-42-4

Обозначения: Пусть $\|f\| = \max\limits_{x\in\overline{\Omega}} |f(x)|$ – норма непрерывной функции, а $\|f^N\|_N = \max\limits_{0\leqslant i\leqslant N} |f_i^N|$ – норма для сеточной функции. Пусть $[u]_\Omega$ – проекция функции $u(x)$ на сетку Ω. Здесь C, иногда с индексом, обозначает положительную константу, которая не зависит от малых параметров ε, μ и размера шага сетки.

1 Постановка задачи

Рассмотрим краевую задачу:

$$\begin{aligned} &Lu(x) = \varepsilon u''(x) + \mu a(x)u'(x) = f(x,u(x)), \qquad x\in\Omega=(0,1),\\ &u(0)=A, \quad u(1)=B, \end{aligned} \tag{1}$$

где функции a и f достаточно гладкие и

$$0<\varepsilon\leqslant 1,\ 0\leqslant\mu\leqslant 1,\ a(x)\geqslant\alpha>0,\ f'_u(x,u)\geqslant\beta>0,\ (x,u)\in\Omega\times R. \tag{2}$$

Решение $u(x)$ в общем случае имеет два пограничных слоя вблизи $x=0$ и $x=1$.

Если условия (2) выполнены, то решение задачи (1) равномерно ограничено по ε и μ:

$$\|u\|\leqslant C_0 = \beta^{-1}\|f(x,0)\| + \max\{|A|,|B|\}.$$

Предположим, что в дополнение к (2) выполнены следующие условия:

$$\beta^*\geqslant f'_u(x,u)\geqslant\beta>0, \quad (x,u)\in\Omega\times R, \quad \gamma=\min_{x\in\overline{\Omega}}\{f'_u(x,u)/a(x)\}. \tag{3}$$

Запишем задачу (1) в следующем виде:

$$\begin{aligned} &Lu(x) = \varepsilon u''(x) + \mu a(x)u'(x) - b(x)u(x) = f(x,0), \qquad x\in\Omega=(0,1),\\ &u(0)=A, \quad u(1)=B, \end{aligned}$$

где $b(x)=f'_u(x,\theta\,u)$, $0<\theta<1$.

Тогда верна следующая теорема.

Теорема 1. *Пусть $u(x)$ – решение задачи (1) с достаточно гладкими коэффициентами и правой частью. Тогда для некоторой константы C_1, не зависящей от ε и μ, выполнено:*

$$\begin{aligned} &|u'(x)|\leqslant C_1\left(1+\left(\mu/\varepsilon+1/\sqrt{\varepsilon}\right)\left(e^{-l_0x/\varepsilon}+e^{-l_1(1-x)/\varepsilon}\right)\right), \qquad 0\leqslant x\leqslant 1,\\ &l_0=\frac{1}{2}\left(\sqrt{\mu^2\alpha^2+4\,\beta_*\,\varepsilon}+\mu\,\alpha\right), \quad l_1=\frac{1}{2}\left(\sqrt{\mu^2(\alpha^*)^2+4\,\beta_*\,\varepsilon}-\mu\,\alpha^*\right). \end{aligned} \tag{4}$$

где $\alpha^*\geqslant a(x)\geqslant\alpha>0$, $f'_u(x,u)-\mu\,a'(x)\geqslant\beta_*>0$, $|f'_x(x,u)|\leqslant C$.

Доказательство. Дифференцируя (1), получим

$$\begin{aligned} &\varepsilon\nu''(x) + \mu a(x)\nu'(x) - B(x)\nu(x) = f'_x(x,u), \qquad x\in\Omega=(0,1),\\ &\nu(0)=u'(0), \quad \nu(0)=u'(1), \end{aligned}$$

где $\nu(x)=u'(x)$, $B(x)=f'_u(x,u)-\mu\,a'(x)$.

В соответствие с Принципом сравнения

$$|\nu(x)|\leqslant w(x),$$

где

$$\begin{aligned} &\varepsilon w''(x) + \mu a(x)w'(x) - \beta_* w(x) = -|f'_x(x,u)|, \qquad x\in\Omega=(0,1),\\ &w(0)=|u'(0)|, \quad w(0)=|u'(1)|. \end{aligned}$$

Рассмотрим декомпозицию на регулярную и погранслойные составляющие

$$w(x) = R(x) + p(x) + q(x),$$

где

$$\varepsilon p''(x)+\mu a(x)p'(x)-\beta_* p(x)=0, \qquad x\in\Omega=(0,1),$$
$$p(0)=|u'(0)|, \quad p(0)=0.$$

$$\varepsilon q''(x)+\mu a(x)q'(x)-\beta_* q(x)=0, \qquad x\in\Omega=(0,1),$$
$$q(0)=0, \quad q(0)=|u'(1)|.$$

$$\varepsilon R''(x)+\mu a(x)R'(x)-\beta_* R(x)=-|f'_x(x,u)|, \qquad x\in\Omega=(0,1),$$
$$w(0)=0, \quad R(0)=0.$$

Учитывая, что аналогично [17] выполнено

$$|u'(x)|\leqslant C_1\left(\frac{\mu}{\varepsilon}+\frac{1}{\sqrt{\varepsilon}}\right)\max\{\|u\|,\ f(x,0)\}, \qquad 0\leqslant x\leqslant 1,$$

и, выбирая подходящие барьерные функции, в соответствие с Принципом максимума, можно показать, что

$$|p(x)|\leqslant C\left(\frac{\mu}{\varepsilon}+\frac{1}{\sqrt{\varepsilon}}\right)e^{-l_0\,x/\varepsilon}, \qquad |q(x)|\leqslant C\left(\frac{\mu}{\varepsilon}+\frac{1}{\sqrt{\varepsilon}}\right)e^{-l_1\,x/\varepsilon}, \qquad |R(x)|\leqslant\frac{1}{\beta_*}\|f'_x(x,\,u)\|,$$

где $x\in\overline{\Omega}=[0,1]$.

Объединяя полученные оценки, получим требуемое. □

В соответствие с [4,17] зададим сетку:

$$\Omega_N=\{x_i: x_i=x_{i-1}+h_i, \quad x_0=0, \quad x_N=1, \quad i=1,2,\ldots,N\}, \tag{5}$$

где

$$h_i=\frac{4\sigma_1}{N},\ 1\leqslant i\leqslant N/4,\ h_i=\frac{2(1-\sigma_1-\sigma_2)}{N},\ N/4<i\leqslant 3\,N/4,\ h_i=\frac{4\sigma_2}{N},\ 3\,N/4<i\leqslant N,$$
$$\sigma_1=\begin{cases}\min\left\{\frac{1}{4},\frac{4\sqrt{\varepsilon}}{\sqrt{\gamma\,\alpha}}\ln N\right\}, & \mu^2\leqslant\frac{\gamma\,\varepsilon}{\alpha},\\ \min\left\{\frac{1}{4},\frac{4\,\varepsilon}{\mu\,\alpha}\ln N\right\}, & \mu^2\geqslant\frac{\gamma\,\varepsilon}{\alpha}.\end{cases}\quad \sigma_2=\begin{cases}\min\left\{\frac{1}{4},\frac{4\sqrt{\varepsilon}}{\sqrt{\gamma\,\alpha}}\ln N\right\}, & \mu^2\leqslant\frac{\gamma\,\varepsilon}{\alpha},\\ \min\left\{\frac{1}{4},\frac{4\,\mu}{\gamma}\ln N\right\}, & \mu^2\geqslant\frac{\gamma\,\varepsilon}{\alpha}.\end{cases}.$$

Рассмотрим конечно-разностную схему, использующую схемы центральных разностей, направленных разностей и средней точки на сетке Шишкина (5):

$$\begin{aligned}&L^N u_j^N=r_j^-\,u_{j-1}^N+r_j^c\,u_j^N+r_j^+\,u_{j+1}^N=Q^N(f(x_j,u_j^N)), \quad 0<j<N,\\ &u_0^N=A, \quad u_N^N=B,\end{aligned} \tag{6}$$

где

$$L^N\equiv L^N_{cd}, \qquad \text{если } 1\leqslant j<\frac{N}{4},$$

$$L^N\equiv\begin{cases}L^N_{cd}, & \sigma_1=0.25,\\ L^N_{mp}, & \sigma_1<0.25\ and\ \beta^*(1-\sigma_1-\sigma_2)/N<\mu\,\alpha,\\ L^N_{up}, & otherwise.\end{cases} \qquad \text{если } j=\frac{N}{4},$$

$$L^N\equiv\begin{cases}L^N_{cd}, & \mu\,\|a\|(1-\sigma_1-\sigma_2)/N<\varepsilon,\\ L^N_{mp}, & \mu\,\|a\|(1-\sigma_1-\sigma_2)/N\geqslant\varepsilon\ and\ \beta^*(1-\sigma_1-\sigma_2)/N<\mu\,\alpha,\\ L^N_{up}, & \mu\,\|a\|(1-\sigma_1-\sigma_2)/N\geqslant\varepsilon\ and\ \beta^*(1-\sigma_1-\sigma_2)/N\geqslant\mu\,\alpha.\end{cases} \qquad \text{если } \frac{N}{4}<j<\frac{3\,N}{4},$$

$$L^N\equiv\begin{cases}L^N_{cd}, & \sigma_2=0.25,\ and\ 2\,\mu\|a\|\sigma_2<\mu\,\alpha,\\ L^N_{mp}, & \sigma_2=0.25\ and\ 2\,\mu\|a\|\sigma_2\geqslant\mu\,\alpha,\\ L^N_{mp}, & \sigma_2<0.25\ and\ 2\,\beta^*\sigma_2/N<\mu\,\alpha,\\ L^N_{up}, & otherwise.\end{cases} \qquad \text{если } j=\frac{3\,N}{4},$$

$$L^N\equiv\begin{cases}L^N_{cd}, & 2\,\mu\,\|a\|\sigma_2/N<\varepsilon,\\ L^N_{mp}, & 2\,\mu\,\|a\|\sigma_2/N\geqslant\varepsilon.\end{cases} \qquad \text{если } \frac{3\,N}{4}<j\leqslant N-1.$$

Здесь используем обозначения из [17]:

$$L^N_{cd}u^N_j \equiv \varepsilon\,\delta^2\,u^N_j + \mu\,a_j\,D^0u^N_j = f(x_j,u^N_j),$$
$$L^N_{up}u^N_j \equiv \varepsilon\,\delta^2\,u^N_j + \mu\,a_j\,D^+u^N_j = f(x_j,u^N_j),$$
$$L^N_{mp}u^N_j \equiv \varepsilon\,\delta^2\,u^N_j + \mu\,\overline{a}_j\,D^+u^N_j = \overline{f}(x_j,u^N_j),$$

где $a_j = a(x_j)$, $\overline{z}_j = (z_j + z_{j+1})/2$ и

$$\delta^2\,u^N_j = \frac{1}{\hbar_j}\left(\frac{u^N_{j+1}-u^N_j}{h_{j+1}} - \frac{u^N_j-u^N_{j-1}}{h_j}\right), \qquad D^0u^N_j = \frac{u^N_{j+1}-u^N_{j-1}}{h_j+h_{j+1}}, \qquad D^+u^N_j = \frac{u^N_{j+1}-u^N_j}{h_{j+1}}.$$

В соответствие с [17] предполагаем, что

$$N\,(\ln N)^{-1} > 8\,\max\{\|a\|/\alpha,\,\beta^*/\,(\alpha\,\gamma)\}\,.$$

Решение разностной схемы (6) может быть получено на основе итераций. Рассмотрим линеаризацию Пикара:

$$L^{m,N}u^{m+1,N}_j = r^-_j\,u^{m+1,N}_{j-1} + \left(r^c_j - \beta^*\right)\,u^{m+1,N}_j + r^+_j\,u^{m+1,N}_{j+1} = Q^N(f(x_j,u^{m,N}_j) - \beta^*\,u^{m,N}_j),\ 0<j<N,$$

$$u^{m+1,N}_0 = A, \quad u^{m+1,N}_N = B, \qquad m \geqslant 0.$$

Аналогично рассмотрим линеаризацию Ньютона.

$$r^-_j\,u^{m+1,N}_{j-1} + \left(r^c_j - f'_u(x_j,u^{m,N}_j)\right)\,u^{m+1,N}_j + r^+_j\,u^{m+1,N}_{j+1} = Q^N(f(x_j,u^{m,N}_j) - f'_u(x_j,u^{m,N}_j)u^{m,N}_j), \quad 0<j<N,$$

$$u^{m+1,N}_0 = A, \quad u^{m+1,N}_N = B, \qquad m \geqslant 0.$$

В случае, когда задача (1) линейна: $f(x,u) = b(x)u + g(x)$, $b(x) \geqslant \beta > 0$, в соответствие с [17] верна следующая теорема.

Теорема 2. *Пусть $u(x)$ – решение линейной задачи (1) с достаточно гладкими коэффициентами и правой частью, и пусть u^N – решение разностной схемы (6). Тогда на сетке Шишкина (5) для некоторой константы C выполнено:*

$$\|\,[u]_{\Omega_N} - u^N\,\|_N \leqslant C\,\Delta_N = \begin{cases} C\,\ln^3 N/N^2, & \gamma\,\varepsilon < \alpha\,\mu^2 \\ C\,\ln^2 N/N^2, & \gamma\,\varepsilon \geqslant \alpha\,\mu^2. \end{cases} \tag{7}$$

2 Двухсеточный метод с экстраполяцией Ричардсона

Для уменьшения необходимого количества арифметических действий для решения разностной схемы предложен двухсеточный метод. В соответствие с идеей двухсеточного алгоритма [9] сначала задача (1) решается на грубой сетке. Затем, найденное сеточное решение интерполируется в узлы исходной сетки и используется как начальное приближение для следующих итераций. Это приводит к сокращению числа итераций на исходной сетке, а, следовательно, и к уменьшению количества арифметических действий. Отметим, что интерполяционная формула должна быть равномерна точная по малому параметру, иначе точность найденного сеточного решения может быть потеряна.

Таким образом, пусть Ω_n – сетка Шишкина, соответствующая (5) и содержащая n сеточных интервалов, где $n \ll N$. Сначала задача (1) предварительно решается на сетке Ω_n. Итерации Пикара или Ньютона на сетке Ω_n выполняются пока условие

$$\|u^{m_n,n} - u^n\|_n \leqslant \Delta_n.$$

не будет выполнено. Тогда сеточное решение $u^{m_n,n}$, найденное на сетке Ω_n, интерполируется в узлы исходной сетки Ω_N с помощью подходящей интерполяции, которая равномерно точна по ε и μ, и для некоторой константы C

$$\|Int([u]_{\Omega_n},x) - u(x)\| \leqslant C\,\Delta_n.$$

Следовательно, для некоторой константы C

$$\|[u]_{\Omega_n} - u^{m_n,n}\|_n \leqslant C\Delta_n.$$

Теперь определим начальное приближение для итераций на исходной сетке Ω_N, и тогда для некоторой константы C

$$\|u^{0,N} - u\|_N \leqslant C\Delta_n, \qquad u^{0,N} = [Int(u^{m_n,n}, x)]_{\Omega_N}.$$

Таким образом, используя итерации на грубой сетке и подходящую интерполяцию, построено начальное приближение $u^{0,N}$ для итераций на исходной сетке Ω_N с точностью $O(\Delta_n)$. Далее итерации выполняются на сетке Ω_N пока не будет достигнута точность $O(\Delta_N)$.

Для повышения точности разностной схемы в двухсеточном методе можно применить экстраполяцию Ричардсона [6, 10–12, 15, 16]. Поэтому сетка $\overline{\Omega}_n$ должна быть с теми же значениями параметров σ_1 и σ_2 как и сетка $\overline{\Omega}_N$. Таким образом эти сетки вложены, так что $\overline{\Omega}_n = \{X_j\} \subset \overline{\Omega}_N = \{x_i\}$. Пусть $N = kn$, где k – некоторое целое число. Тогда очевидно, что исходная сетка $\overline{\Omega}_N$ может быть получена из грубой сетки $\overline{\Omega}_n$ делением каждого интервала на k равных частей.

Пусть u^n – решение разностной схемы на сетке $\overline{\Omega}_n$. В соответствие с методом Ричардсона определим u^{Nn} на сетке $\overline{\Omega}_N$. Сначала зададим u^{Nn} на сетке $\overline{\Omega}_n$ как

$$u^{Nn}(X_j) = k_n u^n(X_j) + k_N u^N(X_j), \qquad X_j \in \overline{\Omega}_n,$$

где $k_n = -n^2/(N^2 - n^2) = -1/(k^2 - 1)$, $k_N = N^2/(N^2 - n^2) = k^2/(k^2 - 1)$.

В узлах исходной сетки, не совпадающих с узлами грубой сетки, определим $u^{Nn}(x_i)$ для любого $x_i \in \overline{\Omega}_N$, используя подходящую интерполяцию. Заметим, что интерполяционная формула должна быть равномерно точна по параметрам ε и μ, иначе точность найденного сеточного решения может быть потеряна [10–12, 20].

Таким образом, построено сеточное решение u^{Nn} на сетке $\overline{\Omega}_N$ с использованием экстраполяции Ричардсона.

3 Результаты численных экспериментов

Рассмотрим следующую краевую задачу:

$$\begin{aligned} &\varepsilon\, u'' + 2\,\mu\, u' = 4\, u - 1 + u e^u + g(x), \quad 0 < x < 1, \\ &u(0) = 0, \quad u(1) = 1, \end{aligned}$$

где $g(x)$ соответствует точному решению:

$$u(x) = 0.25 - \frac{1 + 3\, e^{-\lambda_1/\varepsilon}}{4\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)}\, e^{-\lambda_0\, x/\varepsilon} + \frac{3 + e^{-\lambda_0/\varepsilon}}{4\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)}\, e^{-\lambda_1\,(1-x)/\varepsilon},$$

где $\lambda_0 = \sqrt{\mu^2 + 4\,\varepsilon} + \mu$, $\lambda_1 = \sqrt{\mu^2 + 4\,\varepsilon} - \mu$.

Здесь $\alpha = \alpha^* = \|a\| = 2$, $\beta = \beta_* = 4$, $\gamma = 2$, $l_0 = \lambda_0$, $l_1 = \lambda_1$, $\beta^* = 4 + 2\,e$.

Тогда

$$\begin{aligned} u'(x) = {} & \frac{\lambda_0}{\varepsilon}\, \frac{1 + 3\, e^{-\lambda_1/\varepsilon}}{4\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)}\, e^{-\lambda_0\, x/\varepsilon} + \frac{\lambda_1}{\varepsilon}\, \frac{3 + e^{-\lambda_0/\varepsilon}}{4\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)}\, e^{-\lambda_1\,(1-x)/\varepsilon} = \\ = {} & \frac{\mu}{4\,\varepsilon} \left(e^{-\lambda_0\, x/\varepsilon}\, \frac{\left(1 - e^{-(\lambda_0+\lambda_1)(1-x)/\varepsilon}\right)}{\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)} + 3\, e^{-\lambda_1\,(1-x)/\varepsilon}\, \frac{\left(1 - e^{-(\lambda_0+\lambda_1)\, x/\varepsilon}\right)}{\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)} \right) + \\ & + \frac{1}{8}\, \frac{(\lambda_0 + \lambda_1)/\varepsilon}{\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)} \left(e^{-\lambda_0\, x/\varepsilon} \left(1 + 3\, e^{-\lambda_1/\varepsilon}\right) + e^{-\lambda_1\,(1-x)/\varepsilon} \left(3 + e^{-\lambda_0/\varepsilon}\right) \right), \end{aligned}$$

заметим, что $u'(x) > 0$, $0 \leqslant x \leqslant 1$, следовательно, получаем $0 \leqslant u(x) \leqslant 1$. Учитывая, что

$$\frac{x}{1 - e^{-x}} \leqslant 1 + x,\ x > 0, \quad 0 \leqslant \frac{\left(1 - e^{-(\lambda_0+\lambda_1)(1-x)/\varepsilon}\right)}{\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)} \leqslant 1, \quad 0 \leqslant \frac{\left(1 - e^{-(\lambda_0+\lambda_1)\, x/\varepsilon}\right)}{\left(1 - e^{-(\lambda_0+\lambda_1)/\varepsilon}\right)} \leqslant 1,\ 0 \leqslant x \leqslant 1,$$

получаем, что

$$|u'(x)| \leqslant \frac{\mu}{4\,\varepsilon}\left(e^{-\lambda_0 x/\varepsilon} + 3\,e^{-\lambda_1 (1-x)/\varepsilon}\right) + \frac{1}{2}\left(1 + \frac{\lambda_0+\lambda_1}{\varepsilon}\right)\left(e^{-\lambda_0 x/\varepsilon} + e^{-\lambda_1 (1-x)/\varepsilon}\right) \leqslant$$
$$\leqslant \left(e^{-\lambda_0 x/\varepsilon} + e^{-\lambda_1 (1-x)/\varepsilon}\right)\left(\frac{3}{4}\,\frac{\mu}{\varepsilon} + \frac{\sqrt{\mu^2+4\,\varepsilon}}{\varepsilon} + \frac{1}{2}\right) \leqslant 2\left(1 + \left(\frac{\mu}{\varepsilon} + \frac{1}{\sqrt{\varepsilon}}\right)\left(e^{-\lambda_0 x/\varepsilon} + e^{-\lambda_1 (1-x)/\varepsilon}\right)\right),$$

что соответствует полученной оценке (4).

Принимая во внимание погрешность схемы (7) останавливаем итерации на сетке $\overline{\Omega}_N$ когда выполнено следующее условие:

$$\|L^N u^{m,N} - f^N\|_N \leqslant \beta\Delta_N,$$

где β соответствует (2). Тогда

$$\|u^{m,N} - u^N\|_N \leqslant \beta^{-1}\|L^N u^{m,N} - f^N\|_N \leqslant \Delta_N.$$

В табл. 1 в верхней строке приведено число итераций двухсеточного метода в случае линеаризации Пикара (слева) и линеаризации Ньютона (справа) при $\mu = 2^{-5}$ на сетке $\overline{\Omega}_N$ в случае $n = N/2$ для различных значений N и ε. Число итераций на сетке $\overline{\Omega}_n$ дано в скобках. Число итераций односеточного метода в зависимости от N дано в нижней строке таблицы.

Таблица 1: Число итераций односеточного и двухсеточного методов при $\mu = 2^{-5}$ в случае линеаризации Пикара (слева) и линеаризации Ньютона (справа).

ε	N			
	64	256	1024	8192
1	1(4)	1(5)	1(6)	1(8)
	4	5	7	8
2^{-4}	2(4)	1(6)	2(7)	1(10)
	5	7	8	10
2^{-9}	4(4)	3(6)	3(8)	3(10)
	5	7	9	11
2^{-14}	3(3)	4(6)	5(8)	5(10)
	5	7	9	11

ε	N			
	64	256	1024	8192
1	1(2)	1(3)	1(3)	1(3)
	3	3	3	3
2^{-4}	1(3)	1(3)	1(3)	1(4)
	3	3	3	4
2^{-9}	2(3)	1(3)	1(3)	1(4)
	3	3	3	4
2^{-14}	2(2)	2(3)	1(3)	1(4)
	3	3	4	4

Таким образом, основная часть итераций в двухсеточном методе выполняется на грубой сетке. Это приводит к существенному уменьшению числа арифметических действий.

В табл. 2 приведена норма погрешности для односеточного метода (слева) и двухсеточного метода с экстраполяцией Ричардсона (справа) в случае $n = N/2$ для различных значений N и ε при $\mu = 1$.

Таблица 2: Норма погрешности для односеточного (слева) и для двухсеточного (справа) методов при $\mu = 1$

ε	N			
	64	256	1024	8192
2^{-4}	$2.59e{-3}$	$1.64e{-4}$	$1.01e{-5}$	$1.58e{-7}$
2^{-7}	$1.31e{-2}$	$1.30e{-3}$	$1.26e{-4}$	$3.31e{-6}$
2^{-9}	$1.32e{-2}$	$1.31e{-3}$	$1.26e{-4}$	$3.33e{-6}$
2^{-14}	$1.32e{-2}$	$1.31e{-3}$	$1.26e{-4}$	$3.34e{-6}$

ε	N			
	64	256	1024	8192
2^{-4}	$3.71e{-4}$	$2.01e{-6}$	$3.57e{-7}$	$6.89e{-9}$
2^{-7}	$5.08e{-3}$	$6.92e{-5}$	$5.63e{-7}$	$2.64e{-9}$
2^{-9}	$5.16e{-3}$	$7.25e{-5}$	$5.68e{-7}$	$2.10e{-9}$
2^{-14}	$5.19e{-3}$	$7.23e{-5}$	$5.99e{-7}$	$1.44e{-9}$

В табл. 3 приведена норма погрешности для односеточного метода (слева) и двухсеточного метода с экстраполяцией Ричардсона (справа) в случае $n = N/2$ для различных значений N и ε при $\mu = 2^{-5}$.

Из табл. 2–3 следует, что применение экстраполяции Ричардсона в двухсеточном методе повышает точность разностной схемы до порядка $O(\ln^3 N/N^3)$ равномерно по обоим малым параметрам.

Список литературы

[1] Ильин А. М. Разностная схема для дифференциального уравнения с малым параметром при старшей производной // Математические заметки. 1969. Т. 6, № 2. С. 237–248.

Таблица 3: Норма погрешности для односеточного (слева) и для двухсеточного (справа) методов при $\mu = 2^{-5}$

ε	N				ε	N			
	64	256	1024	8192		64	256	1024	8192
2^{-4}	$1.85e{-4}$	$1.01e{-5}$	$3.62e{-7}$	$3.37e{-9}$	2^{-4}	$9.30e{-5}$	$2.04e{-6}$	$3.11e{-7}$	$2.57e{-9}$
2^{-7}	$1.20e{-3}$	$7.38e{-5}$	$4.14e{-6}$	$7.24e{-8}$	2^{-7}	$1.30e{-4}$	$2.44e{-6}$	$2.23e{-8}$	$2.55e{-9}$
2^{-9}	$4.12e{-3}$	$2.66e{-4}$	$1.66e{-5}$	$2.62e{-7}$	2^{-9}	$6.75e{-4}$	$2.47e{-6}$	$2.63e{-7}$	$1.56e{-9}$
2^{-14}	$1.62e{-2}$	$2.03e{-3}$	$1.04e{-4}$	$2.18e{-6}$	2^{-14}	$5.43e{-3}$	$7.22e{-4}$	$4.08e{-4}$	$7.15e{-9}$

[2] Емельянов К. В. Разностная схема для трехмерного эллиптического уравнения с малым параметром при старшей производной // Краевые задачи для уравнений математической физики. Екатеринбург: УрО РАН, 1973. С. 30–42.

[3] Бахвалов Н. С. К оптимизации методов решения краевых задач при наличии пограничного слоя // Журнал вычислительной математики и математической физики. 1969. Т. 9, № 4. С. 841–859.

[4] Шишкин Г. И. Сеточные аппроксимации сингулярно возмущенных эллиптических и параболических уравнений. Екатеринбург: УрО РАН, 1992.

[5] Roos H.-G., Stynes M., Tobiska L. Robust Numerical Methods for Singularly Perturbed Differential Equations. Volume 24 of Springer series in Computational Mathematics. Berlin: Springer-Verlag, 2008.

[6] Shishkin G. I., Shishkina L. P. Difference Methods for Singular Perturbation Problems. Volume 140 of Monographs and Surveys in Pure and Applied Mathematics. Boca Raton: Chapman & Hall/CRC, 2009.

[7] Kopteva N., Linss T. Uniform second-order pointwise convergence of a central difference approximation for a quasilinear convection-diffusion problem // Journal of Computational and Applied Mathematics. 2001. V. 137, № 2. P. 257–267.

[8] Angelova I. T., Vulkov L. G. Comparison of the Two-grid Method on Different Meshes for Singularly Perturbed Semilinear Problems. // AIP Conference Proceedings. 2008. V. 1067. P. 305–312.

[9] Vulkov L. G., Zadorin A. I. Two-grid algorithms for an ordinary second order equation with exponential boundary layer in the solution // International Journal of Numerical Analysis and Modeling. 2010. V. 7, № 3. P. 580–592.

[10] Задорин А. И., Тиховская С. В. Решение нелинейного сингулярно возмущенного уравнения второго порядка на основе схемы Самарского // Сиб. журн. вычисл. матем. 2013. Т. 16, № 1. С. 11–25.

[11] Задорин А. И., Тиховская С. В. Двухсеточный метод для нелинейной сингулярно возмущенной краевой задачи на сетке Шишкина // Сиб. журн. индустр. матем. 2013. Т. 16, № 1(53). С. 42–55.

[12] Tikhovskaya S. V., Zadorin A. I. A two-grid method with Richardson extrapolation for a semilinear convection-diffusion problem // AIP Conference Proceedings. 2015. V. 1684. P. 090007-1–090007-8.

[13] Axelsson O., Layton W. A two-level discretization of nonlinear boundary value problems // SIAM J. Numer. Anal. 1996. V. 33. P. 2359–2374.

[14] Xu J. A novel two-grid method for semilinear elliptic equation // SIAM J. Sci. Comput. 1994. V. 15. P. 231–237.

[15] Natividad M. C., Stynes M. Richardson extrapolation for a convection-diffusion problem using a Shishkin mesh // Applied Numerical Mathematics. 2003. V. 45, № 2-3. P. 315–329.

[16] Шишкин Г. И., Шишкина Л. П. Метод Ричардсона высокого порядка точности для квазилинейного сингулярно возмущенного эллиптического уравнения реакции-диффузии // Дифференциальные уравнения. 2005. Т. 41, № 7. С. 980-–989.

[17] Gracia J. L., O'Riordan E., Pickett M. L. A parameter robust second order numerical method for a singularly perturbed two-parameter problem // Applied Numerical Mathematics. 2006. V. 56, № 7. P. 962–980.

[18] Çakir M., Amiraliev G. M. Numerical solution of a singularly perturbed three-point boundary value problem // International Journal of Computer Mathematics. 2007. V. 84, № 10. P. 1465—1481.

[19] Tikhovskaya S. V., Korbut M. F. Two-grid algorithm for the solution of singularly perturbed two-parameter problem on Shishkin mesh // Journal of Physics: Conference Series. 2019. V. 1210. P. 012142-1–012142-8.

[20] Задорин А. И. Метод интерполяции на сгущающейся сетке для функции с погранслойной составляющей // Журнал вычислительной математики и математической физики. 2008. Т. 48, № 9. С. 1673–1684.

Тиховская Светлана Валерьевна — к.ф.-м.н., ст. науч.сотр. Института математики им. С. Л. Соболева СО РАН;
e-mail: s.tihovskaya@yandex.ru
Дата поступления — 31 мая 2019 г.

ВЛИЯНИЕ САМООРГАНИЗАЦИИ ПОВЕРХНОСТНЫХ ЗАРЯДОВ НА КВАНТОВЫЙ МИКРОКОНТАКТ

О. А. Ткаченко[1], Д. Г. Бакшеев[2,3], О. П. Сушков[4], В. А. Ткаченко[1,3,5]

[1] *Институт физики полупроводников им. А.В.Ржанова СО РАН, 630090, Новосибирск*

[2] *Яндекс, 630090, Новосибирск*

[3] *Новосибирский государственный университет, 630090, Новосибирск*

[4] *School of Physics, University of New South Wales, Sydney, 2052, Australia*

[5] *Новосибирский государственный технический университет, 630073, Новосибирск*

УДК 004.942:621.382-022.532:530.145
DOI: 10.24411/9999-016A-2019-10081

Моделированием трехмерной электростатики и двумерного квантового транспорта исследуются микроконтакты, создаваемые в затворно-индуцированном двумерном газе электронов, либо дырок. При небольших дистанциях от двумерного газа до поверхности полупроводника и верхнего металлического затвора (30, 60 нм) необходимо учитывать влияние беспорядочно расположенных на границе полупроводника с диэлектриком локализованных зарядов. Расчетом найдено, что при полном беспорядке в координатах зарядов квантование кондактанса достаточно длинных контактов (>400 нм) искажается резонансами, которые возникают из-за когерентного рассеяния носителей на флуктуациях потенциала. Самоорганизация поверхностных зарядов подавляет эти флуктуации даже при большой температуре беспорядка (500 К). Проведено сравнение расчетов и измерений кондактанса дырочных микроконтактов длиной 100 и 600 нм.

Ключевые слова: наноустойство, трехмерная электростатика, поверхностный заряд, самоорганизация, температура беспорядка, затворно-индуцированный двумерный газ, электронный микроконтакт, дырочный микроконтакт, двумерный квантовый транспорт, резонансы, квантование кондактанса.

Введение

В настоящее время исследуются свойства микроконтакта, создаваемого в конфигурации сплошной металлический затвор-диэлектрик-расщепленный затвор Шоттки — *нелегированная* гетероструктура GaAs/AlGaAs [1,2]. Такой микроконтакт или квантовый точечный контакт формируется в затворно-индуцированном двумерном газе (ДГ) электронов, либо дырок (рис. 1). При соответствующем напряжении на верхнем затворе V_{TG} эти носители вытягиваются электрическим полем из омического контакта и заполняют при низкой температуре (~ 1 К) рабочий слой GaAs. В свою очередь напряжение на расщепленном затворе V_{SG} полностью вытесняет двумерный газ под его половинами, оставляя короткий квантовый канал между двумя морями подвижных носителей. Интересно рассмотреть сочетание электрофизических и квантовых свойств таких устройств с учетом зарядки точечных дефектов на поверхности полупроводника, роль которой может быть существенной при небольшой толщине диэлектрика ($d = 20 - 30$ нм) и глубине залегания рабочего слоя под поверхностью полупроводника ($z = 30 - 60$ нм). Структуры с такими параметрами d, z и дырочным микроконтактом длиной от 100 до 600 нм изучаются экспериментально нашими коллегами в University of New South Wales, Australia.

Заметим, что для планарно-однородных структур металл-диэлектрик-нелегированный полупроводник задача о самоорганизации зарядов на поверхности полупроводника решалась нами аналитически для всей плоскости, либо численно на достаточно большом прямоугольнике с соответствующими граничными условиями. Учтено, что точечные заряды на поверхности полупроводника возникают в термодинамическом равновесии при температуре выше комнатной и нейтрализованы зарядами изображения в верхнем затворе. Использовалась самая простая модель, в которой беспорядок в системе поверхностных зарядов определялся лишь основными параметрами структуры и температурой в равновесном состоянии. Аналитические и

Работа выполнена при поддержке Российского научного фонда (проект 19-72-30023).

ISBN 978-5-901548-42-4

численные результаты оказались в хорошем согласии в широком диапазоне температур определяющих беспорядок. Моделирование позволило объяснить результаты измерений транспортного и квантового времен рассеяния в структурах с двумерным электронным и дырочным газом. Для планарно-однородных структур аналитические решения были возможны благодаря тому, что изменения потенциала в двумерном газе, вызванные беспорядком в поверхностных зарядах, были малыми в сравнении с энергией Ферми, Фурье образ флуктуаций потенциала имел аксиальную симметрию и достаточно было знать его поведение в длинноволновом пределе.

Эти условия нарушены в геометрически сложной структуре с микроконтактом (рис. 1a). Присутствие в структуре затвора Шоттки усложняет задачу о самоорганизации, поскольку под этим затвором заряды на поверхности полупроводника (тонкого защитного слоя GaAs) полностью экранированы примыкающим металлом (Ti/Au). В данном случае поверхностные заряды при решении задачи методом Монте-Карло заполняют прямоугольник с двумя встречными вырезами в местах, где находятся половинки расщепленного затвора Шоттки (SG) (рис. 1b). По заданной трехмерной геометрии твердотельной структуры и найденному положению поверхностных зарядов численно решается трехмерное уравнение Пуассона в приближении Томаса-Ферми. Далее определяется эффективный двумерный потенциал для электронов/дырок ДГ и численно решается задача двумерного квантового рассеяния баллистической частицы с заданной полной энергией. Параметры модели устройства задаются структурой и условиями эксперимента.

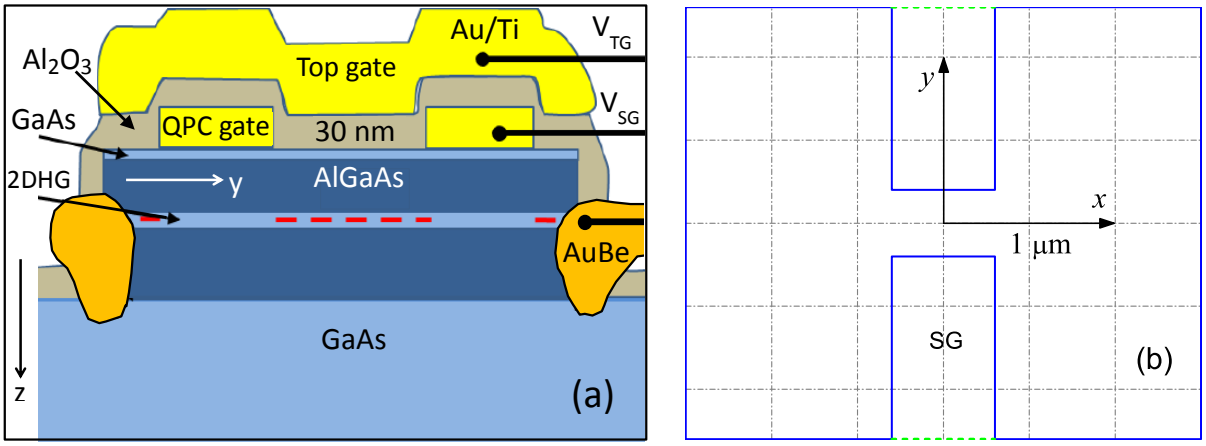


Рис. 1: Схематическое изображение структуры с микроконтактом в затворно-индуцированном двумерном дырочном газе (2DHG): (a) — разрез по вертикали, указаны верхний затвор (Top gate), диэлектрик (Al_2O_3), расщепленный затвор (QPC gate), тонкий защитный слой (GaAs) и омические контакты (AuBe), красным пунктиром отмечены места присутствия дырок в рабочем слое GaAs; (b) —вид сверху на поверхность тонкого защитного слоя GaAs, области SG контактируют с расщепленным затвором.

1 Моделирование самоорганизации поверхностных зарядов

Концентрация точечных зарядов n_σ на границе между диэлектриком Al_2O_3 и тонким защитным слоем GaAs найдена из решения задачи одномерной электростатики для планарно-однородной структуры, имеющей в термодинамическом равновесии общий уровень Ферми в омических контактах к рабочему слою, в верхнем затворе Au/Ti и во всем нелегированном полупроводнике. В таком случае электрическое поле в полупроводнике практически отсутствует, и из теоремы Гаусса получается

$$n_\sigma = -\frac{\epsilon_0 \epsilon_{ins} A_e}{qd}, \tag{1}$$

где q — заряд заряженной точки, d — толщина слоя изолятора, A_e — разница работ выхода электронов в вакуум из полупроводника и из металла, ϵ_0 — диэлектрическая постоянная, ϵ_{ins} — диэлектрическая проницаемость изолятора: для Al_2O_3 $\epsilon_{ins} \approx 8$. В данном случае металлом служит Ti, поэтому $A_e \approx 0.3$ эВ, что дает $n_\sigma = 7.5 \cdot 10^{11}$ см$^{-2}$ при $q = -e$ (0<e — элементарный заряд).

Для разных температур методом Метрополиса [3, 4] вычислялись равновесные состояния системы поверхностных зарядов на прямоугольнике длины 3 мкм и ширины 2.6 мкм. Учитывалось, что часть площади

прямоугольника занята затвором Шоттки (области SG на рис. 1b). В моделируемых образцах длина расщепленного затвора Шоттки менялась от 100 до 600 нм, а дистанция между его половинами равна 400 нм. Считалось, что под этим затвором нет поверхностных зарядов (экранирование металлом, нанесенным на GaAs). На остальной площади прямоугольника $s = S_{rect} - S_{SG}$ предполагались сохранение числа заряженных частиц ($N = n \cdot s \approx 50000$) и континуальность возможных перемещений точечных зарядов при поиске наиболее вероятного значения внутренней энергии системы. Учитывался вклад в полную энергию взаимодействия от зарядов изображения в верхнем затворе. Считалось, что найденное алгоритмом алгоритмом Метрополиса для некоторой температуры беспорядка T_{dis} распределение поверхностных зарядов замораживается при погружении образца в криостат при нулевых затворных напряжениях [5]. Значения T_{dis} от 500 K до 1000 K были найдены из сравнения теоретических и экспериментальных значений транспортного и квантового времен рассеяния в разных планарно-однородных структурах с затворно-индуцированным ДГ электронов или дырок. Здесь мы используем $T_{dis} = 500$ K. Для этого значения вычисленные равновесные распределения поверхностных зарядов мало чем отличались от случая планарно-однородной структуры, кроме мест непосредственной близости к расщепленному затвору, т.е. тонких (< 50 нм) полосок, окаймляющих области SG на рис. 1b. Мы считали, что границы с расщепленным затвором являются непроницаемыми для поверхностных зарядов, и зарядам энергетически выгодно выстраиваться вдоль границы.

2 Моделирование электростатики устройства с микроконтактом

С вычисленным распределением поверхностного заряда решались задачи трехмерной электростатики (3D) и двумерного (2D) квантового транспорта для микроконтакта при затворных напряжениях V_{TG}, V_{SG} отвечающих формированию двумерного газа носителей нужной плотности и микроконтакта с заданным числом мод [2, 6].

Электрический потенциал $\phi(x, y, z)$ вне металла описывается 3D-уравнением Пуассона: $\nabla\epsilon\nabla\phi = -qn/\epsilon_0$, где $n(\boldsymbol{r})$ — объемная концентрация отрицательных ($q = -e$), либо положительных ($q = e$) зарядов, $\boldsymbol{r} = (x, y, z)$, ϵ — диэлектрическая проницаемость материала. При расчете зонных диаграмм удобно перейти к потенциальной энергии электрона $U = -e\phi$:

$$\nabla\epsilon\nabla U = n(\boldsymbol{r})qe/\epsilon_0. \tag{2}$$

На границе GaAs с изолятором считается известным распределение двумерной плотности поверхностного заряда $n_\sigma(x, y)$, дающее скачок z-компоненты электрической индукции [2, 6]:

$$\frac{\epsilon_{GaAs}dU}{dz}|_{+0} - \frac{\epsilon_{ins}dU}{dz}|_{-0} = n_\sigma(x, y)qe/\epsilon_0. \tag{3}$$

Плотность $n_\sigma(x, y)$ находится из вычисленного распределения поверхностных зарядов: точечные заряды $q = -e$ заменяются однородными квадратиками со стороной 2.5 нм, и их центры смещаются в ближайшие точки прямоугольной расчетной сетки, заданной шагом $h_x = h_y = 2.5$ нм и покрывающей область счета — прямоугольник 3×2.6 мкм^2 из рис. 1b.

В зависимости от типа рассматриваемых подвижных носителей мы считаем, что в слоях GaAs потенциал U совпадает с дном зоны проводимости E_c, либо потолком валентной зоны E_v в GaAs. В слое с ДГ концентрация подвижных носителей ($n > 0$) вычисляется в подходе Томаса-Ферми, т.е. в локальном приближении по двумерной квазиклассической плотности состояний с учетом знака q :

$$n(\boldsymbol{r}) = \Psi(z)^2 n_{2D}(\boldsymbol{\rho}) = \Psi(z)^2 \frac{qm^*}{e\pi\hbar^2} \left[U_{\text{eff}}(\boldsymbol{\rho}) - E_F\right]. \tag{4}$$

Здесь $\Psi(z)^2$ — плотность вероятности для основного состояния в квантовой яме по z, $\boldsymbol{\rho} = (x, y)$, $m^* = 0.067m_e$, либо $m^* = 0.3m_e$ — эффективные массы электрона и дырки в GaAs, $E_F \equiv 0$ — уровень Ферми в ДГ (2DG), U_{eff} — эффективный самосогласованный потенциал. Формула (4) применяется, когда $U_{\text{eff}} - E_F > 0$ для дырок и наоборот для электронов, иначе $n(\boldsymbol{r}) = 0$. Здесь $U_{\text{eff}} = U(x, y, z_{2DG}) \pm E_0$, E_0 — основной уровень в квантовой яме по z, знак плюс и $E_0 = 0.02$ эВ относятся к электронному ДГ, а знак минус и $E_0 = 0.01$ эВ — к дырочному ДГ: значения отвечают тому, что в моделируемой структуре ДГ находится в квантовой яме толщиной 16 нм. Заметим, что надежность метода Томаса-Ферми проверена моделированием разных двумерных наносистем с беспорядком [7–11].

Для решения уравнения Пуассона ставятся следующие дополнительные условия. Потенциал в металлических затворах при моделировании электронного, либо дырочного микроконтакта задается как

$$U^e_{TG} = -eV^e_{Tg} + W_e,$$
$$U^e_{SG} = -eV^e_{SG} + E_g/2 + \delta_M, \quad (5)$$
$$U^h_{TG} = -eV^h_{Tg} + W_e - E_g,$$
$$U^h_{SG} = -eV^h_{SG} - E_g/2 + \delta_M. \quad (6)$$

Здесь $E_g = 1.52$ эВ — ширина запрещенной зоны в GaAs, $W_e \approx 0.26$ эВ — разность работы выхода из металла (Ti) и электронного сродства GaAs, $\delta_M \approx 0.15$ эВ — учитывает сдвиг центра запрещенной зоны GaAs относительно общего уровня Ферми на поверхности GaAs и в прилегающем металле.

На боковых плоских поверхностях, ограничивающих область расчетов (прямоугольник 3×2.6 мкм^2), ставятся периодические гранусловия. Однородным предполагается замороженный заряд на границе эпитаксиальной гетероструктуры с полуизолирующей подложкой GaAs, расположенной ниже слоя с ДГ на 1 мкм. Концентрация этого заряда и условие отсутствия электрического поля в подложке определяют производную потенциала U по нормали к нижней плоской границе области расчетов.

Решением задачи 3D электростатики мы нашли, что особенности в расположении поверхностных зарядов порождают на глубине двумерного газа максимумы или минимумы потенциала характерного размера 200 нм (рис. 2).

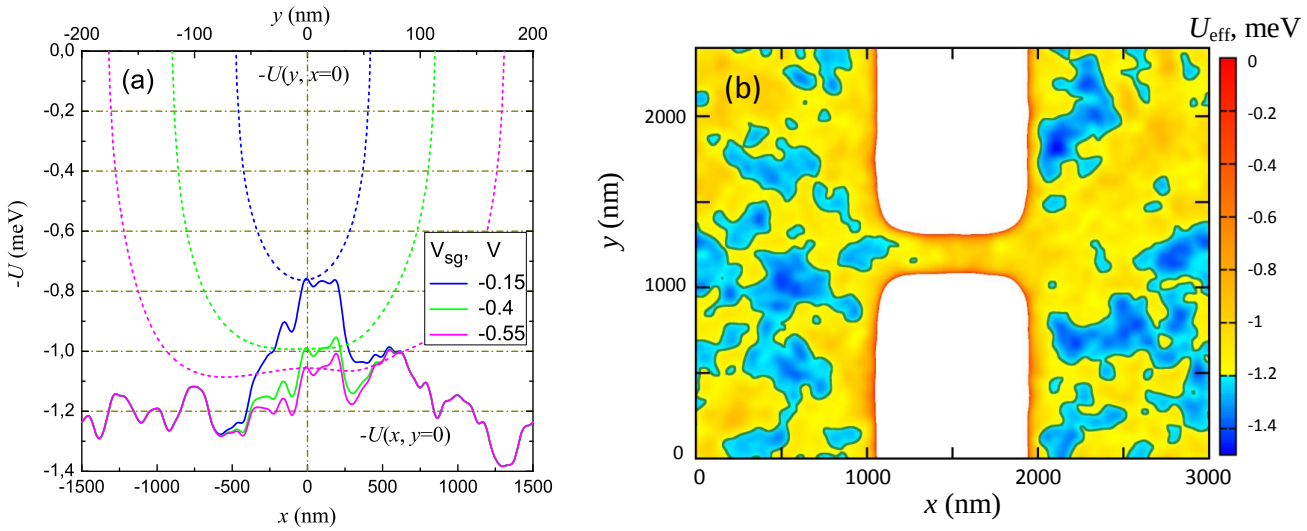


Рис. 2: Пример поведения электростатического потенциала $U(x, y)$ на глубине 48 нм от поверхности гетероструктуры GaAs/AlGaAs, т.е. в середине квантовой ямы с дырочным ДГ, имеющем при $V_{TG} = -1.1$ В плотность $n_h = 1.5 \cdot 10^{11}$ см$^{-2}$. Распределение локализованных зарядов вычислено при температуре беспорядка $T_{dis} = 500$ К, длина расщепленного затвора $L = 600$ нм. (a) — Зависимости $-U(x, y = 0)$, $-U(y, x = 0)$ при $U > E_F = 0$ и указанных V_{SG}; (b) — эффективный потенциал при $U_{\text{eff}} < E_F = 0$ и $V_{SG} = -0.4$ В.

3 Моделирование 2D квантового транспорта

Зависимость $U_{\text{eff}}(\boldsymbol{\rho})$, которая была найдена из решения задачи 3D электростатики, обуславливает транспорт подвижных носителей [6–11]. В низкотемпературном пределе и при тянущем напряжении $V \to 0$ между двумя морями ДГ этот транспорт описывается 2D уравнением Шредингера

$$(\hat{\mathbf{p}}^2/2m)\psi + (E - U_{\text{eff}}(\boldsymbol{\rho}))\psi = 0, \quad (7)$$

где $\hat{\mathbf{p}} = -i\hbar\nabla_{2D}$.

Наиболее важной характеристикой микроконтакта, как субмикронной системы, является кондактанс $G \equiv I/V$, где I — ток между морями ДГ при $V \to 0$. Кондактанс в низкотемпературном пределе дается формулой

Ландауэра:

$$G = (2e^2/h)\sum_n T_n(E), \tag{8}$$

где $2e^2/h$ — квант кондактанса, T_n — коэффициент прохождения для волн, падающих на наноструктуру из n-ой одномерной подзоны подводящего канала, E — полная энергия баллистического электрона. Расчет T_n сводится к решению задачи двумерного квантового рассеяния частицы в бесконечно длинном канале шириной $\Delta y = 2.6$ мкм, имеющем переходную область длиной $\Delta x = 3$ мкм, в которой был задан вычисленный $U_{\rm eff}(x, y)$. Внутри этой области сетка по x, y при решении уравнения (7) была той же самой, как при решении уравнения (2). За пределами данной области зависимости $U_{\rm eff}(y)$ предполагались такими же, как в крайних точках x_{min}, x_{max}. При любых x в точках y_{min}, y_{max} устанавливается $U_{\rm eff} = \infty$. Проверено, что задача двумерного квантового рассеяния в данном случае могла решаться любым из методов, предложенных в [12–15].

Суммарный коэффициент прохождения вычисляется как

$$T_{\rm tot} \equiv \sum_n T_n(E) = \sum_n \sum_m |t_{nm}(E)|^2, \tag{9}$$

где t_{nm} — комплексные амплитуды прохождения с переходом из падающей волны в моде n в прошедшую волну в моде m, и суммирование идет по номерам открытых мод входного и выходного (однородных по x) каналов.

В каждой точке $x_{min} \le x_i \le x_{max}$ решается одномерное уравнение Шредингера для потенциальной ямы $U_{\rm eff}(x = x_i, y)$, т.е. вычисляются соответствующие квантовые уровни $E_n(x_i)$, где n отвечает заданному числу нулей одномерной волновой функции по y. Зависимости $E_n(x_i)$ при фиксированном n =1, 2, 3, и т.д. называются одномерными подзонами. Хотя в общем случае переменные x, y в уравнении (7) не разделяются, графики $E_n(x)$ помогают понять результаты расчета кондактанса. Флуктуации потенциала, показанные на рис. 2, передаются одномерным подзонам. В длинных микроконтактах (L=400–600 нм) на вершине подзоны могут возникать потенциальные ямки (рис. 3a), которые на кривых кондактаса дают резонансы (пики или провалы)(рис. 3b). На показанных графиках такая ямка в $E_3(x)$ и соответствующий широкий пик в $G(E)$ при $E = E_F = 0$ видны для случая $V_{SG} = -0.15$ В. В $E_n(x)$ с $n = 6, 9$ доминирует правый широкий максимум, и соответствующие $G(E)$ не имеют резонансов.

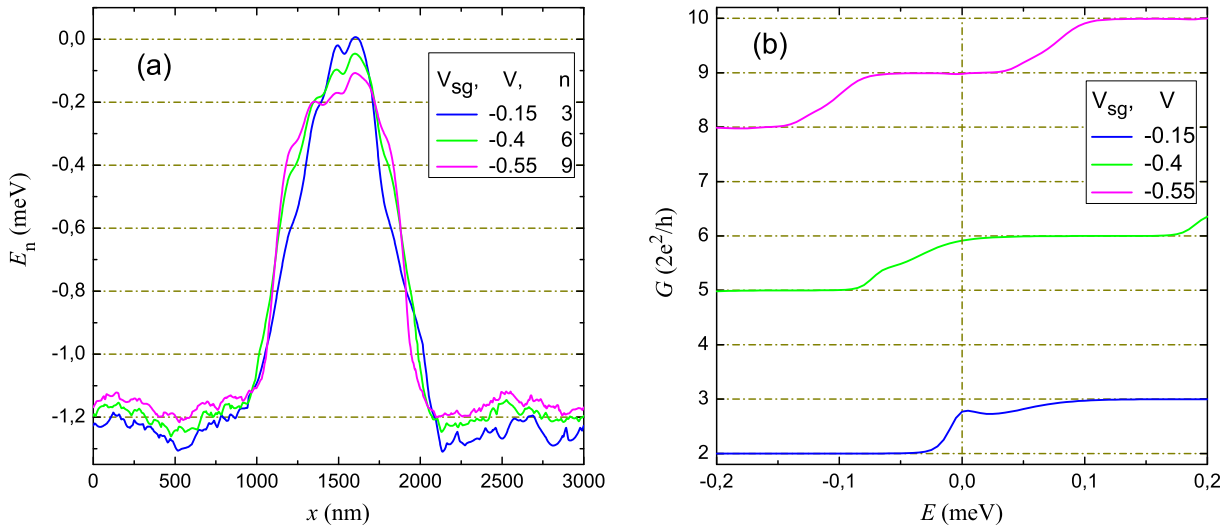


Рис. 3: (a) — Графики 1D подзон E_n для указанных V_{SG} и n; (b) — кондактанс, вычисленный по формулам (8), (9) при указанных V_{SG}. Температура беспорядка поверхностных зарядов 500K, другие параметры указаны в подписи к рис. 2.

При некоторой реализации совершенно случайного распределения поверхностных зарядов лестница кондактанса для высоких n оказывается практически разрушена резонансами (рис. 4). На показанных графиках хорошо видны глубокая и широкая яма на вершине $E_8(x)$ и соответствующие узкие провалы в $G(E)$. В $E_n(x)$ с $n = 2, 4$ доминирует левый широкий максимум, и при $E \approx E_F = 0$ резонансы на соответствующих

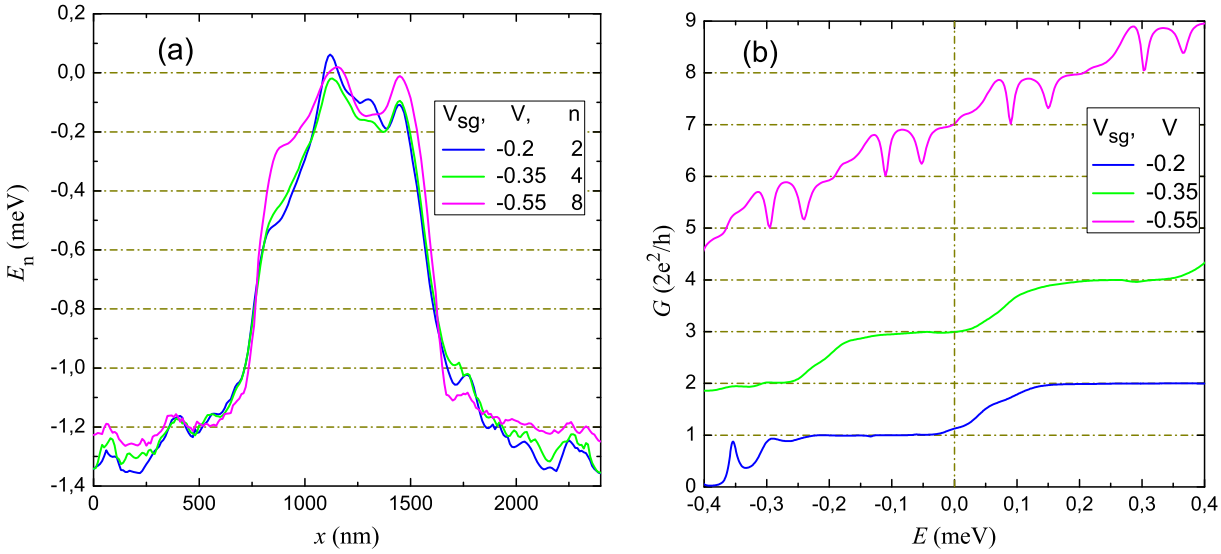


Рис. 4: (a) — Графики 1D подзон E_n для указанных V_{SG} и n в случае совершенно случайного распределения поверхностных зарядов; (b) — кондактанс для этого случая, вычисленный по формулам (8), (9) при указанных V_{SG}. Другие параметры указаны в подписи к рис. 2.

кривых $G(E)$ сглажены. В коротких каналах сохраняется только одно узкое место, и ступени кондактанса получаются гладкими.

Самоорганизация поверхностных зарядов существенно подавляет флуктуации потенциала на глубине двумерного газа, даже при большой температуре беспорядка (500 K). Это хорошо видно из сравнения рис. 3a и 4a. Соответственно, резонансные особенности на высоких ступенях квантования кондактанса становятся менее выраженными или исчезают, что находится в качественном согласии с измерениями кондактанса. Мы считаем, что в реальных образцах имеется примерно такой же беспорядок, как на рис. 2, но температура замораживания беспорядка ниже той, по которой мы находили беспорядок. В нашей упрощенной модели самоорганизации принято во внимание лишь кулоновское взаимодействие зарядов, их положения не ограничены реальной сеткой поверхностных дефектов, и не учитывался разброс энергий связи зарядов с дефектами. Эти неучтенные факторы могут только понижать реальную температуру замораживания по сравнению с температурой, которая найдена из сравнения упрощенной модели самоорганизации с низкотемпературным экспериментом.

Заключение

Применительно к моделированию экспериментов с микроконтактом в затворно-индуцированном двумерном газе дырок численно изучено влияние самоорганизации локализованных поверхностных зарядов на двумерный эффективный потенциал, одномерные подзоны и кондактанс микроконтакта. Показано, что в принятой нами модели беспорядка самоорганизация поверхностных зарядов существенна даже при температурах T_{dis} гораздо выше комнатной.

Расчеты выполнялись с использованием вычислительных ресурсов МСЦ РАН. Мы благодарны нашим коллегам O. Klochan, A. Srinivasan и A.R. Hamilton за предоставленные экспериментальные данные и их обсуждение.

Список литературы

[1] Chen J. C. H., Klochan O., Micolich A. P., Das Gupta K., Sfigakis F., Ritchie D. A., Trunov K., Reuter D., Wieck, A. D., Hamilton A. R. Fabrication and characterisation of gallium arsenide ambipolar quantum point contacts// Appl. Phys. Lett. 2015. V.106. P. 183504.

[2] Miserev D.S., Srinivasan A., Tkachenko O.A., Tkachenko V.A., Farrer I., Ritchie D.A., Hamilton A.R., and Sushkov O.P. Mechanisms for Strong Anisotropy of In-Plane g-Factors in Hole Based Quantum Point Contacts// Phys. Rev. Lett. 2017. V.119. P. 116803.

[3] Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E. Equation of State Calculations by Fast Computing Machines// J. Chem. Phys. 1953. V.21. P. 1087- 1091.

[4] Gann R. C., Chakravarty S., Chester G. V. Monte Carlo simulation of the classical two-dimensional one-component plasma//Phys. Rev. B. 1979. V.20. P. 326-344.

[5] Davies J. H., Larkin I. A., and Sukhorukov E. V. Modeling the patterned two-dimensional electron gas: Electrostatics, J. Appl. Phys. 1995. V.77. P. 4504.

[6] Tkachenko O. A., Tkachenko V. A., Baksheyev D. G., Pyshkin K. S., Harrell R. H., Linfield E. H., Ritchie D. A., and Ford C. J. B. Electrostatic potential and quantum transport in a one-dimensional channel of an induced two-dimensional electron gas// J. Appl. Phys. 2001. V.89. P. 4993.

[7] Ткаченко В.А., Быков А.А., Бакшеев Д.Г., Ткаченко О.А., Литвин Л.В., Латышев А.В., Гаврилова Т.А., Асеев А.Л., Портал Ж.К. Одноэлектронная зарядка треугольных квантовых точек кольцевого интерферометра// ЖЭТФ. 2003. Т.124. С.351.

[8] Tkachenko O. A., Tkachenko V. A., Terekhov I. S., and Sushkov O. P. Effects of Coulomb screening and disorder on an artificial graphene based on nanopatterned semiconductor// 2D Mater. 2015. V.2. P. 014010.

[9] Ткаченко В.А., Ткаченко О.А, Миньков Г.М., Шерстобитов А.А. Перколяция и электрон-электронное взаимодействие в решетке антиточек// Письма в ЖЭТФ. 2016. Т.104. С. 501–506.

[10] Ткаченко О.А, Ткаченко В.А. Суперкомпьютерное моделирование полупроводниковых квантовых наносистем// Вычислительные методы и программирование. 2012. Т. 13. С. 253–262.

[11] Tkachenko O., Tkachenko V., Kvon Z., Sheglov D., and Aseev A. Modeling of Quantum Transport and Single-Electron Charging in GaAs/AlGaAs-Nanostructures, Chapter 6 in "Advances in Semiconductor Nanostructures, Growth, Characterization, Properties and Applications" Eds: A. Latyshev, A. Dvurechenskii, and A. Aseev, Elsevier. 2017. P. 131–155.

[12] Takagaki Y., Ferry D. K. Conductance of quantum waveguides with a rough boundary// J. Phys.: Condens. Matter. 1992. V.4. P. 10421–10432.

[13] Ando T. Quantum point contacts in magnetic fields// Phys. Rev. B 1991. V.44. P. 8017–8027.

[14] Usuki T., Saito M., Takatsu M., Kiehl R. A., Yokoyama N. Numerical analysis of ballistic-electron transport in magnetic fields by using a quantum point contact and a quantum wire// Phys. Rev. B. 1995. V.52. P. 8244–8255.

[15] Cresti A., Farchioni R., Grosso G., and Parravicini G. P. Keldysh-Green function formalism for current profiles in mesoscopic systems// Phys. Rev. B. 2003. V.68. P. 075306.

Ткаченко Ольга Александровна — к.ф.-м.н., ст. науч.сотр. Института физики полупроводников СО РАН;
e-mail: otkach@isp.nsc.ru;
Бакшеев Дмитрий Георгиевич — к.ф.-м.н., ст. разработчик программного обеспечения компании Яндекс, ассистент кафедры общей информатики факультета информационных технологий Новосибирского государственного университета
e-mail: d.baksheev@g.nsu.ru;
Сушков Олег Петрович — д.ф.-м.н., professor (full) University of New South Wales, Sydney ;
e-mail: suskov@unsw.edu.au
Ткаченко Виталий Анатольевич — к.ф.-м.н., ст. науч.сотр. Института физики полупроводников СО РАН;
доцент Новосибирского государственного университета, доцент Новосибирского государственного технического университета;
e-mail: vtkach@isp.nsc.ru.
Дата поступления — 24 мая 2019 г.

МОДЕЛИРОВАНИЕ САМООРГАНИЗАЦИИ ЛОКАЛИЗОВАННЫХ ЗАРЯДОВ НА ГРАНИЦЕ ПОЛУПРОВОДНИКА С ПОДЗАТВОРНЫМ ДИЭЛЕКТРИКОМ

О. А. Ткаченко[1], Д. Г. Бакшеев[2,3] В. А. Ткаченко[1,3,4] О. П. Сушков[5]

[1] *Институт физики полупроводников им. А.В.Ржанова СО РАН, 630090, Новосибирск*

[2] *Яндекс, 630090, Новосибирск*

[3] *Новосибирский государственный университет, 630090, Новосибирск*

[4] *Новосибирский государственный технический университет, 630073, Новосибирск*

[5] *School of Physics, University of New South Wales, Sydney, 2052, Australia*

УДК 004.942:537.311.322-0.22.532:537.29+538.911:530.145
DOI: 10.24411/9999-016A-2019-10082

Рассмотрены корреляции в распределении локализованных зарядов на плоской поверхности полупроводниковой структуры с двумерным электронным или дырочным газом и влияние этих корреляций на транспортное и квантовое время рассеяния носителей заряда. Для моделирования выбраны структуры, где это влияние оказывается максимальным: поверхность полупроводника отделена от металлического затвора тонким слоем диэлектрика (20-30 нм), двумерный газ приближен к поверхности (50 нм) и специальное легирование отсутствует. Предполагалось, что в равновесии лишь температура вносит беспорядок в систему точечных зарядов на идеальной плоскости, имитирующей границу между полупроводником и подзатворным диэлектриком. При расчете внутренней энергии системы учтены заряды изображения в металле. Методом Монте-Карло найдены равновесные распределения точечных зарядов при разных температурах. Вычислена зависимость структурного фактора от волнового числа q. В пределе малых q, а также сравнением с теорией однокомпонентной плазмы, обнаружено понижение структурного фактора в два–три раза при температурах $T = 500 - 1000$ K по сравнению со случаем совершенно случайного беспорядка. Предположение о «замораживании» при $T \sim 1000$ K распределений локализованных зарядов позволило объяснить, почему экспериментальные времена рассеяния двумерных носителей в соответствующих структурах гораздо больше теоретических значений, полученных без учета корреляций.

Ключевые слова: метод Монте-Карло, алгоритм Метрополиса, поверхностный заряд, однокомпонентная плазма, термодинамическое равновесие, внутренняя энергия, самоорганизация, температура беспорядка, квантовое и транспортное времена рассеяния.

Введение

Для экспериментального изучения квантового транспорта двумерных носителей заряда особый интерес представляют *нелегированные* полупроводниковые гетероструктуры, в которых двумерный газ расположен на небольшой глубине (30–60 нм) под поверхностью [1, 2]. В отличие от обычных структур, рабочий слой заполняется двумерным газом (ДГ) электронов, либо дырок при низких температурах (~ 1 K) под действием электрического напряжения соответствующего знака между сплошным металлическим затвором и металлическими контактами к рабочему слою. Подвижность носителей ДГ ограничена рассеянием на зарядах, которые локализованы на поверхности полупроводника. Система поверхностных зарядов возникает в термодинамически равновесном состоянии структуры при температуре выше комнатной, благодаря обмену

Работа выполнена при поддержке Российского научного фонда (проект 19-72-30023).

ISBN 978-5-901548-42-4

термически возбужденными электронами между полупроводником и металлом (Ti), нанесенным на тонкий (20–30 нм) подзатворный диэлектрик (Al_2O_3). Часть этих электронов захватывается точечными дефектами (ловушками) на поверхности полупроводника (GaAs) из-за разницы работ выхода электрона из GaAs и Ti. При температурах ниже некоторой «температуры замерзания» распределение поверхностных зарядов уже не меняется. Значение температуры замерзания обусловлено энергией выброса электрона на уровень протекания из глубоких ловушек, в которых оказались электроны, однако в структурах данного типа температура замораживания неизвестна.

В этой работе численно и теоретически изучается самоорганизация поверхностных зарядов в нелегированных структурах с затворно-индуцированным двумерным газом. Оценка концентрации поверхностных зарядов в структуре металл-диэлектрик-GaAs/AlGaAs/GaAs показывает, что она может быть значительной ($5 \cdot 10^{11}$ см$^{-2}$). При таких концентрациях в результате электростатического расталкивания зарядов их расположение оказывается не чисто случайным, а скоррелированным. При температуре ниже температуры замерзания распределение поверхностных зарядов фиксируется. Беспорядок в распределении порождает флуктуации электростатического потенциала в слое с ДГ, на которых рассеиваются электроны или дырки ДГ. При моделировании беспорядка в расположении поверхностных зарядов мы игнорировали детали химического строения поверхности с большим числом точечных дефектов, а также кинетики зарядки дефектов, которыми можно в первом приближении пренебречь, и нашли температуру замораживания, которая определяет уровень беспорядка в структуре.

Известно, что концентрация ловушек на поверхности GaAs (и почти всех полупроводников) огромна в сравнении с концентрацией поверхностных зарядов. Энергия ионизации заполненных ловушек в GaAs близка к полуширине запрещенной зоны данного материала, т.е. гораздо больше тепловой энергии ($E_g/2 \gg k_BT$) даже при комнатной температуре T, а разброс глубин ловушек не столь велик [3]. Следовательно, процесс обмена электронами между соседними ловушками на поверхности гораздо интенсивнее, чем между поверхностью и объемом полупроводника. Поэтому для простоты моделирования данной ситуации мы пренебрегли обменами с объемом и считали, что поверхностные заряды способны перемещаться в любое положение на идеальной плоскости. В такой модели все факторы беспорядка, кроме температуры в термодинамическом равновесии, выключены. Заметим, что для изолированных на плоскости, либо в объеме систем частиц с любым потенциалом межчастичного взаимодействия давно предложен эффективный вариант метода Монте-Карло, позволяющий находить наиболее вероятное значение внутренней энергии и равновесные расположения частиц при заданной температуре [4]. В пределе бесконечно высокой температуры заряды на плоскости распределены случайным образом, но с понижением температуры расталкивание одноименных зарядов ведет к корреляциям. В данной модели, даже при очень низких температурах (0.1 K) система переходит в равновесное состояние, отвечающее минимуму внутренней энергии [5]. Это состояние в виде треугольной решетки точечных зарядов называется двумерным вигнеровским кристаллом. Именно так выстраиваются электроны на поверхности жидкого гелия, и подобное выстраивание предсказывалось даже для ДГ подвижных носителей возле границы Si/SiO_2, которая практически свободна от точечных зарядов [6]. Однако получить двумерный вигнеровский кристалл на поверхности твердых тел не удается, поскольку с понижением температуры решающим становится другой фактор беспорядка в виде нерегулярной сетки ловушек неодинаковой глубины. В нашей простой модели уровень беспорядка в системе поверхностных зарядов вычисляется численно алгоритмом Метрополиса [4], либо аналитически (в рамках теории двумерной однокомпонентной плазмы [5]) при любой заданной температуре T. Оба подхода адаптированы к изучаемой ситуации.

Центральным при рассмотрении любой однокомпонентной плазмы является способ, которым обеспечивается суммарный нулевой заряд, т.е. электрическая нейтральность системы. Как правило, это делается формальным введением однородного нейтрализующего фонового заряда в тот же объем, либо ультра-тонкий слой, который занимает однокомпонентная плазма [5, 7–10]. Напротив, мы учли, что электронейтральность в изучаемой нами ситуации при нулевом затворном напряжении обеспечивается существованием заряда изображения в металле (Ti/Au) для каждого точечного заряда на поверхности полупроводника. От описания расположения точечных зарядов радиус-векторами мы перешли преобразованием Фурье к структурному фактору F_q. Для чисто случайного расположения ($T = \infty$) при любых q выполняется $F_q = 1$. Для достаточно высоких температур $T > 5$ K величина F_q зависит только от модуля q волнового вектора $\mathbf{q}$. Заметим, что для обычных структур, создаваемых методом удаленного легирования, разработана теория влияния ультратонких слоев заряженных примесей на низкотемпературные характеристики высокоподвижных двумерных носителей заряда [10–12]. В ее основе лежит знание F_q для распределения заряженных примесей. Для изучаемой нами системы была построена аналогичная теория, в которой учтено присутствие металлического

затвора. С ее помощью найдено, что для предсказания влияния беспорядка в расположении поверхностных зарядов на затворно-индуцированный двумерный газ важны лишь значения F_q в пределе малых волновых чисел q. Зависимости $F_q(T)|_{q\to 0}$ при $\infty > T > 100$ K получились близкими при численном расчете с помощью алгоритма Метрополиса и при расчете по формулам теории однокомпонентной плазмы. Значения $F_q|_{q\to 0}$ оказались в два-три раза меньше единицы, как это требуется для соответствия эксперименту, при T, близкой к 1000 K. Таким образом, самоорганизация поверхностных зарядов в предложенной термодинамической модели становится существенной, даже при столь высоких T.

1 Определения и конечные формулы модели

Запишем Фурье преобразование плотности локализованных поверхностных зарядов, определяемой радиус-векторами $\boldsymbol{r}_i$ (i-номер заряда):

$$\rho_{\boldsymbol{q}} = \sum_i e^{i\boldsymbol{q}\cdot\boldsymbol{r}_i} \ . \tag{1}$$

Если заряды распределены совершенно случайно, то

$$|\rho_{\boldsymbol{q}}|^2 = N n_\sigma (2\pi)^2 \delta(\boldsymbol{q}) + |\tilde{\rho}_q|^2 \ , \tag{2}$$

где n_σ — поверхностная концентрация локализованных зарядов и полное число зарядов $N \to \infty$, q — волновое число. Для случайного распределения справедливо

$$|\tilde{\rho}_q|^2 = n_\sigma A \ , \tag{3}$$

где полная площадь $A \to \infty$. Формула (3) представляет собой белый шум. Член с дельта-функцией в (2) дает постоянный потенциал, который мы отбрасываем. Нас интересуют флуктуации потенциала в двумерном газе, который находится на глубине z от плоскости с локализованными зарядами. Для этого мы решаем уравнение Пуассона и находим Фурье образ потенциала в плоскости ДГ с учетом самоэкранировки

$$U_{\boldsymbol{q}} = \rho_{\boldsymbol{q}} \frac{(2\pi e^2/\epsilon)}{q + \kappa D_z} e^{-qz} C_d \ . \tag{4}$$

Здесь $\epsilon = (\epsilon_1 + \epsilon_2)/2$, ϵ_1, ϵ_2 — диэлектрические проницаемости Al_2O_3 и GaAs, d — толщина слоя диэлектрика и e — заряд электрона

$$\begin{aligned} C_d &= \frac{1 - e^{-2qd}}{1 - \lambda e^{-2qd}} \\ D_z &= 1 + e^{-2qz} \frac{\lambda - e^{-2qd}}{1 - \lambda e^{-2qd}} \\ \lambda &= \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} \\ \kappa &= \frac{2m^* e^2}{\epsilon_1} \ . \end{aligned}$$

Значения κ зависят от эффективной массы m^*. Для электронов $m^* = 0.07 m_e$ получаем $\kappa = 0.25\text{нм}^{-1}$, для дырок $m^* = 0.3 m_e$ и $\kappa = 1\text{нм}^{-1}$. Одночастичная волновая функция носителя в ДГ есть

$$\psi = \frac{1}{\sqrt{A}} e^{i\mathbf{p}\cdot\mathbf{r}} \ . \tag{5}$$

Тогда квадрат модуля матричного элемента взаимодействия имеет вид:

$$|V_{\boldsymbol{q}}|^2 = \frac{|\rho_{\boldsymbol{q}}|^2}{A^2} \frac{(2\pi e^2/\epsilon)^2}{(q + \kappa D_z)^2} e^{-2qz} C_d^2 \ . \tag{6}$$

Золотое правило Ферми позволяет вычислить время жизни одночастичных состояний (квантовое время рассеяния τ_q) и транспортное время рассеяния τ_t:

$$\begin{aligned} &\dot{W}_q = \frac{1}{\tau_q} = 2\pi \int |V_{\boldsymbol{q}}|^2 \delta(\epsilon_i - \epsilon_p) \frac{A d^2 p}{(2\pi)^2} = 2\pi m^* \frac{e^4}{\epsilon^2} \int \frac{(|\rho_{\boldsymbol{q}}|^2/A)}{(q + \kappa D_z)^2} e^{-2qz} C_d^2 d\theta \\ &q = 2p_F \sin(\theta/2) \end{aligned} \tag{7}$$

$$\dot{W}_t = \frac{1}{\tau_t} = 2\pi m^* \frac{e^4}{\epsilon^2} \int \frac{(|\rho_{\boldsymbol{q}}|^2/A)}{(q+\kappa D_z)^2} e^{-2qz} C_d^2 (1-\cos\theta) d\theta$$

Определим зарядовые флуктуации на поверхности полупроводника через структурный фактор F_q, предполагая, что распределение $|\tilde{\rho}_q|^2$, подобно (2), зависит только от волнового числа, но не от направления волнового вектора

$$|\tilde{\rho}_q|^2 = n_\sigma A \times F_q \ . \tag{8}$$

Ниже будет установлена область применимости этого предположения. Для совершенно случайного распределения $F_q = 1$. Используя (8), перепишем (7) и (8) как

$$\begin{aligned} \tau_q^{-1} &= \frac{\pi n_\sigma}{2m^*} I_q \\ I_q &= \frac{2}{p_F} \int_0^\infty \frac{F_q}{(q/\kappa + D_z)^2} e^{-2qz} C_d^2 dq \ . \end{aligned} \tag{9}$$

$$\begin{aligned} \tau_t^{-1} &= \frac{\pi n_\sigma}{2m^*} I_t \\ I_t &= \frac{1}{p_F^3} \int_0^\infty \frac{F_q}{(q/\kappa + D_z)^2} e^{-2qz} C_d^2 q^2 dq \end{aligned} \tag{10}$$

Кроме F_q в этих формулах все известно. В рамках теории однокомпонентной плазмы с учетом самоэкранирования в системе поверхностных зарядов и зарядов изображения в металле структурный фактор получился следующим

$$\begin{aligned} F_q &= \left[1 - \frac{\tilde{\kappa}_T}{q + \tilde{\kappa}_T}\right] \\ \tilde{\kappa}_T &= \kappa_T C_d \\ \kappa_T &= \frac{2\pi e^2 n_\sigma}{\epsilon T}. \end{aligned} \tag{11}$$

2 Моделирование самоорганизации методом Монте-Карло

Целью моделирования было нахождение равновесного расположения поверхностных локализованных зарядов с концентрацией n_σ при заданной температуре T и вычисление зависимости структурного фактора F от волнового числа q, которая необходима для расчета квантового и транспортного времен рассеяния подвижных носителей в двумерном газе по формулам (9) и (10). Поскольку полная плотность поверхностных состояний в запрещенной зоне много больше, чем n_σ, можно считать, что заряды могут перемещаться в любое положение. Таким образом, мы используем континуальную, а не решеточную модель самоорганизации локализованных зарядов.

Рассмотрим конечную систему точечных зарядов на прямоугольнике площадью A. Полное число зарядов определяется $N = n_\sigma A$. Для моделирования бесконечной системы границы прямоугольника замыкались в тор. Предполагалось, что плоскость с зарядами находится в однородной среде с диэлектрической проницаемостью ϵ, а над плоскостью на расстоянии d находится металлический затвор. Мы пренебрегали разрывами диэлектрической проницаемости между слоями диэлектрика Al_2O_3 и полупроводника GaAs/AlGaAs и заменили ее средним значением ϵ. С учетом зарядов изображения полная энергия взаимодействия сводилась к сумме попарных взаимодействий U_{ij} вертикальных диполей длины $2d$:

$$U_{ij} = \frac{1}{2} \frac{q_0^2}{2\pi\epsilon\epsilon_0} \left(\frac{1}{r_{ij}} - \frac{1}{\sqrt{r_{ij}^2 + 4d^2}} \right) \tag{12}$$

$$E = \frac{1}{2} \sum_{i \neq j} U_{ij}. \tag{13}$$

Отметим, что множитель 0.5 при определении U_{ij} возникает из-за того, что электростатическая энергия взаимодействия диполей из зарядов q_0 и их изображений в два раза меньше энергии взаимодействия диполей, составленных из реальных зарядов $+q_0$, $-q_0$. При суммировании на торе брались наименьшие расстояния r_{ij} в любой паре взаимодействующих частиц. Кинетической энергией зарядов пренебрегалось по сравнению с потенциальной. Наиболее вероятное значение электростатической энергии взаимодействия при заданной температуре находилось алгоритмом Метрополиса [4]. Для каждой итерации любая частица смещалась в произвольном направлении на случайное расстояние не больше среднего между частицами. Полная энергия пересчитывалась и при выигрыше энергии $\Delta E < 0$ новое положение частицы принималось, а при проигрыше новое положение принималось, если $\exp(-\Delta E/k_BT) \geq r$, где r — случайное число от 0 до 1. Для найденных равновесных распределений вычислялась по формуле (1) зависимость $|\rho_{\mathbf{q}}|^2$. Расчеты выполнены при следующих параметрах: $d = 25$ нм, $\epsilon = 10$. Среднее расстояние между частицами в нанометрах можно найти по формуле $r_s = 17.84/\sqrt{n_\sigma}$, где n_σ подставляется в единицах 10^{11}см$^{-2}$. Распределение поверхностных зарядов порождает в плоскости двумерного газа флуктуации потенциала характерного размера 150-200 нм. Для адекватного расчета корреляторов размеры области счета должны быть больше 2 мкм на 2 мкм, при этом число зарядов достигает 60000. Кроме того, чем больше область счета, тем точнее расчеты при малых волновых числах q.

Модель проверялась в двух предельных случаях. Для совершенно случайного распределения частиц (бесконечно высокая температура) структурный фактор F_q вел себя в согласии с формулами (1–3, 8). При температуре ниже 1K наступала вигнеровская кристаллизация, т.е. заряды выстраивались в треугольную решетку. Расчет $|\rho_{\mathbf{q}}|^2$ давал яркие пики в вершинах шестиугольника — первую зону Бриллюэна и множество рефлексов, отвечающих расширенной зоне Бриллюэна. При температурах выше 4K зависимость $|\rho_{\mathbf{q}}|^2$ становится аксиально-симметричной (рис. 1a). Вычислялась усредненная по колечку радиусом q и шириной

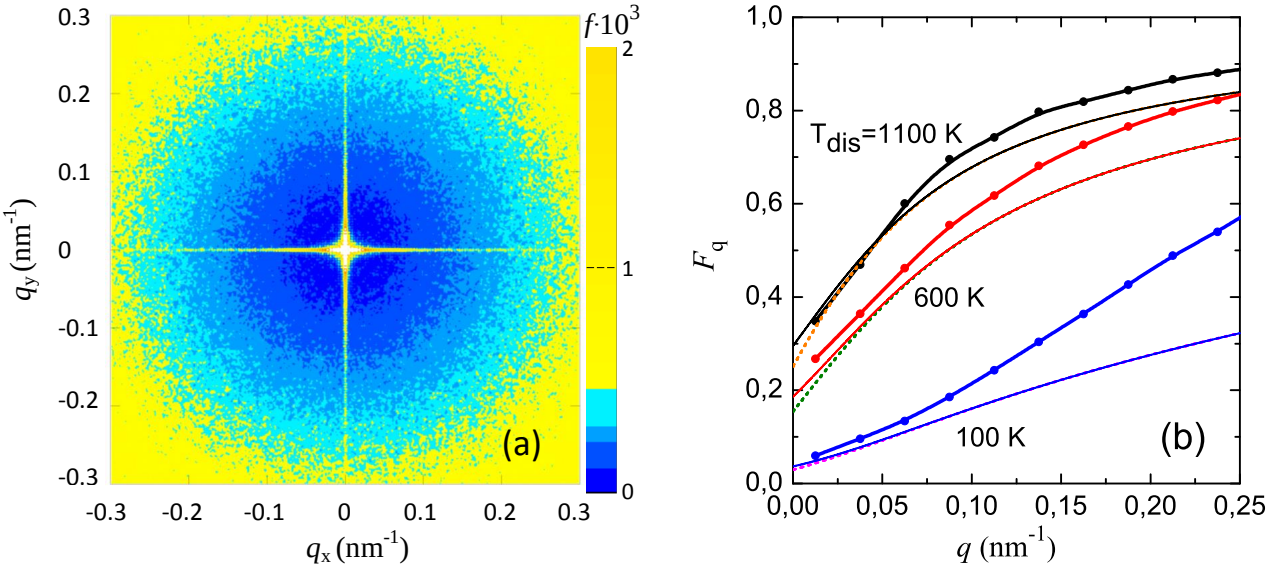


Рис. 1: (a) Пример преобразования Фурье $f \equiv |\rho(q_x, q_y)|^2$ от распределения поверхностных зарядов (1) на плоскости (x, y) при электростатической самоорганизации в случае $n_\sigma = 5 \cdot 10^{11}$ см$^{-2}$ в термодинамически равновесном состоянии с температурой $T = 100$ K: расчет методом Монте-Карло, белый крест с уширением в центре связан с δ-функцией в формуле (2) и конечным размером области счета на плоскости (x, y). (b) Зависимость структурного фактора $F(q)$ для той же n_σ и трех температур беспорядка T_{dis}. Толстые сплошные линии получены методом Монте-Карло для $\epsilon_{1,2} = 10$. Тонкие сплошные линии для $\epsilon_{1,2} = 10$ и тонкие точечные линии для $\epsilon_1 = 12$, $\epsilon_2 = 8$ построены по формулам (11).

Δq радиальная зависимость структурного фактора от q для разных температур. При этом центральный пик, связанный с δ-функцией в формуле (2) и уширенный за счет конечного размера области счета, вычитался и не давал вклада в расчет F_q. Однако сделать это аккуратно в области очень малых $q < 0.1$ нм$^{-1}$, которые дают основной вклад в интегралы (9), (10), удается только с увеличением области счета. Сравнение структурных факторов, полученных в численном моделировании и в приближении слабо неидеальной плазмы (11), показано на рис. 1b. Видно, что при низких $q < 0.1$ нм$^{-1}$ есть неплохое соответствие для всех

температур, и формула (11) может быть использована для учета корреляции при оценке времен рассеяния.

На рис.2a, b для тех же температур показаны зависимости от q подинтегральных выражений в I_q, I_t, предназначенных для расчета времен рассеяния τ_q, τ_t по формулам (9) и (10) на характерном примере структуры с ДГ. На этом примере видно, что для вычисления интегралов достаточен интервал $0 < q < 0.1$ нм$^{-1}$. Следовательно, Монте-Карло моделированием обоснована возможность использования формул (9) и (10) вместе с (11) для расчета времен рассеяния с учетом самоорганизации поверхностных зарядов. Из рис.1, 2 видно еще, что в этом расчете можно с хорошей точностью вместо диэлектрических проницаемостей Al_2O_3 и GaAs использовать среднее значение $\epsilon = (\epsilon_1 + \epsilon_2)/2$.

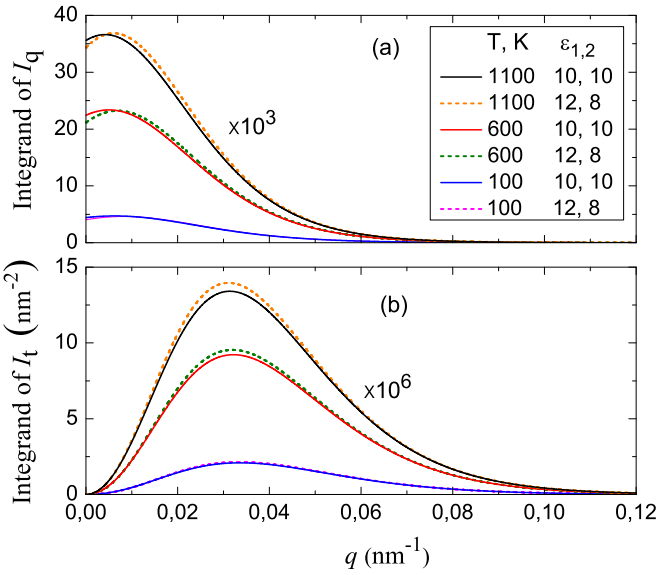


Рис. 2: Для трех температур двумя типами линий показано поведение подинтегральных выражений в формулах (9) и (10) в случае образца с электронным ДГ, в котором $d = 25$ нм, $z = 45$ нм, $n_\sigma = 5 \cdot 10^{11}$ см$^{-2}$ при вычислении F_q по формулам (11).

Наши коллеги в University of New South Wales (Australia) экспериментально изучили образцы с затворно-индуцированным ДГ электронов ($d = 25$ нм, $z = 45$ нм) и дырок ($d = 20$ нм, $z = 68$ нм). Мы сравнили теоретические и экспериментальные времена рассеяния τ_t, τ_q для этих образцов. Из сравнения найдена температура замерзания беспорядка $T_{dis} \approx 1000$ K в случае образца с двумерным электронным газом и $T_{dis} \approx 700$ K для образца с двумерным дырочным газом. Гипотеза замораживания вычисленного для T_{dis} распределения поверхностных зарядов позволила объяснить, почему измеренные времена рассеяния двумерных носителей гораздо больше теоретических значений, полученных без учета корреляций. Поскольку в данной модели пренебрегалось всеми факторами беспорядка, кроме температуры, найденные значения T_{dis} являются оценкой сверху для реальной температуры замораживания беспорядка, которая для образцов данного типа неизвестна.

Заключение

Численно и теоретически изучена самоорганизация локализованных зарядов на границе полупроводник-подзатворный диэлектрик в структурах металл-изолятор-нелегированный полупроводник GaAs/AlGaAs. С помощью алгоритма Метрополиса для разных температур численно найдены равновесные распределения локализованных поверхностных зарядов и соответствующие структурные факторы, характеризующие корреляции. Предложены аналитические формулы для транспортного и квантового времен рассеяния с учетом корреляций зарядов, отвечающей заданной температуре замерзания беспорядка T_{dis}. Показано, что в области малых волновых чисел q приближение слабонеидеальной однокомпонентной плазмы для структурного фактора хорошо описывает численные расчеты. В формулах учтена специфика структур с индуцированным двумерным газом: экранировка локализованных поверхностных зарядов верхним затвором, самоэкранировка носителей двумерного газа, а также различие диэлектрических проницаемостей диэлектрика и полупро-

водника. Из сравнения теоретического и экспериментального транспортного времени рассеяния показано, что самоорганизация поверхностных зарядов существенна при температурах T_{dis} гораздо выше комнатной.

Расчеты выполнялись с использованием вычислительных ресурсов МСЦ РАН. Мы благодарны сотрудникам университета Нового Южного Уэльса (Австралия) A.R. Hamilton, O. Klochan и D. Q.Wang за предоставленные экспериментальные данные и их обсуждение.

Список литературы

[1] Chen J. C. H., Wang D. Q., Klochan O., Micolich A. P., Das Gupta K., Sfigakis F., Ritchie D. A., Reuter D., Wieck A. D., and Hamilton A. R. Fabrication and characterization of ambipolar devices on an undoped AlGaAs/GaAs heterostructure // Appl. Phys. Lett. 2012. V.100. P. 052101.

[2] Mak W. Y., Das Gupta K., Beere H. E., Farrer I., Sfigakis F. , and Ritchie D. A. Distinguishing impurity concentrations in GaAs and AlGaAs using very shallow undoped heterostructures //Appl. Phys. Lett. 2010. V.97. P. 242107.

[3] Darling R. B. Defect-state occupation, Fermi-level pinning, and illumination effects on free semiconductor surfaces// Phys. Rev. B. V.43. P. 4071.

[4] Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E. Equation of State Calculations by Fast Computing Machines// J. Chem. Phys. 1953. V.21. P. 1087–1091.

[5] Gann R. C., Chakravarty S., Chester G. V. Monte Carlo simulation of the classical two-dimensional one-component plasma//Phys. Rev. B. 1979. V.20. P. 326–344.

[6] Chaplik A. V. Possible Crystallization of Charge Carriers in Low-Density Inversion Layers// Zh. Eksp. Teor. Fiz. 1971. V.62. P.746.

[7] Pikus F. G. and Efros A. L. Large-scale potential fluctuations in plane layers with impurities// Sov. Phys. JETP. 1989 V.69. P. 553.

[8] Efros A. L., Pikus F. G. and Samsonidze G. G. Maximum low-temperature mobility of two-dimensional electrons in heterojunctions with a thick spacer layer// Phys. Rev. B. 1990. V.41. P.8295 (1990).

[9] Das Sarma S. and Kodiyyalam S. Intrinsic dopant correlations and transport properties of mesoscopic modulation-doped heterostructures, Semicond. Sci. Technol., **13**, A59 (1998).

[10] Das Sarma S. , Hwang E. H., Kodiyalam S., Pfeiffer L. N., and West K. W. Transport in two-dimensional modulation-doped semiconductor structures, Phys. Rev. B, **91**, 205304 (2015).

[11] Das Sarma S., and Stern F. Single-particle relaxation time versus scattering time in an impure electron gas, Phys. Rev. B., **32**, 8442 (1985).

[12] Gold A. Transport time and single-particle relaxation time in two-dimensional semiconductors, Proc. SPIE Vol. 1362 Physical Concepts of Materials for Novel Optoelectronic Device Applications II: Device Physics and Applications (1991) P. 309–312.

Ткаченко Ольга Александровна — к.ф.-м.н., ст. науч.сотр. Института физики полупроводников СО РАН;
e-mail: otkach@isp.nsc.ru;
Бакшеев Дмитрий Георгиевич — к.ф.-м.н., ст. разработчик программного обеспечения компании Яндекс, ассистент кафедры общей информатики факультета информационных технологий Новосибирского государственного университета
e-mail: d.baksheev@g.nsu.ru;
Ткаченко Виталий Анатольевич — к.ф.-м.н., ст. науч.сотр. Института физики полупроводников СО РАН;
доцент Новосибирского государственного университета;
доцент Новосибирского государственного технического университета;
e-mail: vtkach@isp.nsc.ru;
Сушков Олег Петрович — д.ф.-м.н., professor (full) University of New South Wales, Sydney;
e-mail: suskov@unsw.edu.au

Дата поступления — 24 мая 2019 г.

ЛУЧ СВЕТА В ТУРБУЛЕНТНОЙ СРЕДЕ

В. В. Учайкин[1]

[1] *Ульяновский государственный университет, 432017, Ульяновск*

УДК 52-17
DOI: 10.24411/9999-016A-2019-10083

Рассматриваются несколько моделей распространения излучения (радиоволн, света, рентгеновского и гамма-излучения) в турбулентной среде с применением к задачам астрономии и астрофизики. В отличие от стандартных моделей случайно-неоднородных сред, построенных на основе теории возмущений и по этой причине применимых к описанию рассеяния на мелкомасштабных флуктуациях среды гауссова типа, турбулентная среда характеризуется пульсациями больших амплитуд, крайне нерегулярно распределённых в пространстве. Учёт этих особенностей приводит к необходимости использовать уравнения с нелокальными операторами дробного типа. В докладе даётся краткий обзор свойств их решений.

Ключевые слова: рассеяние, турбулентность, диффузия, устойчивые законы.

Введение

Излучение космических источников (радиоволны, свет, рентгеновы лучи, гамма-излучение) на пути до его регистрации испытывает рассеяние, обусловленное случайными искривлениями лучей, проходящих через неоднородности межзвездной плазмы. Последствия этого рассеяния проявляются в увеличении измеряемых угловых размеров локальных источников, флуктуациях интенсивности (мерцаниях) при измерениях в ограниченных частотных полосах, увеличении длительности и изменении формы сигналов от импульсных источников. Эффект запаздывания является основным источником оценки расстояний до пульсаров [1]. В свою очередь, искажение формы изображений, получаемых на современных земных и внеземных телескопах, является нежелательным эффектом, для компенсации разрабатываются специальные коды [2], основанные на теории переноса излучения в межзвездной среде, основным свойством которой является турбулентная неоднородность [3]. Она характеризуется сильными пульсациями, далекими корреляциями, степенными законами, в силу чего клаассическая гауссова модель случайной среды оказывается неприменима, равно как и теория возмущений для описания диффузии в такой среде, предполагающая неоднородности мелкомасштабными, а их амплитуды малыми. В настоящей работе даётся краткий обзор моделей, свободных от этих ограничений и примениых к расчётам сигналов от далёких источников космического излучения — пульсаров, квазаров, космических всплесков.

1 Волновая модель

В основе волновой модели переноса изображения в мелкодисперсной (газ, плазма) однородной среде лежит классическое волнововое уравнение

$$\triangle u - \frac{1}{c_0^2}\frac{\partial^2 u}{\partial t^2} = 0.$$

Представляя его решение $u(\mathbf{r},t)$ в виде произведения быстро меняющегося гармонического множителя $\exp(i\omega t)$ на медленно меняющуюся амплитуду $\phi(\mathbf{r},t)$, так, что

$$\left|\frac{\partial^2\phi}{\partial t^2}\right| \ll 2\omega\left|\frac{\partial\phi}{\partial t}\right|,$$

и

$$\frac{\partial^2 u}{\partial t^2} \cong \exp(i\omega t)\left(2i\omega\frac{\partial\phi}{\partial t} - \omega^2\phi\right),$$

ISBN 978-5-901548-42-4

приходим к параболическому уравнению

$$(\triangle + k^2)\phi = \frac{2ik}{c_0}\frac{\partial\phi}{\partial t}$$

с волновым числом $k = \omega/c_0$, где c_0 — скорость света. Комбинируя его с комплексно-сопряжённой его версией

$$(\triangle + k^2)\phi^* = -\frac{2ik}{c_0}\frac{\partial\phi^*}{\partial t},$$

после несложных преобразований получаем для распределения интенсивности $I(x,y;t) = \phi^*\phi \equiv |\phi|^2$ в плоскости xy диффузионное уравнение

$$\frac{\partial I(x,y;t)}{\partial t} = D\triangle I(x,y;t) \tag{1}$$

с оператором

$$\triangle \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2},$$

коэффициентом диффузии $D = c_0/(2k_1)$ и начальным условием $I(x,y;0) = I_0(x,y)$, представляющим выходящее из источника изображение $I_0(x,y)$. Результирующее изображение в момент t дается двумерной свёрткой первичного изображения и двумерного гауссиана

$$G^{(1)}(x,y;t) = \frac{1}{4\pi Dt}\exp\left[-\left(\frac{x^2+y^2}{4Dt}\right)\right], \tag{2}$$

являющегося функцией Грина диффузионного уравнения (1).

Точка на плоскости xy, представляющая собой проекцию на нее положения луча в момент t, совершает на этой плоскости броуновское движение, физические адекватное предельному блужданию частицы в случайном распределении *независимых* и *равномерно распределённых* атомов (типа атомов идеального газа). Вот эти-то два фактора этой модели и мешают принять её в качестве рабочей для описания процесса распространения фотонов в межзвёздной среде. Броуновская модель (в её допредельном варианте) предполагает экспоненциальное распределение интервалов времени между последовательными столкновениями, формирующее пуассоновский (во времени) поток событий, и определённую ограниченность элементарных смещений, обеспечивающую существование среднего квадрата смещения. Отношение среднего квадрата смещения к среднему интервалу времени между событиями и даёт величину коэффициента диффузии.

Главным препятствием на пути применения такой модели является расходимость указанных характеристик, порождаемых турбулентным характером межзвёздной среды, а именно, степенными законами. В [4], в частности, указывается, что основной причиной турбулентности является развитие гравитационной МГД-неустойчивости, приводящей к формированию иерархической последовательности — суперкластеры, кластеры, молекулярные облака, протозвёздные облака, которые и идентифицируются по различию в показателях степенных законов, связывающих скорости с магнитными полями, плотностями и размерами наблюдаемых облаков.

2 Дробно-дифференциальная супердиффузия

В теории многократного рассеяния света в турбулентной оптически прозрачной среде используется кинетическая модель, в рамках которой поток света рассматривается как поток частиц (фотонов). Дифференциальное сечение рассеяния на единице длины пути $w(\theta)$ выражается через спектральную плотность (трансформанту Фурье) корреляционной функции диэлектрической постоянной $\varepsilon(\mathbf{r})$

$$\Phi_\varepsilon(\mathbf{q}) = \frac{1}{8\pi^3}\int\limits_{\mathrm{R}^3} e^{-i\mathbf{qr}}\langle[\varepsilon(\mathbf{r}_1) - \overline{\varepsilon}][\varepsilon(\mathbf{r}_1+\mathbf{r}) - \overline{\varepsilon}]\rangle d\mathbf{r}.$$

В статистически однородной и изотропной турбулентной среде спектральная плотность зависит только от абсолютной величины векторного аргумента $q = |\mathbf{q}|$, и эта связь выражается соотношением

$$w(\theta) = (1/2)\pi k_0^4\Phi_\varepsilon(q)$$

где $q = 2k_0 \sin(\theta/2)$, а k_0 — волновое число в однородной среде с диэлектрической постоянной $\bar{\varepsilon}$. В инерционном интервале волновых чисел $(1/L, 1/l)$, согласно колмогоровскому закону двух третей,

$$\Phi_\varepsilon(q) = Cq^{-11/3}, \quad C = \text{ const}$$

(уравнение (26.31) книги [5]). Наиболее вероятно рассеяние на малые углы, для которых

$$w(\theta) = (1/2)\pi k_0^4 C[2k_0 \sin\theta/2]^{-11/3} \sim A\theta^{-\alpha-2}, \qquad \theta \to \infty,\ A = \text{const},\ \alpha = 5/3. \tag{3}$$

Пусть в начале координат находится точечный мононаправленный источник света. Выберем координатную ось Z совпадающей с этим первичным направлением $\mathbf{\Omega}_0$. В малоугловом приближении теории рассеяния координата z фотона отождествляется с пройденным им путем, а для характеристики отклонения его от первоначального направления используется двумерный вектор $\mathbf{u} = \{\Omega_x, \Omega_y\}$, область изменения которого расширяется на всю плоскость: $-\infty < u_x < \infty,\ -\infty < u_y < \infty,\ 0$. Соответствующее данному приближению кинетическое уравнение для углового распределения имеет вид

$$\frac{\partial f(\mathbf{u},t)}{\partial t} = \mu \int\limits_{\mathrm{R}^2} [f(\mathbf{u}-\mathbf{u}',t) - f(\mathbf{u},t)] w(|\mathbf{u}'|) d\mathbf{u}', \qquad f(\mathbf{u},0) = \delta(\mathbf{u}),$$

подобный уравнению, описывающему многократное рассеяние заряженных частиц в веществе. Разлагая уменьшаемое в ряд по $\mathbf{u}'$ и выполняя интегрирование (то есть, усредняя по углу рассеяния), приходят к известному уравнению типа Фоккера-Планка для главной асимптотической части решения $f^{\mathrm{as}}(\mathbf{u},t) \equiv I(\mathbf{u},t)$:

$$\frac{\partial I(\mathbf{u},t)}{\partial t} = \mu \frac{\langle u^2 \rangle}{2} \triangle I(\mathbf{u},t), \tag{4}$$

где $\langle u^2 \rangle = \int\limits_{\mathrm{R}^2} w(|\mathbf{u}|) u^2 d\mathbf{u}$ — средний квадрат угла рассеяния в единицу времени. В турбулентном случае этот интеграл расходится, и вместо разложения в ряд мы воспользуемся асимптотической подстановкой (3):

$$\frac{\partial I(\mathbf{u},t)}{\partial t} = \mu A J(\mathbf{u}), \quad J(\mathbf{u}) = \int\limits_{\mathrm{R}^2} [I(\mathbf{u}-\mathbf{u}',t) - I(\mathbf{u},t)] |\mathbf{u}'|^{-\alpha-2} d\mathbf{u}'.$$

При $\alpha \geq 1$ интеграл столкновений расходится, и для устранения этой расходимости воспользуемся процедурой регуляризации Адамара, заменяющей расходящийся интеграл $J(\mathbf{u})$ его конечной (в смысле Адамара) частью

$$\text{p.f. } J(\mathbf{u}) \equiv \int\limits_{\mathrm{R}^2} [I(\mathbf{u}-2\mathbf{u}',t) - 2I(\mathbf{u}-\mathbf{u}',t) + I(\mathbf{u},t)] |\mathbf{u}'|^{-\alpha-2} d\mathbf{u}'$$

Регуляризованный таким образом интеграл выражается через дробную степень двумерного лапласиана соотношением

$$\text{p.f. } J(\mathbf{u}) = -c(\alpha)(-\triangle)^{\alpha/2} I(\mathbf{u},t), \qquad c(\alpha) = \frac{\pi^2(1-2^{1-\alpha})}{[\Gamma(1-\alpha/2)]^2 \sin(\alpha\pi/2)}.$$

Следовательно, асимптотическое поведение углового распределения интенсивности $I(\mathbf{u},t)$ фотонов, регистрируемых в момент t в турбулентной среде, описывается уравнением с двумерным лапласианом дробного порядка

$$\frac{\partial I(\mathbf{u},t)}{\partial t} = -B(-\triangle)^{\alpha/2} I(\mathbf{u},t), \quad I(\mathbf{u},0) = \delta(\mathbf{u}), \tag{5}$$

где $B = c(\alpha)\mu A$. Решение уравнения (5) выражается через двумерную изотропную устойчивую плотность Леви-Фельдгейма с характеристическим показателем $\alpha = 5/3$

$$\Psi_2^{(\alpha)}(r) = \frac{1}{2\pi} \int\limits_0^\infty e^{-k^\alpha} J_0(kr) k dk$$

соотношением:

$$I(\mathbf{u},t) = [Bt]^{-2/\alpha} \Psi_2^{(\alpha)}([Bt]^{-1/\alpha}|\mathbf{u}|). \tag{6}$$

Отметим три характерных отличия рассеяния в турбулентной среде $\alpha < 2$ от аналогичного процесса в среде с мелкомасштабными флуктуациями коэффициента преломления ($\alpha = 2$): ширина углового распределения рассеянных фотонов растет со временем пропорционально $t^{1/\alpha}$, а не $t^{1/2}$, хвосты распределения имеют степенной, а не гауссов вид (в силу чего дисперсия бесконечна), и проекции вектора $\mathbf{u}$ на оси x и y не являются независимыми. Вследствие того, что диффузионный пакет в данной модели расширяется быстрее, чем в случае нормальной диффузии, её часто называют *аномальной супердиффузией* [6]. Заметим также, что при $\alpha = 2$ она превращается в нормальную диффузию, при $\alpha = 1$ превращается в двумерный процесс Коши, а при $\alpha > 2$ решение соответствующего уравнения теряет смысл плотности вероятности. Нельзя не указать на одно важное свойство распределения со степенными хвастами: медленное (по сравнению с привычной экспонентой) их спадание приводит к тому, что на крыльях наблюдается эффект *перемежаемости*: появляются малочисленные группы событий, перемежаемые пустотами. В астрономических измерениях они проявляются как гало, окружающие основное изображение источника и состоящие из одиночных нерегулярно распределённых вспышек — *сцинтилляций*. Именно в этом природа сцинтилляций, а не в особом механизме поглощения. На связь сцинтилляций со степенной (негауссовой) формой крыльев распределения в актах рассеяния указывается в работах [7]- [9].

3 Дробно-дифференциальная субдиффузия

Рассмотренные выше модели обладают одним общим свойством: они *марковские*. Очевиднейшим признаком этого является первая производная по времени, с которой начинаются эти уравнения. Собственные функции этого оператора — экспоненты, описывающие распределение времен между событиями такого процесса. Как и в случае рассмотрения пробегов, экспоненты эти отражают независимость процессов, но теперь уже *временную независимость*, иными словами — *отсутствие памяти*. Чтобы "включить память" в такой процесс, достаточно заменить в диффузионном уравнении (1) первую производную по времени производной дробного порядка:

$$\frac{\partial^\nu}{\partial t^\nu}\left[I(\mathbf{u},t) - I(\mathbf{u},0)1_+(t)\right] = D\triangle I(\mathbf{u},t). \tag{7}$$

Здесь $1_+(t)$ — "ступенька" Хэвисайда, первая производная от которой — дельта-функция, а производная порядка $\nu \in (0,1)$ — её дробный аналог: $\delta_\nu(t) = \frac{t_+^{-\nu}}{\Gamma(1-\nu)}$. Второй момент распределения, даваемого решением уравнения (7),

$$\int u^2 I(\mathbf{u},t)d\mathbf{u} = \frac{4D}{\Gamma(\nu+1)}t^\nu, \tag{8}$$

указывающей на замедленное по сравнением с нормальным расплывание диффузионного пакета, что и оправдывает термин "субдиффузия" [10] (прилагательное "дробно-дифференциальная" указывает на способ реализации замедления, которое может быть достигнуто и другим способом и иметь другие свойства).

Блэкледж также вводит в процесс описания распространения изображения и уравнение с показателем $\nu \in (1,2)$, полагая его некоторым промежуточным вариантом между диффузионным и волновым уравнениями. В обоснование этого он пишет: "Дробно-дифференциальная модель применима к процессам рассеяния в слабой и крайне разреженной среде. В прикладной оптика, один из самых распространенных примеров этого явление имеет место в астрономии и связан с с рассеянием света на межзвёздной пыли, частицы которой представляют собой кластеры от нескольких молекул до размеров порядка 10^{-4} м" [2]. Решение этого уравнения так же записывается в форме (6), только вместо устойчивой плотности $\Psi_2^{(\alpha)}$ фигурирует *дробно-устойчивая плотность* $\Psi_2^{(2,\nu)}$. При $\nu = 1$ она превращается в гауссово распределение, а при меньших значениях имеет острую вершину и характеризуется дисперсией (подробности см. в [10]– [12]). Несколько таких распределений показано на рис. 1.

Если проследить за временной последовательностью событий, то при $\nu = 1$ мы увидим стандартный пуассоновский процесс, в больших масштабах превращающийся в однородный поток с исчезающе малыми флуктуациями, а вот при меньших значениях этого параметра процесс станет *дробно-пуассоновским* [13]: случайные интервалы с более или менее плотным содержанием событий будут перемежаться с почти пустыми, и чем ближе будет ν к единице, тем меньше будет доля пустых интервалов. С математической точки зрения причина такого поведения процесса та же, что и в случае угловых распределений: степенной характер законов, в данном случае — распределений интервалов времени между событиями. Физическая же интерпретация в наличии крупномасштабных корреляций в турбулентной среде, учет которых и приводит к

появлению распределений Миттаг-Леффлера, заменяющих традиционные экспоненты. Эти распределения, при $\nu = 1$ превращающиеся в экспоненты, при других значениях ν характеризуются степенными хвостами.

Наклнец, вполне естественно ввести уравнение, содержащее оба рассмотренных выше дробных оператора. Получаемый в результате процесс называют *дробно-дифференциальной аномальной диффузией* [12, 13].

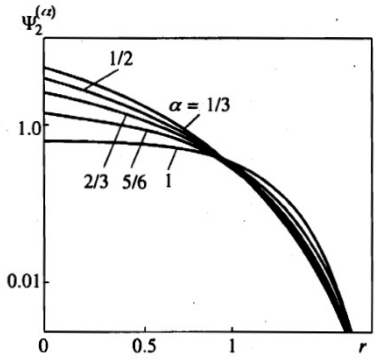


Рис. 1: Распределения $\Psi_2^{(\alpha)}(r)$ для указанных значений α

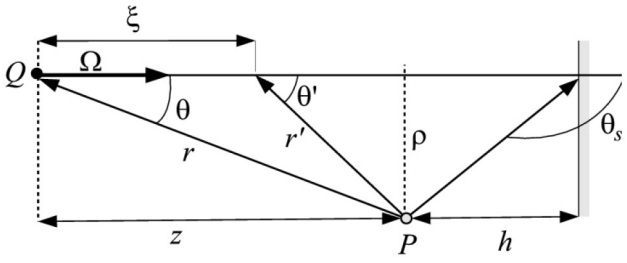


Рис. 2: Геометрия узкого луча (Q — положение источника, P — точка наблюдения)

4 Поле вблизи оси

Рассмотренные выше диффузионные уравнения относились к асимптотически удалённым областям, в больших масштабах отдельные столкновения становились почти неразличимыми, отсюда и диффузионный характер уравнений. В работе [14], рассматривающей комптоновское рассеяние, показано, что в окрестности тонкого луча, испускаемого направленным источником, *на любой глубине* будет доминировать вклад однократного рассеяния. Ограничившись для этой цели рассмотрением интенсивности в точке Q на границе однородной среды, запишем спектр однократно рассеянной интенсивности в координатах, изображённых на рис.2:

$$I_1(r,\theta,E) = \frac{\sigma_s(E_0)}{\rho}\int\limits_{\theta}^{\theta_s} \exp\left\{-\frac{r}{\sin\theta'}\left[\sigma(E_0)\sin(\theta'-\theta)+\sigma(E_1)\sin\theta\right]\right\} w(\cos\theta')E_1(\cos\theta')d\theta'. \tag{9}$$

Здесь E_0 и $E_1 = E_0/[1+(1-\cos\theta')E_0/m_ec^2]$ — энергии первичного и рассеянного на угол θ' квантов, σ и σ_s — полное сечение взаимодействия и сечение комптоновского рассеяния на единице длины пути, $w(\cos\theta')$ — угловая индикатриса комптоновского рассеяния ($\int_{4\pi} w(\cos\theta')d\Omega' = 1$). Разлагая подынтегральную экспоненту в ряд, $I_1 = \sum_i I_1^i$, получим:

$$I_1^{(0)} = \frac{\sigma_s(E)}{\rho}\, e^{-\sigma(E)z}\int\limits_{\theta}^{\theta_s} w(\cos\theta')E_1 d\theta', \tag{10}$$

$$I_1^{(1)} = \sigma_s(E)e^{-\sigma(E)z}\int\limits_{\theta}^{\theta_s} [\sigma(E)\operatorname{ctg}\theta' - \sigma(E_1)\operatorname{cosec}\theta']w(\cos\theta')E_1 d\theta', \tag{11}$$

$$I_1^{(2)} = \frac{\sigma_s(E)}{2}\rho e^{-\sigma(E)z}\int\limits_{\theta}^{\theta_s} [\sigma(E)\operatorname{ctg}\theta' - \sigma(E_1)\operatorname{cosec}\theta']^2 w(\cos\theta')E_1 d\theta', \tag{12}$$

и т.д.

Если детектор находится достаточно далеко от источника и от границы (т.е. $z \gg \rho$ и $h \gg \rho$), то $\theta \approx \rho/z \to 0$, $\theta_s \approx \pi - \rho/h \to \pi$, и следовательно,

$$I_1^0 = \frac{\sigma_s(E)}{\rho}\, e^{-\sigma(E)z}\int\limits_{0}^{\pi} w(\cos\theta')E_1 d\theta', \tag{13}$$

а в случае расположения детектора на границе $\theta_s = \pi/2$.

Пусть $I_0(x,y)$ первичный профиль излучаемого моноэнергетическим источником пучка квантов, а $I(x,y,z)$ — совокупное распределение интенсивности плоскости, перпендикулярной оси пучка и находящейся на расстоянии z от источника. Как сказано выше, оно состоит из нерассеянной компоненты, повторяющей форму излучаемого источником пучка, но ослабленную на экспоненциальный множитель, учитывающий выход квантов за пределы пучка вследствие рассеяния,

$$I^{\mathrm{nsc}}(x,y,z) = I_0(x,y)e^{-\mu z},$$

и рассеянной компоненты

$$I^{\mathrm{sc}}(x,y,z) = \mu_s \Theta e^{-\mu z} \int_A \int \frac{I_0(x',y')dx'dy'}{\sqrt{(x-x')^2+(y-y')^2}}, \tag{14}$$

где

$$\Theta(E) = \frac{2\pi}{E} \int_0^{\pi/2} w(\cos\theta')E_1(\cos\theta')d\theta',$$

а интегрирование ведётся по поперечному сечению A выходящего из источника луча.

В случае однородного распределения нерассеянного излучения по круговому сечению радиуса a

$$\sigma(r) = \begin{cases} \sigma_0, & r < a; \\ 0, & r > 0, \end{cases}$$

and

$$\varphi(r) = \int_0^a \int_0^{2\pi} \frac{\sigma_0 r'dr'd\vartheta}{\sqrt{r^2+r'^2-2rr'\cos\vartheta}}.$$

Интегрирование по r' даёт

$$\varphi(r) = 2\sigma_0 \int_0^{\pi} \Big(\sqrt{r^2+a^2-2ra\cos\vartheta} - r + r\cos\vartheta \ln \frac{a - r\cos\vartheta + \sqrt{r^2+a^2-2ra\cos\vartheta}}{r(1-\cos\vartheta)} \Big) d\vartheta.$$

Далее, вводя $\chi = \frac{\pi-\vartheta}{2}$ и интегрируя член с логарифмом по частям, получаем

$$\varphi(r) = 2\sigma_0 \left[(a+r)E(k) + (a-r)K(k) \right], \quad k = \frac{2\sqrt{ra}}{r+a}, \tag{15}$$

где K и E — полные эллиптические интегралы первого и второго рода соответственно. При $r = 0 \quad E = K = \pi/2$, и формула (15) принимает вид $\varphi(0) = 2\pi a\sigma_0$.

Заключение

Мощные и сверхмощные источники электромагнитных излучений, знакомством с которыми одарил нас минувший век, эти маяки, большей частью пульсирующие или ометающие вращающимися лучами Вселенную, несут колоссальный объём информации о межзвёздной и межгалактической среде, извлечение которой требует совершенствования математических моделей как самих этих сред, так и процессов распространения излучений в них. Всего лишь полвека назад модель, предложенная Шойером [15], заменившим всю пересекаемую радио-излучением среду одним тонким экраном со случайно распределенным коэффициентом преломления, казалась достаточной для оценки формы и времени задержки измеряемых радио-импульсов таких источников, но в наши дни наземная и внеземная астрономические службы способны представить информацию, требующую для её осмысления более совершенных математических методов и моделей.

В настоящем докладе показано, каким образом учет важнейшей (хотя, разумеется, не единственной) особенности межзвёздной среды — многокомпонентности и турбулентности — модифицирует математический аппарат, выводит на сцену нелокальные в пространстве-времени дробно-дифференциальные операторы, степени которых связаны с показателями степенных законов развитой турбулентности, сформулированных А.Н.Колмогоровым и подтверждаемых сегодня многочисленными астрофизическими наблюдениями. Изложенный здесь аппарат вполне пригоден для изучения влияния многократного рассеяния на величину и форму наблюдаемых импульсов, так и для восстановления свойств среды, сквозь которую проходит регистрируемое излучение [16,17].

Список литературы

[1] Taylor J. H. andCordes J.M. Pulsar distances and the galactic distribution of free electrons // Astrophys. J. 1993. V. 441, P.674–684.

[2] Blackledge J.M. Diffusion and Fractional Diffusion Based Models for multiple light scattering and image analysis // Isast Trans. Electron and Signal Processing. 2007. V. 1, iss. 1.P. 1–22.

[3] Kardashev N.S., Khartov V.V., Abramov V.V., et al. RadioAstron'-A telescope with a size of 300 000 km: Main parameters and first observational results // Astronomy Reports. 2013. V. 57. P. 153–194.

[4] Dudorov A.E. Properties of the hierarchy of interstellar magnetic clouds // Astron. Zh. 1991. V. 68, P. 695–708.

[5] Рытов С.М., Кравцов Ю.А., Татарский В.И. Введение в статистическую радиофизику. Том 2. М.: Наука, 1978.

[6] Золотарев В.М., Учайкин В.В., Саенко В.В. Супердиффузия и устойчивые законы // ЖЭТФ. 1999. V. 115, P. 1411–1425.

[7] Boldyrev S., Koenigl A. Non-Gaussian radio-wave scattering in the interstellar medium // The Astrophysical Journal. 2006. V. 640, P. 344–352.

[8] Boldyrev S., Gwinn C. Scintillations and Lévy flights through the interstellar medium // The Astrophysical Journal. 2008. V. 584. iss. 2. P.791–796.

[9] Scalo J., Elmegreen B.G. Interstellar turbulence II: Implication and Effects // Annu. Rev. Astron. Astrophys. 2004. V. 42. P. 275–316.

[10] Учайкин В.В. Субдиффузия и устойчивые законы // ЖЭТФ. 1999. V. 115, P. 2113–2132.

[11] Учайкин В.В., Коробко Д.А. Фрактальная модель переноса: малоугловое приближение // ЖТФ. 2004. V. 74. P. 12–19.

[12] Учайкин В.В. Аномальная диффузия и дробно-устойчивые распределения // ЖЭТФ/ 2003. V. 124, P. 903–920.

[13] Учайкин В.В. Метод дробных производных, Ульяновск: Артишок, 2008.

[14] Uchaikin V.V. A gamma-ray in a uniform Compton scatterer // Phys. Astron. Int. J. 2018. V. 2. iss. 6, P. 573–576.

[15] Scheuer P.A.G. Amplitude variations in pulsed radio sources // Nature. 1968. V. 218. P. 920–922.

[16] Siyao Xu, Bing Zhang, Scatter broadening of pulsars and implications on the interstellar medium turbulence // The Astrophysical Journal. 2017. V. 835. iss. 2, P. 1–11.

[17] Ue-Li Pen, Levin Yu. Pulsar scintillations from corrugated reconnection sheets in the interstellar medium // Monthly Notices of the Royal Astronomical Society. 2014, V. 442, P. 3338–3346.

Учайкин Владимир Васильевич — д.ф.-м.н., заведующий кафедрой теоретической физики;
Ульяновский государственный университет;
e-mail: vuchaikin@gmail.com.
Дата поступления — 30 апреля 2019 г.

ДРОБНО-ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ В СТАТИСТИКЕ АФТЕРШОКОВ

В. В. Учайкин, Е. В. Кожемякина

Ульяновский государственный университет, 432017, Ульяновск

УДК 517.925
DOI: 10.24411/9999-016A-2019-10084

Доклад посвящен статистической модели последовательности повторных толчков (афтершоков) после сильного землетрясения. В основу её положены три основных закона, относящихся к распределению временных промежутков между афтершоками, энерговыделениям и фрактальным свойствам пространственного распределения толчков. Основное внимание уделяется временным характеристикам последовательности. Для математического описания используются уравнения с дробными производными. Излагаются основные свойства дробно-пуассоновского процесса. Выведено предельное распределение числа событий в таком процессе.
Ключевые слова: Афтершок, дробно-пуассоновская модель.

Введение

Сильные землетрясения всегда сопровождаются многочисленными *афтершоками* — повторными толчками меньшей амплитуды, нерегулярно распределёнными во времени и пространстве. Частота и магнитуда афтершоков со временем уменьшаются, а продолжительность проявления может длиться месяцами. Особенно велика вероятность сильных афтершоков в первые часы после главного толчка. Известно много случаев, когда повреждённые главным ударом здания рушились именно при повторных, менее сильных толчках. Афтершоки представляют существенную угрозу и при проведении спасательных работ.

Качественно, афтершоки можно считать локальными во времени и (с некоторой степенью условности) в пространстве, и рассматривать как случайный процесс, который можно сформулировать следующим образом.

1. Запускающий первичный толчок вызывает множество вторичных толчков, следующих друг за другом в случайные моменты времени $t_1, t_2, t_3, \dots$

2. Вторичные толчки распределяются в пространстве случайным образом, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots$, удаляясь (в среднем) от точки начального толчка $\mathbf{x}_0 = 0$.

3. В каждом из событий выделяется случайная энергия $E_i,\ i = 1, 2, 3, \dots$

Будем, для краткости, обозначать набор переменных $\{E, \mathbf{x}, t\}$ одной буквой x. Каждая последовательность перечисленных случайных элементов порождает в фазовом пространстве последовательность случайных точек $\{x_1, x_2, x_3, \dots\}$ Пренебрегая, в первом приближении, корреляциями между последовательными приращениями фазовых координат процесса, можно описать каждый такой процесс как траекторию марковской (возможно, обрывающейся) цепи в энерго-координатно-временном фазовом пространстве (рис. 1). Эта программа реализована в серии работ Хелмстеттера и Сорнетта (см. статью [1] и библиографию в ней). В рамках данной модели средняя плотность толчков $N(x)$ в однородной безграничной сейсмо-активной среде удовлетворяет интегральному уравнению

$$n(x) = q \int dx' n(x') P(x' \to x) + \delta(x - x_0), \tag{1}$$

ISBN 978-5-901548-42-4

описывающему случайное блуждание «частицы» по траектории типа изображённой на рис. 1 «Частицы» эти не являются реальными образованиями. Будучи сугубо виртуальными (воображаемыми), они просто представляют собой простейший способ введения реально существующих корреляций между последовательными толчками. Множитель $q \in (0,1)$ задаёт вероятность обрыва последовательности в данном узле.

Имеющихся в настоящее время данных недостаточно для определения ядра этого уравнения, да и не может быть какого-то общего представления его безотносительно к географическому положению процесса. Качественное же его описание основывается на трёх ниже-перечисленных законах, относящихся к маргинальным распределениям, и могущих быть использованными для факторизации ядра, как приближённого представления интегрального члена, записанного в пренебрежении корреляциями как между *приращениями* пяти компонент фазовой переменной x на каждом из отрезков, так и между их приращениями на разных отрезках:

$$P(E', \mathbf{x}', t' \to E, \mathbf{x}, t) = P(E' \to E) P(\mathbf{x}' \to \mathbf{x}) P(t' \to t).$$

Полагая (иключительно из прагматических соображений) каждый из сомножителей трансляционно-инвариантным, описывающим однородное блуждание по своей переменной,

$$P(E', \mathbf{x}', t' \to E', \mathbf{x}', t') = \chi(E - E')\phi(\mathbf{x} - \mathbf{x}')\psi(t - t'), \tag{2}$$

получим три уравнения восстановления: временное уравнение

$$f(t) = \int_0^t \psi(t - t') f(t') dt' + \psi(t), \tag{3}$$

пространственное уравнение

$$f(\mathbf{x}) = \int_{R^3} \phi(\mathbf{x} - \mathbf{x}') f(\mathbf{x}') d\mathbf{x}' + \phi(\mathbf{x}), \tag{4}$$

и энергетическое уравнение

$$f(E) = \int_0^E \chi(E - E') f(E') dE' + \chi(E). \tag{5}$$

Выполненная выше факторизация (принудительное разделение уравнений), переводит задачу в разряд качественных, решение которых по понятным причинам не нуждается в разработке специальных аналитических или численных методов: слишком неопределёнными являются граничные и начальные условия, строение среды, да и сам механизм возникновения и влияния друг га друга повторных толчков строго говоря, неизвестен. Можно лишь ориентироваться на некоторые частные, но понятные нам модели, которыми часто сопровождается изложение основных законов динамики афтершоков, установленных сугубо эмпирически на основе анализа большого объёма статистических данных. Наш доклад и посвящён обсуждению некоторых задач в очень элементарной постановке, позволяющей, впрочем, непосредственно опираться на эти основные законы.

1 Основные законы динамики афтершоков.

1. **Закон Омори–Утсу** [2], [3]: среднее число толчков в единицу времени убывает при $t \to \infty$ по степенному закону $t^{-\nu-1}$, $\nu \approx 0,25$ [4].

2. **Закон Гутенберга–Рихтера** [5] : среднее число толчков с энерговыделением в единичном интервале около E убывает по степенному закону $E^{-\beta}$, $\beta \cong 2/3$ [1].

3. **Фрактальность пространственного распределения**: пространственное распределение афтершоков в области 10–20 эпицентра имеет вид фрактального кластера с фрактальной размерностью $D_f \approx 2,2$ [6]. Последнее означает, что в пространстве целой размерности d плотность афтершоков должна спадать с расстоянием от первичного толчка как $r^{D_f - d}$ В статье [4], впрочем, отмечается, что большинство данных указывает на то, что $D_f < 2$.

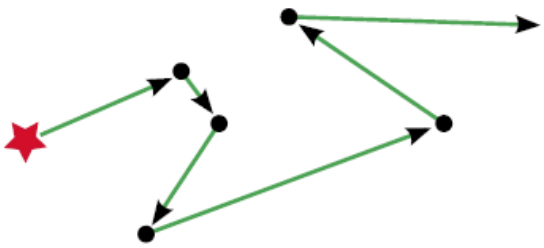
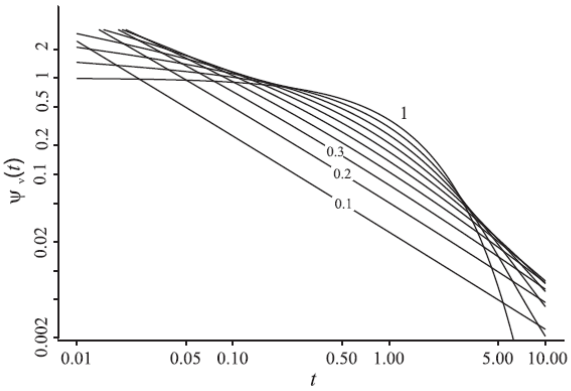


Рис. 1: Афтершоки как узлы траектории



$\propto t^{\nu-1},\ t \to 0,$ $\propto t^{-\nu-1},\ t \to \infty$

Рис. 2: Плотность распределения временных промежутков между толчками

По-существу, все три закона являются проявлением свойства *фрактальности* в разных переменных. Напомним, что множество случайных точек в R^d образует случайный (стохастический) фрактал (кластер) с фрактальной размерностью $D_f < d$, если среднее число точек, находящихся в сферической окрестности U_R радиуса R с центром в одной из точек этого множества, $\langle N(R)\rangle \propto R^{D_f}$.

В приведённых выше законах речь идет о распределениях случайных величин, но не говорится явно об их корреляциях — как между собой в отдельном событии, так и в последовательных событиях. Тем не менее, фрактальный характер последовательностей недвусмысленно указывает на наличие сильных корреляций. Простейший способ ввести корреляции заключается в использовании *модели скачкообразных блужданий* с независимыми (в статистическом смысле) приращениями этих характеристик, иными словами, модели *однородного процесса восстановления*. Эту модель, применительно к отдельным характеристикам процесса, и описывают интегральные уравнения (3)–(5). Остаётся согласовать ядра этих уравнений с основными законами динамики афтершоков. Это можно сделать на основе работ: [7], [8] и [9]. В них показано, что для получения решений уравнений (3)–(5), обладающих фрактальными свойствами с размерностью D_f, достаточно, чтобы плотность распределения случайных расстояний между соседними узлами имела степенную асимптотику порядка $-D_f - d,\ (d-1 < D_f < d,\ d = 2, 3.)$.

2 Дробно-пуассоновская модель последовательности афтершоков

Наиболее популярной в теоретической физике моделью потока случайных событий (импульсов, толчков), своего рода стандартной точкой отсчёта в пространстве случайных моделей одномерных точечных процессов является *пуассоновский процесс* (ПП). Он принадлежит к семейству процессов восстановления, определяемому уравнением (3) с ядром $\psi_1(t) = \mu_1 e^{-\mu_1 t}$. Сделав эту подстановку в уравнение (3),

$$f_1(t) = \int_0^t \psi_1(t-t')f_1(t')dt' + \psi_1(t),$$

получим равномерное по времени (в среднем) распределение толчков $f_1 = \mu_1$, что противоречит наблюдаемым последовательностям афтершоков. Впрочем, и использованное нами экспоненциальное распределение случайных промежутков противоречит 1-му закону динамики афтершоков. Существует, однако, семейство распределений, основанное на двухпараметрической функции Миттаг-Леффлера и включающее в себя экспоненциальное как частный случай, и в то же время, содержащее континуум распределений, характеризующихся степенными хвостами (см.рис. 2):

$$\psi_\nu(t) = \mu t^{\nu-1} \sum_{j=0}^{\infty} \frac{(-\mu t^\nu)^j}{\Gamma(\nu j + \nu)} = \mu t^{\nu-1} E_{\nu,\nu}(-\lambda t^\nu), \quad t > 0. \tag{6}$$

Используя это распределение в качестве переходной вероятности,

$$f_\nu(t) = \int\limits_0^t \psi_\nu(t-t')f_\nu(t')dt' + \psi_\nu(t), \tag{7}$$

и выполняя преобразование Лапласа полученного уравнения (6), найдём трансформанту его решения в окрестности нуля по лапласовской переменной:

$$\hat{f}_\nu(\lambda) = \int\limits_0^\infty e^{-\lambda t} f_\nu(t)dt \sim \mu\nu\Gamma(\nu)\lambda^{-\nu}, \quad \lambda \to 0.$$

Применяя, наконец, Тауберовы теоремы, получим в асимптотике больших времён:

$$f_\nu(t) \sim \mu\nu t^{\nu-1}, \quad \langle N_\nu(t)\rangle \sim \mu t^\nu, \quad t \to \infty.$$

Случайный процесс, описываемый уравнением восстановления (6) с переходной плотностью вероятности (6) называют *дробно-пуассоновским процессом* порядка ν ([10], [11]). Происхождение этого термина связано с дробно-дифференциальным характером уравнений, описывающих этот процесс:

3 Дробно-дифференциальные уравнения

Случайный процесс, описываемый уравнением восстановления (6) с переходной плотностью вероятности (7) называют *дробно-пуассоновским процессом* (ДПП) порядка ν ([10], [11], [2]). Происхождение этого термина связано с дробно-дифференциальным характером уравнений, описывающих этот процесс:

$${}_0\mathsf{D}_t^\nu \psi_\nu(t) + \mu\psi_\nu(t) = \mu\delta(t), \tag{8}$$

где

$${}_0\mathsf{D}_t^\nu \psi_\nu(t) \equiv \frac{d}{dt}\int\limits_0^t \left\{\frac{(t-t')^{-\nu}}{\Gamma(1-\nu)}\right\} f(t')dt', \qquad 0<\nu<1,$$

— *производная Римана–Лиувилля дробного порядка* ν. Содержимое фигурных скобок здесь представляет обобщённую функцию, при $\nu \to 1$ переходящую в дельта-функцию $\delta(t-t') = \frac{d}{dt}1_+(t)$. Аналогичным образом, известные уравнения для вероятностей в пуассоновском процессе

$$p_n(t) \equiv \mathsf{P}(N(t)=n), \ n=0$$

в ПП

$$\frac{d\left[p_n(t) - \delta_{n0}1_+(t)\right]}{dt} = \mu[p_{n-1}(t) - p_n(t)]$$

обобщаются на случай ДПП,

$${}_0\mathsf{D}_t^\nu\left[p_n(t) - \delta_{n0}1_+(t)\right] = \mu[p_{n-1}(t) - p_n(t)]$$

или, в эквивалентном виде

$${}_0\mathsf{D}_t^\nu p_n(t) = \mu[p_{n-1}(t) - p_n(t)] + \frac{t^{-\nu}}{\Gamma(1-\nu)}\delta_{n0}, \qquad 0<\nu<1.$$

4 Дробно-пуассоновский процесс

Аналитические исследования свойств ДПП выполнялись рядом авторов (см. библиографию в [11]). В этом рпзделе мы кратко опишем обладаемые большей наглядностью численные результаты моделирования методом Монте–Карло с помощью **алгоритма В.В. Учайкина**, полученного модификацией известного алгоритма Кантера [12].

4.1 Моделирование интервала времени

Случайная величина $T(\nu)$ с плотностью распределения (6), удовлетворяющей уравнению (8) моделируется с помощью алгоритма

$$T = \frac{|\ln U_1|^{1/\nu}}{\mu^{1/\nu}} S(\nu) = \frac{|\ln U_1|^{1/\nu}}{\mu^{1/\nu}} \frac{\sin(\nu\pi U_2)[\sin((1-\nu)\pi U_2)]^{1/\nu-1}}{[\sin(\pi U_2)]^{1/\nu}[\ln U_3]^{1/\nu-1}}, \tag{9}$$

где $S(\nu)$ — субординатор порядка ν (положительная случайная величина с односторонним устойчивым распределением с характеристическим показателем ν), U_1, U_2 и U_3 — независимые, равномерно распределенные на $[0,1]$ случайные числа [12]. Стоит заметить, что при $\nu = 1$ этот алгоритм превращается в известную формулу моделирования эгспоненциально распределённой случайной величины: $T = \frac{|\ln U_1|}{\mu}$.

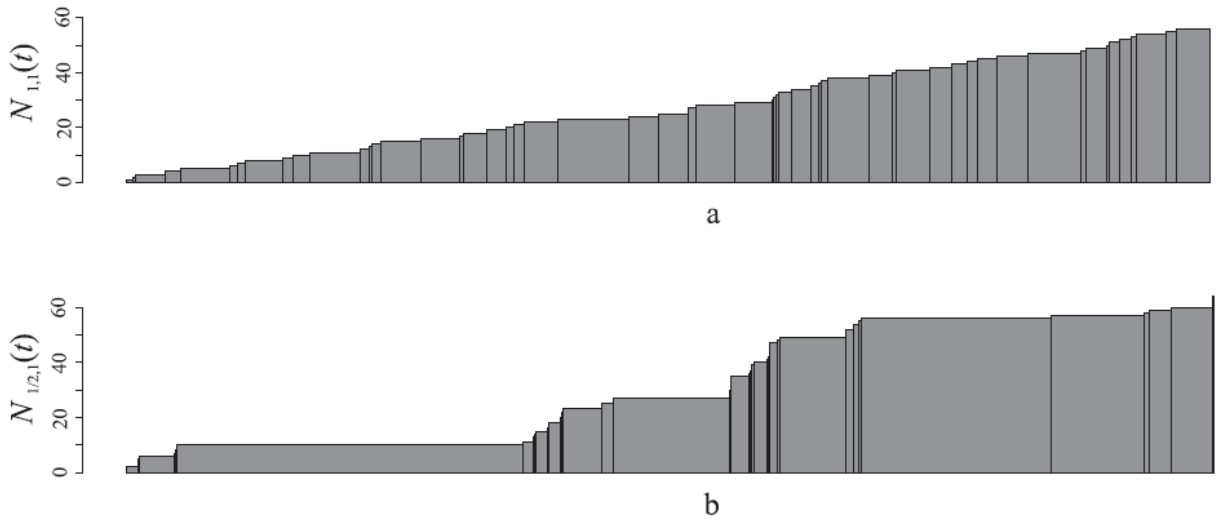


Рис. 3: Случайные реализации $N_{\nu,\mu}(t)$ пуассоновского (a) и дробно-пуассоновского (b) процессов

Графики реализаций $N_{\nu,\mu}(t)$ ПП $(\nu = 1, \mu = 1)$ (a) и ДПП $(\nu = 1/2, \mu = 1)$ (b), представленные на рис. 3, имеют вид ступенчатых кривых с единичными ступеньками, расположенными в случайных точках. При $\nu = 1$ длины разных ступенек разброс, но разброс этот не критический: нет ни очень коротких ступенек (точнее они маловероятны), ни очень длинных, и, если шагать по ним, то можно как-то приноровиться. При $\nu < 1$ ситуация резко меняется: у плотности распределения длины ступеньки появляются особенности в нуле, порождая сгущения ступенек (кластеры), и тяжелый степенной хвост, ответственный за появление длинных ступенек, усыпляющих бдительность бегущего. Именно по этой причине математическое ожидание случайного интервала между событиями ДПП бесконечно: $\mathsf{M}T = \infty$.

4.2 Время n-го толчка

В рамках рассматриваемой модели распределение времени n-го толчка определяется n — кратной свёрткой распределений первого толчка: в ПП это было бы гамма-распределение

$$\psi_1^{*n}(t) = \mu \frac{(\mu t)^{n-1}}{(n-1)!}\, e^{-\mu t},$$

а в ДПП-версии —

$$\psi_\nu^{*n}(a_n t) \sim \Psi_n(t;\nu), \qquad n \to \infty,$$

где $\Psi_n(t;\nu)$ односторонняя устойчивая (в смысле Леви) плотность распределения с характеристическим показателем ν, задаваемая трансформантой Лапласа

$$\hat{\Psi}_n(t;\nu) = e^{-n(\lambda\tau)^\nu},$$

где τ — характерный масштаб времени одного интервала.

4.3 Перемежаемость

В обычном пуассоновом потоке событий укрупнение масштаба (эквивалентное увеличению интенсивности μ) ведет к сглаживанию флуктуаций, случайная реализация все больше сближается с линейной функцией, представляющей среднее число толчков к моменту времени окончания подсчёта. Начиная с какого-то большого масштаба их различием (то-есть, статистическими флуктуациями процесса) можно пренебречь. Иначе дело обстоит в дробно-пуассоновском потоке: фрактальность есть самоподобие в разных масштабах, и если мы наблюдаем какое-то свойство при одном масштабе, оно должно наблюдаться и при других. Стало быть, и флуктуации не должны исчезать (и не исчезают!) с увеличением масштаба: в любом масштабе наблюдаются сгущения импульсов перемежаемые разреженными, почти пустыми промежутками. Это хорошо видно на рис. 4, показывающем, как влияет изменение масштаба (или, что то же, интенсивности) в 100 раз, на распределение числа событий процесса по равным временным отрезкам (бинам). В случае ПП (левая пара гистограмм) мы видим заметное влияние процедуры изменения масштаба, тогда как в случае LGG (правая пара гистограмм) качественная разница практически не видна.

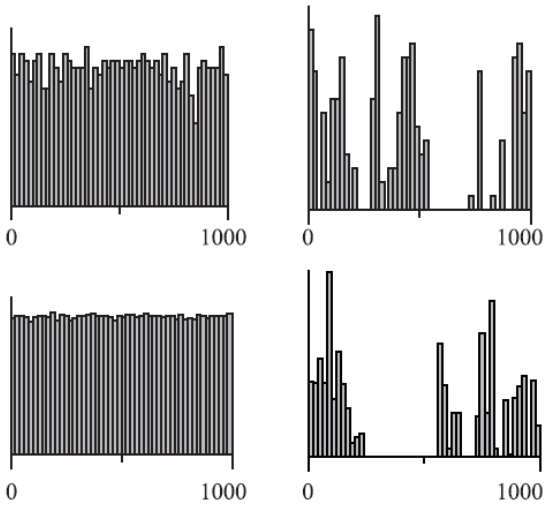


Рис. 4: Изменение интенсивности μ в 100 раз заметно сглаживает гистограмму ПП (левая пара), но мало влияет на ДПП-гистограмму (правая пара)

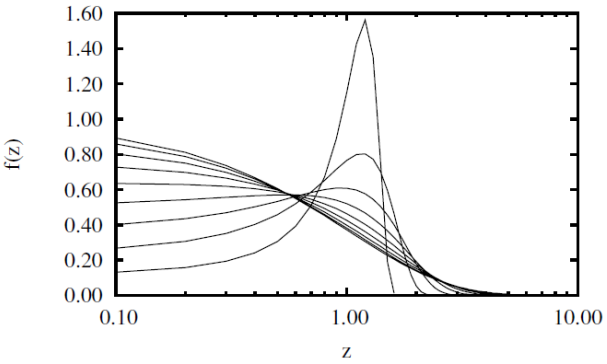


Рис. 5: Плотность предельного (при $t \to \infty$) распределения скейлинговой переменной $Z = N/\bar{n}$ при $\nu = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ (снизу вверх)

4.4 Предельное дробно-пуассоновское распределение.

Напомним, что распределение числа событий в ПП за время t подчиняется закону Пуассона:

$$p_n(t) = \frac{\bar{n}^n}{n!}e^{-\bar{n}}, \qquad \bar{n} \equiv \langle N(t)\rangle = \mu t.$$

Введя квазинепрерывную скейлинговую переменную $Z = N/\bar{n}$, получим для ее плотности

$$f(z;\bar{n}) = \bar{n}\frac{\bar{n}^{\bar{n}z}}{\Gamma(\bar{n}z+1)}e^{-\bar{n}} \sim \sqrt{\frac{\bar{n}}{2\pi}}\exp\left\{-\frac{(z-1)^2}{2/\bar{n}}\right\} \to \delta(z-1), \quad \bar{n}\to\infty.$$

Это просто алгебраический перевод фразы: при указанном условии флуктуации пренебрежимо малы, и переменная Z может считаться детерминированной. Интересно отметить, что в ДПП распределение аналогичной скейлинговой переменной ($\bar{n} = \mu t^{\nu}$) при $t \to \infty$ не вырождается, а выражается через одностороннюю устойчивую плотность $g_{+}^{(\nu)}(x)$ простой формулой

$$f_\nu(z;\bar{n}) \to f_\nu(z) = \left\{ \frac{[\Gamma(\nu+1)]^{1/\nu}}{\nu} g_{+}^{(\nu)} \left(\left(\frac{\Gamma(\nu+1)}{z} \right)^{1/\nu} \right) z^{-1-1/\nu} \right\},$$

допускающей к тому же элементарное представление моментов $\langle Z^k \rangle = [\Gamma(1+\nu)]^k \Gamma(1+k)/\Gamma(1+\nu k)$

$$\langle Z^k \rangle = \frac{[\Gamma(1+\nu)]^k \Gamma(1+k)}{\Gamma(1+\nu k)}.$$

Заключение

Предсказание геологических явлений типа землетрясений, извержений вулканов, возникновения цунами затрудняется сложностью всей системы, существованием далеких и подчас скрытых корреляционных связей, зависимостью процессов не просто от начальных условий, а от предыстории самой системы и её окружения. В отсутствие законченных теорий приходится оперировать с моделями, и модели сами могут формироваться на основе различных подходов. Так, к решению рассматриваемой нами задачи привлекаются механические модели, состоящие из блоков, пружин и демпферов, континуальные модели (относительное движение и взаимная деформация слоев), динамические модели типа теории катастроф, синергетики, нелинейной динамики. В нашей статье продемонстрирована *статистическая модель*, эффективность которой объясняется выходом за пределы гауссовой статистики, использование теории устойчивых (по Леви) законов, изначально приспособленных к описанию катастрофических отклонений, процессов с памятью, и требующих для этого описания специальные математические средства: фрактальный анализ, дробные производные, нелокальные операторы, вероятностные распределения с бесконечными моментами. Некоторые из этих понятий использованы в настоящей работе.

Список литературы

[1] Helmstetter A., Sornette D. Diffusion of epicenters of earthquake aftershocks, Omori's law, and generalized continuous-time random walk models // Phys.Rev. E. 2002. V. 66, P. 061104.

[2] Учайкин В.В., Учайкин Д.В. Дробно-дифференциальная модель землетрясения // Обозрение прикладной и промышленной математики. 2009. Т. 16, вып. 2, С. 392.

[3] Utsu, T., A statistical study on the occurrence of aftershocks // Geophys. Mag. 1961. V. 30, P. 521–605.

[4] Gu C., Schumann A., Baiesi M. et all. Triggering cascades and statistical properties of aftershocks // J. Geophys. Res. Earth. 2013. V. 118, P. 4278–4295.

[5] Gutenberg B., Richter C. F. Magnitude and Energy of Earthquakes // Annali di Geofisica. 1956. V. 9. P. 1–15

[6] Kagan Y. Y., Knopoff L. Statistical study of the occurrence of shallow earthquakes //Geophys. J. R. Astron. Soc. 1978. V. 55. P. 67–86.

[7] Учайкин В.В. Автомодельная аномальная диффузия и устойчивые законы // Успехи физических наук. 2003. Т. 173, вып. 8, С. 524–557.

[8] Слободенюк В.А., Учайкин В.В. О фрактальном характере асимптотики решения интегрального уравнения с устойчивым (по Леви) ядром. В сб.: Теоретическая и экспериментальная физика. Ученые записки УлГУ, вып. 2, Изд-во СВНЦ, С. 70-81, 1995.

[9] Uchaikin V.U., Gusarov G.G., Levy-flight applied to random media problems // J.Math.Phys. 1997. V. 38. iss. 5. P. 2453–2464.

[10] Uchaikin V.V., Cahoy D.O., Sibatov R.T. Fractional processes: From Poisson's to branching ones // Intern. J. Bifurcations and Chaos. 2008. V. 18, P. 2717–2725.

[11] Учайкин В.В. Метод дробных производных, Ульяновск: изд-во Артишок, 2008.

[12] Uchaikin V.V., Zolotarev V.M. Chance and Stability.Netherlands, Utrecht, VSP, 1999.

Учайкин Владимир Васильевич — д.ф.-м.н., заведующий кафедрой теоретической физики;
Ульяновский государственный университет;
e-mail: vuchaikin@gmail.com;
Кожемякина Елена Владиславовна — старший преподаватель кафедры теоретической физики;
Ульяновский государственный университет;
e-mail: elvk@mail.ru.
Дата поступления — 30 апреля 2019 г.

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПЕРЕНОСА ОПТИЧЕСКИХ СИГНАЛОВ ЧЕРЕЗ ВЗВОЛНОВАННУЮ ПОВЕРХНОСТЬ

О. С. Ухинова[1], Б. А. Каргин[1,2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 517.968
DOI: 10.24411/9999-016A-2019-10085

В работе рассмотрены две модели оценки оптической передаточной функции системы океан-атмосфера, предложенные в [5]. После определенных уточнений были проведены численные эксперименты с использованием второй из предложенных моделей (для первой модели численные эксперименты были проведены также в работе [5]). Было проведено сравнение предложенных методов оценки, которое показало преимущество последней модели над предыдущей.

Ключевые слова: метод Монте-Карло, система океан-атмосфера, оптическая передаточная функция.

Введение

Строгое решение подавляющего числа прямых задач оптики системы океан-атмосфера возможно лишь в статистической формулировке. Это обусловлено тем, что поле оптического излучения системы океан-атмосфера в значительной мере формируется под воздействием взволнованной морской поверхности, отражающей и преломляющей приходящее электромагнитное излучение. Характерная особенность этого воздействие обусловлена случайным характером ветрового волнения морской поверхности. В работе для описания взволнованной поверхности моря принята так называемая плоская “фацетная” модель, которая впервые была применена к расчету переноса солнечной радиации в системе атмосфера-океан в [1, 2]. Также в работе рассмотрены алгоритмы метода Монте-Карло для моделирования переноса излучения через взволнованную морскую поверхность в применении его для вычисления оптической передаточной функции (ОПФ) системы океан-атмосфера. Представлены некоторые количественные оценки влияния ряда факторов на ОПФ системы для одной из моделей, а также проведено сравнение двух рассмотренных моделей в применении в вычислению оценок ОПФ для различных скоростей ветра.

1 Постановка задачи

Постановка задачи полностью совпадает с задачей в работе [5]. Верхняя граница $z = H$ плоской атмосферы освещается параллельным потоком фотонов $I_0\delta(\omega - \omega_0)$ в напралении ω_0. Взаимодействие света с веществом атмосферы и океана определяется заданием коэффициентов ослабления $\sigma(z)$ и рассеяния $\sigma_s(z)$, индикатрисы рассеяния $g(z, \mu)$, нормированной условием $\int_{-1}^{1} g(z, \mu)d\mu = 1$, где μ — косинус угла рассеяния. Введем обозначения: $r_\perp = (x, y)$ — проекция радиуса-вектора $r = (x, y, z)$ на горизонтальную плоскость; $\omega = (a, b, c)$ — единичный вектор направления с направляющими косинусами a, b, c; Ω — множество всех направлений с $c \in [-1, 1]$, Ω_- — единичная полусфера с $c \in [0, 1]$, Ω_+ — единичная полусфера с $c \in [-1, 0]$. При попадании света на плоскость $z = 0$ он испытывает отражение, которое характеризуется коэффициентом яркости $Q(r_\perp, \omega, \omega')$, равным отношению яркости света, идущего из точки $(r_\perp, 0)$ в направлении ω к освещенности единичной площадки в точке $(r_\perp, 0)$, ориентированной перпендикулярно к ω'. В плоскости $z = h$

Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2019-0002).

ISBN 978-5-901548-42-4

расположена граница раздела двух сред, при взаимодействии с которой свет испытывает преломление и отражение. Эта граница представляет собой случайную поверхность, составленную из набора элементарных площадок, центры которых лежат в плоскости $z = h$, а нормали к площадкам l — случайные единичные векторы с плотностью распределения вероятности $p(l)$, $l = (l_x, l_y, l_z) \in \Omega_-$ (для такой модели взволнованной морской поверхности в соответствующей литературе утверждено название “фацетная модель”).

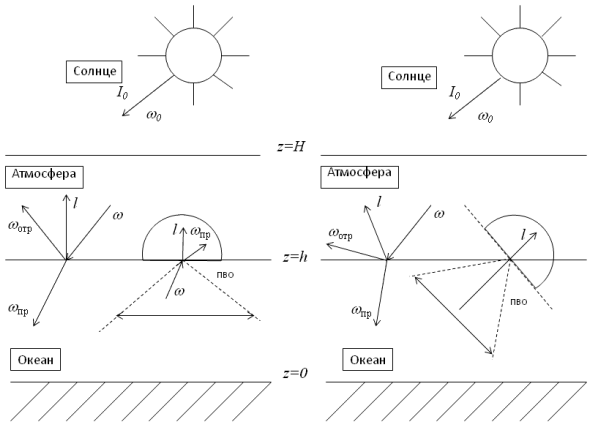


Рис. 1: Схема просесса переноса излучения через границу раздела двух сред 1 — детерминированная плоская граница, 2 — случайная граница.

Обозначим через $\Phi(r_\perp, z, \omega)$ полную яркость светового поля в точке $(r_\perp, z)$ в направлении ω. Световой режим в плоскопараллельной среде, освещенной параллельным потоком фотонов, описывается известным интегро-дифференциальным уравнением переноса с соответствующими условиями (см. [3], стр. 5). Для описания светового поля в системе океан-атмосфера к этим краевым условиям необходимо добавить выражение, определяющее взаимодействие света со случайной границей раздела двух сред:

$$\Phi(r_\perp, h, \omega)\Big|_{\omega\in\Omega} = \int_{\Omega_-} dl \int_{\Omega} p(l)R(\omega', l)\Phi(r_\perp, h, \omega')\delta[\omega' - \omega_{\text{отр}}(\omega, l)]d\omega' + \tag{1}$$
$$+\nu^2 \int_{\Omega_-} dl \int_{\Omega} p(l)[1 - R(\omega', l)]\Phi(r_\perp, h, \omega')\delta[\omega' - \omega_{\text{пр}}(\omega, l)]d\omega'$$

Здесь $R(\omega, l)$ — френелевский коэффициент отражения,

$$\omega_{\text{отр}}(\omega, l) = \omega - 2(\omega, l)l,$$

$$\omega_{\text{пр}}(\omega, l) = \nu(\omega + \gamma l),$$

$$\gamma = \text{sign}[(\omega, l)]\sqrt{\frac{1}{\nu^2} - 1 + (\omega, l)^2} - (\omega, l),$$

$$\nu = \begin{cases} n, & (\omega', l) \le 0; \\ 1/n, & (\omega', l) > 0; \end{cases}$$

n — коэффициент преломления воды относительно воздуха,

В случае, когда поверхность раздела двух сред является статической, $p(l) = \delta(l - k)$ и

$$\Phi(r_\perp, h, \omega)\Big|_{\omega\in\Omega} = R(\omega_{\text{отр}}, k)\Phi(r_\perp, h, \omega_{\text{отр}}) + \nu^2[1 - R(\omega_{\text{пр}}, k)]\Phi(r_\perp, h, \omega_{\text{пр}}),$$

где $k = (0, 0, 1)$.

Яркость $\Phi(r_\perp, z, \omega)$ является суперпозицией световой дымки системы океан-атмосфера $\Phi_0(z, \omega)$ и яркости отраженного плоскостью $z = 0$ излучения $\widetilde{\Phi}(r_\perp, z, \omega)$. На основе теории линейных систем [4] введем ОПФ системы океан-атмосфера $T(p, z, \omega)$ такую, что

$$\widehat{\widetilde{\Phi}}(r_\perp, h, \omega)\Big|_{\omega\in\Omega} = T(p, z, \omega)\Big|_{\omega\in\Omega}\widehat{\widetilde{\Phi}}(p, 0, \omega)\Big|_{\omega\in\Omega_-}, \tag{2}$$

где символом $\widehat{\ }$ обозначен Фурье-образ, $p = (p_x, p_y)$- вектор пространственных частот. Функция $T(p, z, \omega)$ определяется только оптическими и геометрическими свойствами системы океан-атмосфера и не зависит от условий освещения и пространственного спектра яркости $\widetilde{\Phi}(r_\perp, 0, \omega)_{\omega\in\Omega_-}$. Таким образом, функция $T(\cdot)$ позволяет для произвольного распределения яркости $\widetilde{\Phi}(\cdot)\big|_{\omega\in\Omega_-}$ вблизи плоскости $z = 0$ с помощью прямого и обратного преобразования Фурье определить пространственно-угловое распределение яркости $\widetilde{\Phi}(\cdot)$ на любой высоте над поверхностью океана. И наоборот, по измерениям поля яркости системы с помощью функции $T(\cdot)$ можно определить пространственно-угловое распределение яркости вблизи плоскости $z = 0$.

2 Численный алгоритм

В данной работе рассматривается дополнительная сопряженная задача, в которой источником света является светящаяся плоскость $z = 0$. Яркость последней обозначим через $J(r_\perp, \omega), \omega \in \Omega_-$. Световое поле в такой системе описывается упоминавшимся интегро-дифференциальным уравнением переноса с краевыми условиями

$$\widetilde{\Phi}(r_\perp, H, \omega)\Big|_{\omega\in\Omega_+} = 0, \quad \widetilde{\Phi}(r_\perp, 0, \omega)\Big|_{\omega\in\Omega_-} = J(r_\perp, \omega)\Big|_{\omega\in\Omega_-}$$

вместе с соотношением (1).

Один из способов оценки $T(\cdot)$, по существу, основан на алгоритме, изложенном в [6]. Там приведен способ оценки методом Монте-Карло пограничной кривой, описывающей световое поле в плоском слое мутной среды от источника с прямолинейным разрывом яркости. Алгоритм построен на сочетании методов "корелированной выборки" и "сопряженных траекторий". Не вдаваясь в общую схему построения траекторий фотонов и вычисления соответствующих оценок (см., например [3]) отметим лишь специфичные для рассматриваемой задачи особенности.

Выберем для определенности систему координат xyz так, чтобы граница разрыва яркости $J(\cdot)$ была ориентирована перпендикулярно оси OX, т.е.

$$J(r_\perp, \omega) = J(x, \omega) = \begin{cases} 0 & , x < 0; \\ J(\omega) & , x \geq 0. \end{cases}$$

В этом случае $\widetilde{\Phi}(r_\perp, z^*, \omega^*) = \widetilde{\Phi}(x, z^*, \omega^*)$. Зададим на оси OX сетку узлов $\{x_i, i = 1, \ldots, N\}$, в которых вычисляются значения пограничной кривой $\Phi_i = \widetilde{\Phi}(x_i, z^*, \omega^*)$ для направления ω^*. Согласно [6] траектории фотонов строятся из произвольной точки $P = (x_p, y_p, z*)$ плоскости $z = z^*$ в направлении $-\omega^*$ и заканчиваются либо вылетом в полупространство $z > H$, либо попаданием фотона на плоскость $z = 0$. В случае пересечения траектории фотона с плоскостью $z = h$ в направлении ω' вычисляется коэффициент $R(\omega', s)$, где s — случайное направление нормали к элементарной площадке с центром в точке пересечения, выбранное из распределения с плотностью $p(l)$. С вероятностью $R(\omega, s)$ происходит зеркальное отражение и фотон приобретает направление $\omega = \omega' - 2(\omega', s)s$. С вероятностью $1 - R(\omega', s)$ происходит преломление в направлении $\omega = \omega'/\nu - \gamma s$. Коэффициент $R(\omega, l)$ вычисляется по известной формуле Френеля (см., например, [7]).

Использование коррелированных траекторий и локальной оценки "по направлению" [3] позволяет по одним и тем же траекториям оценить функцию $\widetilde{\Phi}(x, z^*, \omega^*)$ для всего набора точек $\{x_i, i = 1, \ldots, N\}$ одновременно. Случайные оценки ξ_i вкладов в функционалы $\widetilde{\Phi}_i(\cdot)$ от столкновения в точке (x, y, z) с направлениями ω' и ω до и после столкновения с учетом (1) приобретают вид:

$$\xi_i = \begin{cases} \dfrac{\sigma_s(z)}{\sigma(z)} W \exp(-\tau)[1 - R(\omega, s)] n^2 J(\rho, -m) & , z \in (h, H]; \\ \dfrac{\sigma_s(z)}{\sigma(z)} W \exp(-\tau) J(\rho, -\omega) & , z \in (0, h]. \end{cases} \tag{3}$$

$$\rho = \begin{cases} x_i + x - x_p + (a, \omega) \dfrac{z - h}{(-\omega, k)} + (a, m) \dfrac{h}{(-m, k)} & , z \in (h, H]; \\ x_i + x - x_p + (a, \omega) \dfrac{z}{(-\omega, k)} & , z \in (0, h], \end{cases}$$

где $m = \omega/n - \gamma s$, $\gamma = [(\omega, s) + \sqrt{n^2 - 1}]/n$, $a = (1, 0, 0)$,

$$\tau = \begin{cases} \displaystyle\int_0^{t_1} \sigma[z + t(\omega, k)] dt + \int_0^{t_2} \sigma[z + t_1(\omega, k) + t(m, k)] dt, \\ \text{где} \quad t_1 = (z - h)/(-\omega, k), t_2 = h/(-m, k) \quad , z \in (h, H]; \\ \displaystyle\int_0^{t_1} \sigma[z + t(\omega, k)] dt, \quad \text{где} \quad t_1 = z/(-\omega, k) \quad , z \in (0, h], \end{cases}$$

при $(\omega, k) < 0$, при $(\omega, k) \geq 0$ значения $\xi_i = 0$.

В (3) W — "вес" фотона до столкновения (см. [6]).
В точке (x, y, z) пересечения траектории фотона с плоскостью $z = h$ в направлении ω имеем:

$$\xi_i = \begin{cases} W[1 - R(\omega, s)] \exp(-\tau) J(\rho, -m) n^2 & , m \quad \text{определяется так же как в (3)}, \omega \in \Omega_+; \\ W R(\omega, s)] \exp(-\tau) J(\rho, -m) & , m = \omega - 2(\omega, s)s, \omega \in \Omega_-. \end{cases} \tag{4}$$

где $\rho = x_i + x - x_p + (am)h/(-mk)$, а τ определяются также, как в (3) для $z \in (0, h]$.

По результатам оценки пограничной кривой $\widetilde{\Phi}(x, z, \omega)$ в узлах сетки $\{x_i, i = 1, \ldots, N\}$ нетрудно вычислить ОПФ системы. В нашем случае $p = (p_x, 0)$ и имеем [4, 8]:

$$T(p_x, z, \omega) = T_a(\cdot) \exp[i T_\Phi(\cdot)],$$

где $T_a(p_x, z, \omega) = [S^2(\cdot) + C^2(\cdot)]^{1/2}$ — амплитудно-частотная и $T_\Phi(p_x, z, \omega) = \arcsin[S(\cdot)/T_a(\cdot)]$ — фазочастотная характеристики системы, а $S(p_x, z, \omega)$ и $C(p_x, z, \omega)$ соответственно синус- и косинус- преобразования Фурье от функции

$$A(x, z, \omega) = \frac{\partial \widetilde{\Phi}(\cdot)}{\partial x} / [\widetilde{\Phi}(+\infty, z, \omega) - \widetilde{\Phi}(-\infty, z, \omega)].$$

Поскольку пограничная кривая вычисляется в конечном числе узлов, то точность вычисления функций $S(\cdot)$ и $C(\cdot)$ существенно зависит от выбора сетки $\{x_i\}$. Вблизи границы раздела яркости рекомендуется брать более мелкую сетку для более точного вычисления функции $A(x, z, \omega)$, поскольку она вычисляется методом аппроксимации. Кроме того, в данной работе при нахождении ОПФ используется дискретное прямое и обратное преобразование Фурье, результат которого является периодической функцией, период которой также напрямую зависит от выбора сетки и вычисляется по формуле $L = 2\pi/\Delta x$, где $\Delta x = x_{i+1} - x_i$. Поэтому необходимым условием для выбора сетки является то, что период полученной после преобразования Фурье функции включал в себя интересующий нас интервал области исследования оптической передаточной функции.

3 Результаты

Здесь приведен пример тестовых расчетов функций $T_a(\cdot)$ с использованием описанной методики. Расчеты выполнены для длины волны λ=0.5 мкм. Индикатриса рассеяния в атмосфере была принята рэлеевской ([10], стр. 18), в воде использовалась линейная интерполяция индикатрисы, заданной таблично по результатам наблюдений для чистой воды ([10], стр. 34).

Показатель преломления воды относительно воздуха был принят равным $n = 1.33$. Плотность распределения нормали к взволнованной поверхности определялась в соответствии с [9] двумерной гауссовской

плотностью распределения уклонов элементарных площадок с параметрами, зависящими от скорости ветра над поверхностью океана:

$$p(z_x, z_y) = \frac{1}{2\pi\alpha_x\alpha_y} \exp\left[-\left(\frac{z_x}{\alpha_x}\right)^2 - \left(\frac{z_y}{\alpha_y}\right)^2\right],$$

$$l_x = -z_x/\sqrt{z_x^2 + z_y^2 + 1},$$

$$l_y = -z_y/\sqrt{z_x^2 + z_y^2 + 1},$$

$$l_z = 1/\sqrt{z_x^2 + z_y^2 + 1},$$

$\alpha_x^2 = 0.001 + 3.16 \cdot 10^{-3} \cdot V$, $\alpha_y^2 = 0.003 + 1.85 \cdot 10^{-3} \cdot V$, V — скорость ветра и м/сек.

На рис. 2 представлены результаты оценок функции T_a, полученные для скоростей ветра 2, 6, 10, 14 м/с и двух направлений. Синей линией обозначены графики функций для направления ветра поперек границы раздела яркости на дне, красной линией — вдоль границы раздела яркости.

Для качественного сравнения двух алгоритмов, описанных в работе [5], были произведены численные эксперименты при различных показателях скорости ветра.

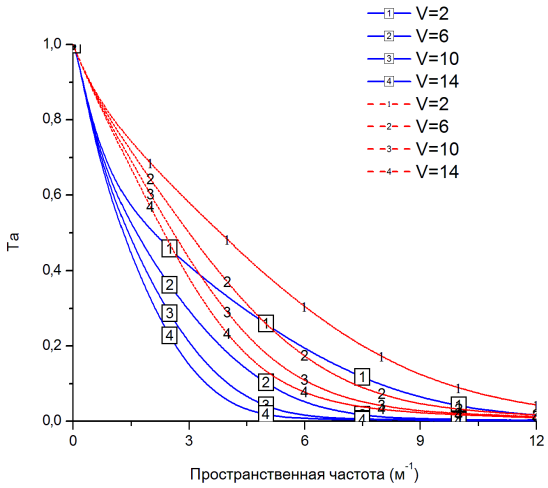


Рис. 2: Функция T_a для глубины 10 м в зависимости от скорости и направления ветра.

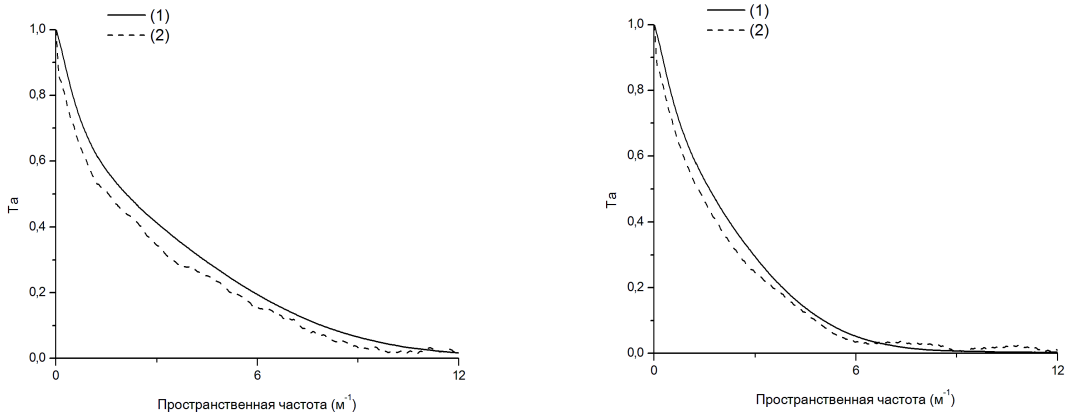


Рис. 3: Функция T_a для глубины 10 м, скорости ветра V=2 м/с (слева) и V=6 (справа), вычисленная с использованием двух различных алгоритмов.

На рис. 3 представлены результаты оценок функции T_a для скоростей ветра 2 м/с и 6 м/с, полученные с использованием двух разных алгоритмов: (1) — оценка, полученная с использованием формул (3) и (4); (2) — оценка, полученная с использованием алгоритма, исследованного ранее в работе [5]. В первом случае моделировалось 10^6 траекторий, а во втором — 10^7. Расчеты показывают, что первый алгоритм существенно превосходит по эффективности второй алгоритм, т.е. при одних тех же затратах времени вычислений статистическая погрешность расчетов первым алгоритмом значительно меньше статистической погрешности расчетов вторым алгоритмом.

Заключение

В работе получены результаты исследования поведения оптической передаточной функции для различных параметров окружающей среды. Результаты исследования показали, что ОПФ уменьшается с увеличением скорости ветра, а также очень сильно зависит от его направления. Также произведено сравнение двух возможных алгоритмов для оценки ОПФ. Результаты показали, что использование в объединении с остальными алгоритмами метода “сопряженных траекторий” при решении сопряженной задачи (в которой источником света считается морское дно) показывает более точную оценку ОПФ при одних и тех же затратах времени вычисления.

Список литературы

[1] Мулламаа Ю.-А.Р. Атлас оптических характеристик взволнованной поверхности моря. Институт физики и астрономии АН ЭССР. Тарту. 1964.Plass G.N., Kattawar G.W. Radiative Transfer in an Atmosphere-Ocean System // Appl. Opt. 1969. V. 8. № 2. P. 455–466.

[2] Plass G.N., Kattawar G.W. Radiative Transfer in an Atmosphere-Ocean System // Appl. Opt. 1969. V. 8, № 2. P. 455–466.

[3] Марчук Г. И., Михайлов Г. А., Назаралиев М. А. и др. Метод Монте-Карло в атмосферной оптике. // Новосибирск: Наука, 1976, с.280.

[4] Папулис А. Теория систем и преобразований в оптике. // М.: Мир, 1971, с. 495.

[5] Каргин Б.А., Ухинова О.С. Расчет методом Монте-Карло оптической передаточной функции системы океан-атмосфера // Труды международной конференции “Вычислительная математика и математическая геофизика”, посвященной 90-летию со дня рождения академика А. С. Алексеева, ИВМиМГ СО РАН, Новосибирск. 8–12 октября, 2018 г., с.189–195.

[6] Каргин Б. А., Кузнецов С. В., Михайлов Г. А. Оценка методом Монте-Карло функции передачи контраста яркости через светорассеивающую среду. // Изв. АН СССР, ФАО, 1979, т.15, №10, с.1027–1035.

[7] Борн М., Вольф Э., Основы оптики.// М.:Наука, 1973, с. 720.

[8] Мишин И. В., Сушкевич Т.А. Оптическая пространственно-частотная характеристика атмосферы и ее приложения. // Исследования Земли из космоса, 1980, №3. с. 58–68.

[9] Cox C., Munk W. H., The measurement of the roughness of the sea curface from photographs of the sun’s glitter. // J. Opt. Soc. America, 1954, 44, No. 11, p. 838–850.

[10] Кондратьев К.Я., Кантер Р.Р., Каргин Б.А., Поздняков Д.В. Численное моделирование в задачах оптического дистанционного зондирования внутренних водоемов
Л.: Наука, 1987, 62 с.

Ухинова Ольга Сергеевна — к.ф.-м.н., науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: olsu@osmf.sscc.ru.
Каргин Борис Александрович — д.ф.-м.н., зав. лабораторией Института вычислительной математики и математической геофизики СО РАН; Новосибирский государственный университет;
e-mail: bkargin@osmf.sscc.ru.
Дата поступления — 30 апреля 2019 г.

МЕТОД КОЛЛОКАЦИИ И НАИМЕНЬШИХ КВАДРАТОВ И ЕГО ПРИЛОЖЕНИЕ К РАСЧЕТУ НАПРЯЖЕННО-ДЕФОРМИРОВАННОГО СОСТОЯНИЯ ПЛАСТИН

В. П. Шапеев[1,2], С. К. Голушко[1,3], В. А. Беляев[2], Л. С. Брындин[1,2]

[1] *Новосибирский государственный университет, 630090, Новосибирск*

[2] *Институт теоретической и прикладной механики им. С. А. Христиановича СО РАН, 630090, Новосибирск*

[3] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*

УДК 519.635
DOI: 10.24411/9999-016A-2019-10086

Предложены и реализованы новые h-, hp- и p-варианты метода коллокации и наименьших квадратов для численного решения краевых задач для уравнений бигармонического типа в канонических и нерегулярных областях. Приведено сравнение полученных численных результатов с известными частными решениями других авторов, использовавших конечно-разностные и спектральные методы повышенного порядка аппроксимации. Дан анализ напряженно-деформированного состояния тонкой пластины на упругом основании, находящейся под действием постоянной поперечной нагрузки.

Ключевые слова: метод коллокации и наименьших квадратов, повышенный порядок аппроксимации, бигармоническое уравнение, нерегулярная область, пластина, напряженно-деформированное состояние.

Введение

Бигармонические уравнения $\Delta^2 u = f$, где Δ — оператор Лапласа, играют важную роль во многих областях науки и техники. К решению бигармонических уравнений сводится моделирование и анализ поведения многих задач механики сплошной среды: гидродинамики, линейной теории упругости, теории тонких пластин и др. Например, в гидродинамике при малых числах Рейнольдса функция тока удовлетворяет бигармоническому уравнению. В линейной теории упругости решение бигармонического уравнения используется для представления функции напряжений Эйри. В теории тонких пластин решение бигармонического уравнения используется для анализа напряженно-деформированного состояния (НДС) изотропных пластин, находящихся под действием поперечной нагрузки [1]. Однако численное решение бигармонических уравнений вызывает трудности из-за наличия производных четвертого порядка в дифференциальном уравнении.

Для численного решения краевых задач для бигармонического уравнения обычно используется метод конечных разностей (МКР), главными преимуществами которого являются простота построения расчетной сетки и скорость решения соответствующих задач линейной алгебры. При этом существует ограниченное количество публикаций, посвященных применению МКР для решения бигармонического уравнения в нерегулярных областях [2, 3]. Другим широко используемым методом для решения таких задач является метод конечных элементов (МКЭ) с использованием неструктурированных сеток [4–8]. Как отмечается в [9], несмотря на то, что техника решения уравнений с частными производными (УЧП) четвертого порядка с помощью МКЭ и МКР хорошо развита, существует не так много результатов, полученных для произвольных нерегулярных областей и различных граничных условий. В последнее время при решении задач механики сплошной среды все более популярными становятся спектральные методы и методы спектральных элементов, главным преимуществом которых является экспоненциальный характер уменьшения погрешности в

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 18-29-18029) и Российского научного фонда (код проекта 18-13-00392).

ISBN 978-5-901548-42-4

случае достаточно гладких решений [9–12]. Однако в случаях, когда решения не являются достаточно гладкими, спектральные методы становятся менее эффективными и менее гибкими, например, по сравнению с МКЭ. Существуют и другие методы решения бигармонического уравнения: метод быстрых мультиполей, метод сплайн-коллокации, миметический метод, смешанные методы. Подробную библиографию о них можно найти, например, в [13].

Настоящая статья является продолжением работ [14,15] и распространяется на случай применения h-, p- и hp-вариантов метода коллокации и наименьших квадратов (КНК) для решения уравнений бигармонического типа в канонических и нерегулярных областях. В методе КНК путем проектирования задачи для УЧП в конечномерное линейное функциональное пространство ставится в соответствие приближенная задача, решение которой сводится к решению системы линейных алгебраических уравнений (СЛАУ). Приближенное решение ищется в виде линейной комбинации с неопределенными коэффициентами базисных элементов, определенных в этом пространстве. В каждой ячейке сетки для нахождения неизвестных коэффициентов выписывается переопределенная СЛАУ, состоящая из уравнений коллокации, краевых условий и условий согласования.

1 Постановка задачи и описание метода КНК

1.1 Постановка задачи

Рассмотрим задачу Дирихле для бигармонического уравнения

$$\Delta^2 u = f(x_1, x_2),\ (x_1, x_2) \in \Omega,\ \ u|_{\delta\Omega} = g_1(x_1, x_2),\ \ u_n|_{\delta\Omega} = g_2(x_1, x_2) \tag{1}$$

в нерегулярной области $\Omega \subset \mathrm{R}^2$ с границей $\delta\Omega$, $u_n = \dfrac{\partial u}{\partial n}$, $\overrightarrow{n}$ внешняя единичная нормаль к границе области, $u(x_1, x_2)$ — искомая функция, $f(x_1, x_2)$, $g_1(x_1, x_2)$ и $g_2(x_1, x_2)$ — заданные функции. Реализация рассматриваемого метода решения других краевых задач для бигармонического уравнения осуществляется аналогично алгоритму, описанному ниже. Аналогично [15,16] в представленной работе внешняя граница $\delta\Omega$ области Ω задана дискретно. Восполним границу области по входным дискретным данным непрерывной векторной сплайн-функцией $(x_1(t), x_2(t))$ в случае, когда граница области является гладкой. Если же граница области имеет точки излома, то следует строить отдельные векторные сплайн-функции на участках, концами которых являются точки излома.

1.2 h-вариант метода КНК

Нерегулярная область включается в фиктивный прямоугольник, который покрывается регулярной сеткой размера $N_1 \times N_2$ с прямоугольными ячейками. На границе области появляется "одинарный" слой нерегулярных ячеек (н-ячеек), отсеченных границей от прямоугольных граничных ячеек начальной регулярной сетки. В данном варианте метода н-ячейка, которая не содержит центр исходной содержащей ее прямоугольной ячейки, присоединяется к соседней ячейке. Более подробное описание методики можно найти в работах [15,16].

Пусть N_{cells} — количество расчетных ячеек в каждом варианте метода КНК. Для удобства реализации метода КНК в каждой j-ой ячейке области введем локальные координаты

$$y_1 = \frac{(x_1 - x_{1j})}{h_1}, \qquad y_2 = \frac{(x_2 - x_{2j})}{h_2}, \tag{2}$$

где (x_{1j}, x_{2j}) — центр ячейки, h_1 и h_2 характерные линейные размеры j-ой ячейки, $j = 1, ..., N_{cells}$, $v(y_1, y_2) = u(x_1(y_1), x_2(y_2))$. Здесь $2h_1$ и $2h_2$ являются размерами прямоугольной ячейки в направлениях x_1 и x_2 соответственно. В каждой j-ой ячейке сетки приближенное решение v_{hj} задачи (1) ищем в виде линейной комбинации с неопределенными коэффициентами базисных элементов пространства полиномов некоторой степени относительно двух переменных. Здесь в качестве базисных элементов рассматриваются ортогональные полиномы Чебышева и мономы. В случае применения полиномов Чебышева представление приближенного решения имеет вид

$$v_{hj}(y_1, y_2) = \sum_{i_1=0}^{K_1-1} \sum_{i_2=0}^{K_2-1} c_{i_1 i_2, j} \phi_{i_1}(y_1) \phi_{i_2}(y_2), \quad (y_1, y_2) \in [-1, 1] \times [-1, 1], \tag{3}$$

где

$$\phi_{i_1}(y_1) = \cos(i_1 \arccos(y_1)), \quad \phi_{i_2}(y_2) = \cos(i_2 \arccos(y_2)).$$

Если в качестве базисных элементов взяты мономы, то приближенное решение имеет вид

$$v_{hj}(y_1, y_2) = \sum_{i_1=0}^{K} \sum_{i_2=0}^{K-i_1} c_{i_1 i_2, j} y_1^{i_1} y_2^{i_2}. \tag{4}$$

1.3 p- и hp-варианты метода КНК

В p-варианте исходная область также включается в фиктивный прямоугольник и в расчетной области строится единый кусок полиномиального решения. Аналогично по формуле (2) вводятся локальные координаты и $N_{cells} = N_1 = N_2 = j = 1$.

В hp-варианте можно добиться повышения точности решения за счет измельчения шагов сетки и/или за счет повышения степени полиномов базиса используемого пространства.

1.4 Уравнения приближенной задачи и их решение

Неизвестные коэффициенты $c_{i_1 i_2, j}$ в методе КНК в каждой ячейке определяются из переопределенной "локальной" системы уравнений, состоящей из уравнений коллокации, условий согласования (h- и hp-вариант) и краевых условий.

Уравнения коллокации, умноженные на $h_1^2 h_2^2$, выписываются в определенных точках коллокации (y_{1c}, y_{2c}) в каждой j-ой ячейке и имеют вид

$$k_c \left(\frac{h_2^2}{h_1^2} \frac{\partial^4 v_{hj}}{\partial y_1^4} + 2 \frac{\partial^4 v_{hj}}{\partial y_1^2 \partial y_2^2} + \frac{h_1^2}{h_2^2} \frac{\partial^4 v_{hj}}{\partial y_2^4} \right) = k_c h_1^2 h_2^2 f(x_1(y_{1c}), x_2(y_{2c})), \tag{5}$$

где k_c положительный весовой множитель.

Условия согласования (умноженные на $h_1 h_2$ в (7)) выписываются в определенных точках согласования в h- и hp-вариантах в каждой j-ой ячейке и имеют вид:

$$k_{m_0} v_{hj} + k_{m_1} \frac{\partial v_{hj}}{\partial n_j} = k_{m_0} \hat{v}_h + k_{m_1} \frac{\partial \hat{v}_h}{\partial n_j}, \tag{6}$$

$$h_1 h_2 \left(k_{m_2} \frac{\partial^2 v_{hj}}{\partial {n_j}^2} + k_{m_3} \frac{\partial^3 v_{hj}}{\partial {n_j}^3} \right) = h_1 h_2 \left(k_{m_2} \frac{\partial^2 \hat{v}_h}{\partial {n_j}^2} + k_{m_3} \frac{\partial^3 \hat{v}_h}{\partial {n_j}^3} \right), \tag{7}$$

где n_j — внешняя нормаль к границе j-ой ячейки, v_{hj} и $\hat{v}_h$ — пределы приближенного решения задачи при стремлении изнутри и извне к границе j-й ячейки соответственно, $k_{m_0}, k_{m_1}, k_{m_2}, k_{m_3}$ — положительные весовые множители.

Граничные условия для задачи Дирихле (1) выписываются в определенных точках $(y_{1b}, y_{2b}) \in \delta\Omega$ и имеют вид

$$k_{b_0} v_{hj} = k_{b_0} g_1(x_1(y_{1b}), x_2(y_{2b})), \tag{8}$$

$$k_{b_1} \left(\frac{n_1}{h_1} \frac{\partial v}{\partial y_1} + \frac{n_2}{h_2} \frac{\partial v}{\partial y_2} \right) = k_{b_1} g_2(x_1(y_{1b}), x_2(y_{2b})), \tag{9}$$

где k_{b_0}, k_{b_1} — положительные весовые множители.

Из-за ограничений на объем статьи здесь только отметим, что в предыдущих работах [14,15], посвященных решению бигармонического уравнения h-вариантами метода КНК было установлено, что для достижения хорошей обусловленности приближенной задачи количество уравнений должно превышать количество неизвестных приблизительно в три раза. Способ равномерной расстановки точек коллокации, согласования и точек записи краевых условий на внешней границе $\delta\Omega$ приведен, например, в [15] для h-варианта при $K = 4$ (4). В [17,18] описан способ расположения точек коллокации и согласования в случае применения ортогональных полиномов Чебышева (3).

Значения весовых множителей в методе КНК в некоторых пределах влияют на точность приближенного решения задачи, скорость сходимости итерационного процесса, обусловленность локальных СЛАУ, т.е. в данной работе применяется взвешенный метод КНК.

В h- и hp-вариантах переопределенная СЛАУ, полученная объединением уравнений (5)–(9), взятых из всех ячеек расчетной области (глобальная система), решается итерационно методом Гаусса-Зейделя. В нем одна "глобальная итерация" состоит из последовательного решения локальных СЛАУ в каждой ячейке области. Матрица локальной системы приводится к верхнетреугольному виду ортогональным методом Хаусхолдера (аналогично и в p-варианте) при построении решения в каждой ячейке. Итерационный процесс продолжается до тех пор, пока выполняется условие

$$\max_{i_1 i_2, j} |c^{n+1}_{i_1 i_2, j} - c^{n}_{i_1 i_2, j}| > \epsilon,$$

где $c^{n}_{i_1 i_2, j}$ — коэффициент полинома, аппроксимирующего решение в ячейке с номером j на n-й итерации. Величина ϵ — заданная малая константа [15, 16].

2 Численные эксперименты и результаты расчетов

2.1 Задача Дирихле

В численных экспериментах на сходимость приближенного решения на последовательности сеток при измельчении шагов сетки в h-варианте и увеличении степени полиномов в hp- и p-варианте рассматривалась задача (1) и другие краевые задачи для бигармонического уравнения. В представленных ниже таблицах приведены значения относительной погрешности приближенного решения в норме L_2 (величина $||E_r||_2$) и абсолютной погрешности приближенного решения в бесконечной норме (величина $||E_a||_\infty$). Порядок сходимости погрешности численного решения в случае h-варианта метода КНК определяется следующим образом

$$R = \log_2 \frac{E_p}{E_c},$$

где E_c — значение погрешности ($||E_r||_2$ или $||E_a||_\infty$) на сетке размера $N_1 \times N_2$, E_p — значение погрешности на сетке размера $N_1/2 \times N_2/2$.

Для значительного уменьшения времени, необходимого для проведения расчета, и количества итераций здесь было использовано комбинированное применение операции продолжения (prolongation) на многосеточном комплексе в методе Федоренко [19], диагонального предобуславливателя [20] и на всех сетках комплекса метода ускорения сходимости итерационного процесса, основанного на подпространствах Крылова [21].

Пример 1. Рассмотрим задачу Дирихле для бигармонического уравнения с тестовым решением $u(x_1, x_2) = x_1^3 \ln(1 + x_2) + \dfrac{x_2}{1 + x_1}$ в области $\dfrac{x_1^2}{0.5^2} + \dfrac{x_2^2}{0.15^2} \leqslant 1$. Возможности вариантов метода КНН показаны в табл. 1. В ней приведены значения абсолютной погрешности $||E_a||_\infty$ приближенного решения этой задачи на последовательности сеток, полученные авторами работ [2, 3]. Далее в табл. 1 приведены значения погрешности, полученные применением h-варианта метода КНК при $K = 4$ [15], далее — новые результаты при $K = 6$ и результаты, полученные применением hp-варианта при фиксированном значении $N_1 \times N_2$. Из анализа таблицы видно, что МКР, реализованный в [3], имеет больший порядок сходимости, чем h-вариант метода КНК при $K = 4$ [15]. Однако, при $K = 6$ вариант метода КНН не уступает указанному МКР по порядку сходимости и значительно превосходит его по точности. В [2,3] решение исходной краевой задачи для бигармонического уравнения (1) было сведено к последовательному решению двух задач Дирихле для уравнений Пуассона. Известно, что задача Дирихле для уравнения Пуассона является лучше обусловленной задачей, чем краевая задача для бигармонического уравнения (1). Поэтому такой подход позволяет относительно просто построить высокоточное решение исходной задачи. Здесь решение с повышенной точностью краевой задачи для бигармонического уравнения получено методом КНК без использования такого дополнительного приема. Результаты, полученные применением hp-варианта методом КНК и приведенные в конце табл. 1, значительно превосходят другие приведенные в ней результаты. Следует отметить также, что в hp-варианте малое значение погрешности $||E_a||_\infty$ достигается с меньшими вычислительными затратами по сравнению с h-вариантом.

Пример 2. Рассмотрим задачу Дирихле для бигармонического уравнения с тестовым решением $u(x_1, x_2) = \sin(2\pi x_1)\cosh(2x_2) - \cos(2\pi x_1)\sinh(2x_2)$ в невыпуклой полигональной области с вершинами: $(1; 1)$, $(-1; 1)$, $(-1; 0)$, $(0; 0)$, $(1; -1)$, $(1; 1)$. В табл. 2 приведены результаты численных экспериментов, полученных применением p-варианта с представлением решения в виде (3) и спектрального метода [11]. Во второй и третьей колонках приведены результаты численных экспериментов в случае, когда приближенное решение в обоих методах строилось во всей области одним полиномом, в пятой и шестой колонках — в трех ячейках.

Таблица 1: Результаты численных экспериментов (пример 1)

	Работа [2]			Работа [3]		Работа [15], $K=4$		$K=6$		$N_1\times N_2 = 4\times 2$	
$N_1\times N_2$	$\|\|E_a\|\|_\infty$	R	$N_1\times N_2$	$\|\|E_a\|\|_\infty$	R	$\|\|E_a\|\|_\infty$	R	$\|\|E_a\|\|_\infty$	R	K	$\|\|E_a\|\|_\infty$
64×32	3.65e-4	—	8×4	—	—	1.97e-5	—	1.34e-7	—	10	2.83e-9
128×64	9.54e-5	1.9	16×8	3.0e-6	—	1.56e-6	3.65	9.26e-9	3.85	12	1.42e-10
256×128	2.08e-5	2.2	32×16	1.1e-7	4.8	1.52e-7	3.35	1.24e-10	6.22	14	7.00e-12
512×256	4.98e-6	2.1	64×32	1.5e-9	6.2	1.24e-8	3.61	4.14e-12	4.90	16	6.56e-14

Таблица 2: Результаты численных экспериментов (пример 2)

	Работа [11]	p-вариант		Работа [11]	hp-вариант
$K_1 = K_2$	$\|\|E_a\|\|_\infty$	$\|\|E_a\|\|_\infty$	$K_1 = K_2$	$\|\|E_a\|\|_\infty$	$\|\|E_a\|\|_\infty$
15	9.00e-2	1.23e-2	12	6.00e-2	1.44e-3
18	9.00e-4	3.26e-5	15	8.00e-4	7.92e-5
22	3.10e-5	3.72e-8	18	2.00e-7	3.20e-9
25	9.10e-5	8.32e-10	21	7.00e-9	1.36e-10

Таблица 3: Результаты численных экспериментов (пример 3)

h-вариант, $K=4$				p-вариант	
$N_1\times N_2$	$\|\|E_r\|\|_2$	R	N_{iter}	$K_1 = K_2$	$\|\|E_r\|\|_2$
10×10	1.80e-1	—	41	20	2.63e-4
20×20	1.43e-2	3.65	90	30	1.02e-5
40×40	2.96e-3	2.27	131	40	1.32e-6
80×80	6.89e-4	2.10	231	50	3.02e-7

2.2 Пластина на упругом основании

Пусть поперечно нагруженная пластина покоится на упругом основании. Дифференциальное уравнение для прогиба в декартовой системе координат в модели Винклера [1] записывается в виде

$$\frac{\partial^4 w(x_1,x_2)}{\partial x_1^4} + 2\frac{\partial^4 w(x_1,x_2)}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 w(x_1,x_2)}{\partial x_2^4} + \frac{k w(x_1,x_2)}{D} = \frac{q(x_1,x_2)}{D},$$

где w — прогиб срединной поверхности пластины, q — поперечная нагрузка, $D = Eh^3/(12(1-\nu^2))$ — жесткость пластины при изгибе, h — толщина пластины, E, ν — модуль Юнга и коэффициент Пуассона изотропного материала пластины, k — коэффициент постели (модуль основания).

Пример 3. Рассмотрим шарнирно закрепленную ($w|_{\delta\Omega} = M_n|_{\delta\Omega} = 0$, где M_n — изгибающий момент [1]) квадратную пластину $[0,10]\times[0,10]$, находящуюся под действием постоянной нагрузки q (Па). В этом случае

$$w(x_1,x_2) = \frac{16q}{\pi^2} \sum_{m=1,3,5,\ldots}^{\infty} \sum_{n=1,3,5,\ldots}^{\infty} \frac{\sin\dfrac{m\pi x_1}{d_1} \sin\dfrac{n\pi x_2}{d_2}}{mn\left[\pi^4 D\left(\dfrac{m^2}{{d_1}^2} + \dfrac{n^2}{{d_2}^2}\right)^2 + k\right]}. \tag{10}$$

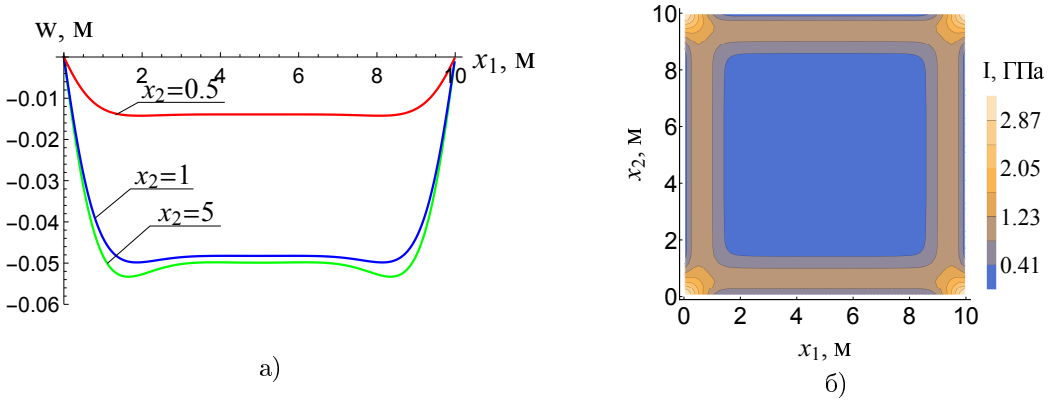
Для описания напряженного состояния пластины используем функцию распределения интенсивности напряжений

$$I = \frac{1}{\sqrt{2}}\sqrt{(\sigma_x - \sigma_y)^2 + \sigma_x^2 + \sigma_y^2 + 6{\sigma_{xy}}^2},$$

где $\sigma_x = -\overline{E}(w_{xx} + \nu w_{yy})$, $\sigma_y = -\overline{E}(w_{yy} + \nu w_{xx})$, $\sigma_{xy} = -\overline{E}(w_{xy})$, $\overline{E} = Ex_3/(1-\nu^2)$, $x_3 \in [-h/2, h/2]$.

Расчеты были проведены при следующих параметрах: $h = 0.1$ м, $E = 200$ ГПа, $\nu = 0.28$, $q = 15$ МПа, $k = 0.3$ ГПа/м, $x_3 = h/2$, $\epsilon = 10^{-10}$. Для оценки погрешности приближенного решения (см. табл. 3) в качестве точного решения взят начальный отрезок ряда (10) так, чтобы соответствующий остаток ряда по модулю гарантированно не превосходил 10^{-8}. В табл. 3 N_{iter} обозначает количество итераций. На рис. 1 показана форма прогиба шарнирно закрепленной пластины в разных сечениях $x_2 = \text{const}$ и интенсивность

напряжений. Здесь для наглядности масштабы вдоль горизонтальной оси и вдоль вертикальной оси отличаются примерно в сто раз, так что величина прогиба пластины меньше ее продольного размера более, чем в сто раз.



а) б)

Рис. 1: Величина прогиба w деформированной пластины в разных сечениях $x_2 = \text{const}$ ($а$) и интенсивность напряжений I ($б$)

Заключение

Предложены и реализованы новые h-, hp- и p-варианты метода КНК повышенной точности для численного решения краевых задач для уравнений бигармонического типа в канонических и нерегулярных областях. Показано, что при решении методом КНК бигармонического уравнения в нерегулярных областях достигается точность не ниже точности решений, полученных МКР и спектральным методом. С другой стороны, показанная здесь возможность применения метода КНК при моделировании и анализе НДС пластин на упругом основании имеет самостоятельное прикладное значение.

Список литературы

[1] Тимошенко С.П., Войновский-Кригер С. Пластины и оболочки. М.: Физматгиз, 1963.

[2] Chen G., Li Z., Lin P. A Fast Finite Difference Method for Biharmonic Equations on Irregular Domains and its Application to an Incompressible Stokes Flow // Advances in Computational Mathematics. 2007. V. 29, N 2. P. 113–133. doi:10.1007/s10444-007-9043-6.

[3] Ben-Artzi M., Chorev I., Croisille J.P., Fishelov D. A Compact Difference Scheme for the Biharmonic Equation in Planar Irregular Domains // SIAM J. Numer. Anal. 2009. V. 47, N 4. P. 3087–3108. doi: 10.1137/080718784.

[4] Brenner S. C. An Optimal-Order Nonconforming Multigrid Method for the Biharmonic Equation // SIAM J. Numer. Anal. 1989. V. 26, N 5. P. 1124–1138. doi: 10.1137/0726062.

[5] Mayo A., Greenbaum A. Fast Parallel Iterative Solution of Poisson's and the Biharmonic Equations on Irregular Regions // SIAM J. Sci. Comput. 1992. V. 13. N 1. P. 101–118. doi: 10.1137/0913006.

[6] Hanisch M.R. Multigrid Preconditioning for the Biharmonic Dirichlet Problem // SIAM J. Numer. Anal. 1993. V. 30. No 1. P. 184–214. doi: 10.1137/0730009.

[7] Davini C., Pitacco I. An Unconstrained Mixed Method for the Biharmonic Problem // SIAM J. Numer. Anal. 2000. V. 38. N 3. P. 820–836. doi: 10.1137/S0036142999359773.

[8] Eymard R., Gallouët T., Herbin R., Linke A. Finite Volume Schemes for the Biharmonic Problem on General Meshes // Math. Comp. 2012. V. 812. N 280. P. 2019–2048. doi: 10.1090/S0025-5718-2012-02608-1.

[9] Shao W., Wu X., Chen S. Chebyshev tau Meshless Method Based on the Integration-Differentiation for Biharmonic-Type Equations on Irregular Domain // Engin. Anal. Bound. Elem. 2012. V. 36. N 12. P. 1787–1798. doi: 10.1016/j.enganabound.2012.06.005.

[10] Mai-Duy N., See H., Tran-Cong T. A Spectral Collocation Technique Based on Integrated Chebyshev Polynomials for Biharmonic Problems in Irregular Domains // Appl. Math. Model. 2009. V. 33. N 1. P. 284–299. doi: 10.1016/j.apm.2007.11.002.

[11] Shao W., Wu X. An Effective Chebyshev tau Meshless Domain Decomposition Method Based on the Integration-Differentiation for Solving Fourth Order Equations // Appl. Math. Modell. 2015. V. 39. N 9. P. 2554–2569. doi: 10.1016/j.apm.2014.10.048.

[12] Zhuang Q., Chen L. Legendre-Galerkin Spectral-Element Method for the Biharmonic Equations and its Applications // Comp. & Math. Appl. 2017. V. 74. N 12. P. 2958–2968. doi: 10.1016/j.camwa.2017.07.039.

[13] Jiang Y., Wang B., Yuesheng X. A Fast Fourier-Galerkin Method Solving a Boundary Integral Equation for the Biharmonic Equation // SIAM J. Numer. Anal. 2000. V. 52, N 5. P. 2530–2554. doi: 10.1137/140955744.

[14] Shapeev V.P., Belyaev V.A., Golushko S.K., Idimeshev S.V. New Possibilities and Applications of the Least Squares Collocation Method // EPJ Web of Conferences. 2018. V. 173. P. 01012-1 — 01012-8. doi: 10.1051/epjconf/201817301012.

[15] Шапеев В.П., Беляев В.А. Решение с повышенной точностью бигармонического уравнения в нерегулярных областях методом коллокации и наименьших квадратов // Выч. мет. программирование. 2018. Т. 19. № 4. С. 340–355. doi: 10.26089/NumMet.v19r431.

[16] Беляев В.А., Шапеев В.П. Решение задачи Дирихле для уравнения Пуассона методом коллокации и наименьших квадратов в области с дискретно заданной границей // Вычисл. технологии. 2018. Т. 23. № 3. С. 15–30.

[17] Голушко С.К., Идимешев С.В., Шапеев В.П. Разработка и применение метода коллокаций и наименьших невязок к задачам механики анизотропных слоистых пластин // Вычисл. технологии. 2014. Т. 19. № 5. С. 24–36.

[18] Голушко С.К., Идимешев С.В., Семисалов Б.В. Методы решения краевых задач механики композитных пластин и оболочек // Уч. пособие по курсу «Прямые и обратные задачи механики композитов». Новосибирск: КТИ ВТ СО РАН, 2014. 131 с.

[19] *Федоренко Р.П.* О скорости сходимости одного итерационного процесса // Ж. вычисл. матем. и матем. физ. 1964. Т. 4. № 3. С. 559–564.

[20] Ramšak M., Škerget L. A Subdomain Boundary Element Method for High-Reynolds Laminar Flow Using Stream Function–Vorticity Formulation // Int. J. Numer. Meth. Fluids. 2004. V. 46. N 8. P. 815–847. doi: 10.1002/fld.776.

[21] Saad Y. Numerical Methods for Large Eigenvalue Problems. Manchester: Manchester University Press, 1992. 346 p.

Шапеев Василий Павлович — д.ф.-м.н., глав. науч. сотр. Института теоретической и прикладной механики им. С. А. Христиановича СО РАН; профессор Новосибирского государственного университета; e-mail: vshapeev@gmail.com;

Голушко Сергей Кузьмич — д.ф.-м.н., глав. науч. сотр., Института вычислительных технологий СО РАН; профессор Новосибирского государственного университета; e-mail: s.k.golushko@gmail.com;

Беляев Василий Алексеевич — аспирант, мл. науч. сотр. Института теоретической и прикладной механики им. С. А. Христиановича СО РАН; e-mail: belyaevasily@mail.ru;

Брындин Лука Сергеевич — магистрант Новосибирского государственного университета; лаборант Института теоретической и прикладной механики им. С. А. Христиановича СО РАН; e-mail: bryndin-1996@mail.ru.

Дата поступления — 24 мая 2019 г.

ОЦЕНКА ВАРИАБЕЛЬНОСТИ РЕШЕНИЯ, ПОЛУЧЕННОГО С ПОМОЩЬЮ ИНТЕРВАЛЬНОЙ РЕГУЛЯРИЗАЦИИ

С. П. Шарый[1,2]

[1] *Институт вычислительных технологий СО РАН, 630090, Новосибирск*
[2] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.9
DOI: 10.24411/9999-016A-2019-10087

В работе представлена одна из возможных конструкций количественной меры вариабельности псевдорешения неточно заданной системы линейных алгебраических уравнений, которое получено с помощью так называемой интервальной регуляризации. Даётся вывод новой меры вариабельности и её неформальная трактовка.
Ключевые слова: неточная система линейных уравнений, интервальная регуляризация, псевдорешение, мера вариабельности.

Введение

Цель этой работы — представить одну из возможных конструкций количественной меры вариабельности для решения системы линейных алгебраических уравнений, которое получено с помощью так называемой *интервальной регуляризации* (см. [1, 2]). Напомним, что это способ регуляризации плохообусловленных и неточно заданных систем линейных алгебраически хуравнений, основанный на методах интервального анализа и развивающий приёмы регуляризации по М.М. Лаврентьеву и регуляризации сдвигом.

Термином «вариабельность» мы называем степень изменчивости и неоднозначности решения, и необходимость её введения диктуется тем обстоятельством, что в задачах обработки интервальных данных ответ, как правило, неединствен. Обычно мы получаем целое множество различных оценок, одинаково пригодных в качестве ответов к задаче и согласующихся с её данными. То, насколько мало или обширно это множество, характеризуется термином «вариабельность». Иногда в подобных случаях говорят также о «разбросе решений».

Ситуация здесь аналогична задачам со случайными ошибками в данных, в которых решения, как известно, сами являются случайными величинами, а мерой их чувствительности и вариабельности может служить дисперсия оценки. Она выражает как меру рассеяния значений оценки, так и меру чувствительности этой оценки к изменениям входных данных задачи.

В отношении вариабельности псевдорешения при традиционном задачии погрешностей ответ на этот вопрос кажется достаточно очевидным: ею может стать любая величина, характеризующая размеры множества решений задачи, если оно непусто. Можно даже просто брать интервальные оценки множества решений, получаемые теми или иными методами интервального анализа. Определённый недостаток этого варианта (хорошо заметный в сравнении с аналогичным понятием дисперсии оценки для задач со случайными данными) — излишняя детализация ответа, выдаваемого в виде бруса в $\mathbb{R}^n$ или интервального вектора какого-нибудь другого вида, большое количество информации, которую ещё необходимо «переварить» и привести к компактной и выразительной форме. Другой недостаток — относительная сложность нахождения такой оценки.

Требуется относительно несложная и эффективно вычислимая величина, выражаемая одним числом, которая давала бы общее агрегированное представление об интересующем нас предмете. Аналогично дисперсии она может служить «прикидочной» характеристикой качества решений при практическом применении интервальной регуляризации.

В работе используется система обозначений, введённая неформальным международным стандартом [3], когда интервалы и интервальные объекты выделяются жирным шрифтом.

ISBN 978-5-901548-42-4

1 Формулировка основного результата

Пусть дана система линейных алгебраических уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad\quad \vdots \qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{1}$$

или, кратко,

$$Ax = b \tag{2}$$

с $m\times n$-матрицей $A = (a_{ij})$ и m-вектором правой части b, причём $b \neq 0$. При интервальной регуляризации [1,2] мы интервализуем систему (1)–(2), «раздувая» элементы матрицы a_{ij} и компоненты вектора правой части b_i, и переходим к интервальной системе линейных алгебраических уравнений (ИСЛАУ)

$$\begin{cases} \boldsymbol{a}_{11}x_1 + \boldsymbol{a}_{12}x_2 + \ldots + \boldsymbol{a}_{1n}x_n = \boldsymbol{b}_1, \\ \boldsymbol{a}_{21}x_1 + \boldsymbol{a}_{22}x_2 + \ldots + \boldsymbol{a}_{2n}x_n = \boldsymbol{b}_2, \\ \quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad\quad \vdots \qquad\quad \vdots \\ \boldsymbol{a}_{m1}x_1 + \boldsymbol{a}_{m2}x_2 + \ldots + \boldsymbol{a}_{mn}x_n = \boldsymbol{b}_m, \end{cases} \tag{3}$$

или, кратко,

$$\boldsymbol{A}x = \boldsymbol{b} \tag{4}$$

с $m\times n$-матрицей $\boldsymbol{A} = (\boldsymbol{a}_{ij})$ и m-вектором правой части $\boldsymbol{b}$, такими что $\boldsymbol{A} \ni A$ и $\boldsymbol{b} \ni b$. Далее рассматривается *допусковое множество решений* построенной ИСЛАУ, т. е. множество

$$\Xi_{tol}(\boldsymbol{A},\boldsymbol{b}) = \big\{\, x \in \mathbb{R}^n \mid\ Ax \in \boldsymbol{b} \text{ для любых матриц } A \in \boldsymbol{A} \,\big\},$$

которое является наиболее устойчивым к возмущениям в матрице (см. подробности в [1,2,4]). Псевдорешением исходной системы уравнений (1)–(2) берётся точка максимума специального «распознающего функционала», который даёт количественную «меру согласования» вектора с интервальной системой (3)–(4) относительно допускового множества решений. Распознающий функционал обычно обозначается символом «Tol», и он имеет вид

$$\operatorname{Tol}(x,\boldsymbol{A},\boldsymbol{b}) \;=\; \min_{1\le i\le m} \left\{ \operatorname{rad}\boldsymbol{b}_i - \left| \operatorname{mid}\boldsymbol{b}_i - \sum_{j=1}^{n} \boldsymbol{a}_{ij}x_j \right| \right\}, \tag{5}$$

где

$$\operatorname{rad}\boldsymbol{b}_i = \tfrac{1}{2}(\overline{\boldsymbol{b}}_i - \underline{\boldsymbol{b}}_i) \qquad \text{и} \qquad \operatorname{mid}\boldsymbol{b}_i = \tfrac{1}{2}(\overline{\boldsymbol{b}}_i + \underline{\boldsymbol{b}}_i)$$

— соответственно радиусы и середины компонент вектора правой части $\boldsymbol{b}$, а модуль интервала — это максимум модулей его точек (или концов). Таким образом, решение исходной СЛАУ сводится к нахождению безусловного максимума по всем x из $\mathbb{R}^n$ —

$$\operatorname{Tol}(x,\boldsymbol{A},\boldsymbol{b}) \to \max,$$

и аргумент $\hat{x} = \arg\max_{x\in\mathbb{R}^n} \operatorname{Tol}(x,\boldsymbol{A},\boldsymbol{b})$ этого максимума принимается за псевдорешение системы (1)–(2).

В качестве величины, характеризующей вариабельность вектора псевдорешения $\hat{x} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$ системы линейных уравнений (1)–(2), мы предлагаем

$$\operatorname{IVE}(\boldsymbol{A},\boldsymbol{b}) \;=\; \sqrt{n}\, \max_{\mathbb{R}^n} \operatorname{Tol} \cdot \Big(\min_{A\in\boldsymbol{A}} \operatorname{cond}_2 A \Big) \cdot \frac{\big\| \arg\max\limits_{\mathbb{R}^n} \operatorname{Tol} \big\|_2}{\|b\|_2}. \tag{6}$$

В этой формуле
- n — размерность вектора неизвестных решаемой СЛАУ,
- $\|\cdot\|_2$ — евклидова норма векторов (2-норма),
- $\operatorname{cond}_2(A)$ — спектральное число обусловленности матрицы A, определяемое как отношение её наибольшего и наименьшего сингулярных чисел [5,6].

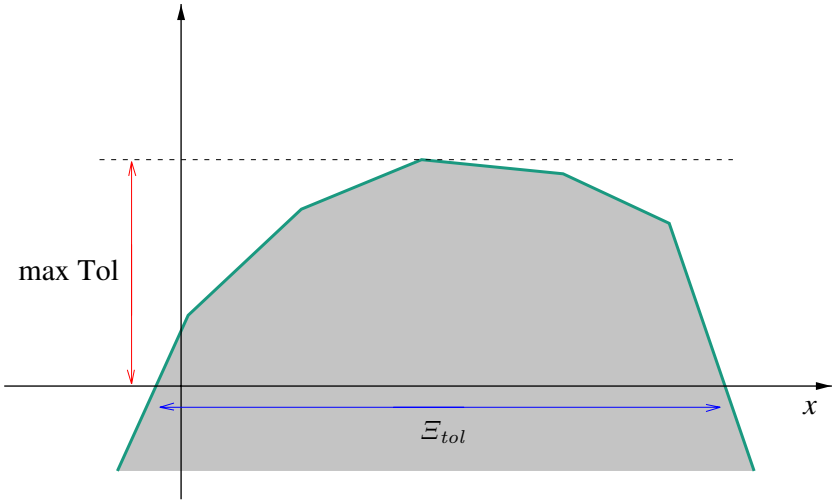


Рис. 1: Величина максимума распознающего функционала даёт представление о размерах допускового множества решений

Сам символ IVE образован как аббревиатура английской фразы «interval variability of the estimate». Ниже мы строго покажем, что величина IVE адекватно характеризует размеры непустого допускового множества решений. Но этот факт может быть также понят на основе наглядных образных соображений.

Допусковое множество решений интервальной системы линейных алгебраических уравнений является, как известно, множеством нулевого уровня распознающего функционала Tol [4], или, другими словами, пересечением подграфика этого функционала с плоскостью $\mathrm{Tol} = 0$ (см. Рис. 1):

$$\Xi_{tol}(\boldsymbol{A}, \boldsymbol{b}) = \left\{ x \in \mathbb{R}^n \mid \mathrm{Tol}\,(x, \boldsymbol{A}, \boldsymbol{b}) \geq 0 \right\}.$$

Как следствие, величина максимума распознающего функционала может служить, при прочих равных условиях, мерилом того, насколько обширно или узко допусковое множество решений. Чем больше $\max \mathrm{Tol}$, тем больше размеры допускового множества решений, и наоборот. Другие факторы, обеспечивающие «прочие равные условия» — это наклон гиперплоскостей, из кусков которых составлен полиэдральный график функционала Tol (прямых в одномерном случае на Рис. 1). Наклон гиперплоскостей определяется коэффициентами задающих их уравнений, которые являются концами интервалов в системе (3)–(4). Величину этого наклона обобщённым образом характеризует число обусловленности матрицы данных. Наконец, множитель

$$\frac{\| \arg\max \mathrm{Tol} \|_2}{\|b\|_2} = \frac{\|\hat{x}\|_2}{\|b\|_2}$$

— это масштабирующий коэффициент, с помощью которого обеспечивается соизмеримость окончательной величины псевдорешению и вектору правой части системы уравнений. Так и получается формула (6).

2 Обоснование меры вариабельности

Будем считать, что количество уравнений m может не совпадать с числом неизвестных переменных n, причём $m \geq n$. Тогда системы линейных алгебраических уравнений (1)–(2) и (3)–(4) являются либо квадратными, либо переопределёнными, т. е. имеющими $m \times n$-матрицы с $m \geq n$.

2.1 Оценка возмущений решения прямоугольных линейных систем

Отправной точкой наших конструкций, обосновывающих выбор меры вариабельности именно в виде величины (6), будет известное неравенство, оценивающее возмущение Δx ненулевого решения системы линейных алгебраических уравнений $Ax = b$ в зависимости от изменений Δb в векторе правой части (см., к примеру, [5–7]):

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \mathrm{cond}_2(A) \cdot \frac{\|\Delta b\|_2}{\|b\|_2}. \tag{7}$$

В случае евклидовой нормы векторов и спектрального числа обусловленности матриц это неравенство справедливо и при $m \geq n$, но лишь если матрица системы имеет полный ранг.

2.2 Интервальные системы с точечной матрицей

Рассмотрим интервальную систему линейных алгебраических уравнений

$$Ax = \boldsymbol{b} \tag{8}$$

с точечной (неинтервальной) $m \times n$-матрицей A и интервальным m-вектором $\boldsymbol{b}$ в правой части, причём $m \geq n$. Предположим также, что её допусковое множество решений непусто, т. е. $\Xi_{tol}(A, \boldsymbol{b}) = \left\{ x \in \mathbb{R}^n \mid Ax \in \boldsymbol{b} \right\} \neq \varnothing$. Как можно быстро и несложно оценить размеры этого множества решений? Форма ответа на этот вопрос может быть различной, и мы будем строить его в виде оценки «типа радиуса» для множества решений. Более точно, мы получим оценку евклидовой нормы $\max \|x' - \hat{x}\|_2$ по всем $x' \in \Xi_{tol}(A, \boldsymbol{b})$ для некоторой специальной фиксированной точки $\hat{x} \in \Xi_{tol}(A, \boldsymbol{b})$, в качестве которой берётся

$$\hat{x} \;=\; \arg\max_{x \in \mathbb{R}^n} \operatorname{Tol}(x, A, \boldsymbol{b})$$

— аргумент безусловного максимума распознающего функционала для системы (8), т. е. вектор псевдорешения исходной СЛАУ. Строго говоря, эта точка может определяться неединственным образом, но тогда $\hat{x}$ может быть любой из точек, на которых достигается рассматриваемый максимум.

Пусть x' — какая-то точка из допускового множества решений $\Xi_{tol}(A, \boldsymbol{b})$. Как оценить $\|x' - \hat{x}\|_2$? Ясно, что x' и $\hat{x}$ являются решениями систем линейных алгебраических уравнений с матрицей A и некоторыми правыми частями b' и $\hat{b}$, соответственно, из интервального вектора $\boldsymbol{b}$. Если $\hat{x} \neq 0$ и $\hat{b} \neq 0$, то можно применить неравенство (7), рассмотрев возмущение решения $\hat{x}$ системы линейных алгебраических уравнений $Ax = \hat{b}$. Тогда $\Delta x = x' - \hat{x}$, $\Delta b = b' - \hat{b}$, и мы имеем

$$\frac{\|x' - \hat{x}\|_2}{\|\hat{x}\|_2} \;\leq\; \operatorname{cond}_2 A \cdot \frac{\|b' - \hat{b}\|_2}{\|\hat{b}\|_2},$$

откуда получается абсолютная оценка

$$\|x' - \hat{x}\|_2 \;\leq\; \operatorname{cond}_2 A \cdot \|\hat{x}\|_2 \cdot \frac{\|b' - \hat{b}\|_2}{\|\hat{b}\|_2}. \tag{9}$$

Точка $\hat{x}$ находится в результате максимизации распознающего функционала Tol, для вычисления обусловленности $\operatorname{cond}_2 A$ существуют хорошо разработанные стандартные процедуры, и потому для практической работы с неравенством (9) нужно как-то оценить значения $\|b' - \hat{b}\|_2$ и $\|\hat{b}\|_2$.

Наиболее просто мы поступим с $\|\hat{b}\|_2$, приближённо взяв $\|\hat{b}\|_2 \approx \|b\|_2$, т. е. как норму правой части b исходной СЛАУ. Ясно, что это «наиболее представительная» точка из интервального вектора $\boldsymbol{b}$, который получен «раздуванием» b. Естественно, что при этом может допускаться некоторое огрубление, так что теперь вместо (9) мы будем писать

$$\|x' - \hat{x}\|_2 \;\lessapprox\; \operatorname{cond}_2 A \cdot \|\hat{x}\|_2 \cdot \frac{\|\Delta b\|_2}{\|b\|_2}. \tag{10}$$

Теперь необходимо определить приращение правой части $\Delta b = b' - \hat{b}$. Его оценкой сверху является вектор радиусов $\operatorname{rad} \boldsymbol{b}$ правой части, но она слишком груба. Чтобы получить более тонкую оценку Δb, наряду с системой (8) рассмотрим также систему линейных алгебраических уравнений

$$Ax = \tilde{\boldsymbol{b}}, \tag{11}$$

у которой правая часть $\tilde{\boldsymbol{b}}$ получена «сжатием» интервального вектора $\boldsymbol{b}$, т. е. как

$$\tilde{\boldsymbol{b}} := \left[\underline{\boldsymbol{b}} + M, \overline{\boldsymbol{b}} - M \right], \tag{12}$$

где величина «сжатия» M определяется следующим образом:

$$M := \max_{x \in \mathbb{R}^n} \operatorname{Tol}(x, A, \boldsymbol{b}) \;\geq\; 0.$$

Поскольку этот максимум M достигается функционалом Tol при конкретном значении аргумента $\hat{x}$, то

$$M = \operatorname{Tol}(\hat{x}, A, \boldsymbol{b}) = \min_{1\le i\le m}\left\{ \operatorname{rad}\boldsymbol{b}_i - \left| \operatorname{mid}\boldsymbol{b}_i - \sum_{j=1}^{n} \boldsymbol{a}_{ij}\hat{x}_j \right| \right\} \le \min_{1\le i\le m} \operatorname{rad}\boldsymbol{b}_i.$$

Как следствие, $\underline{\boldsymbol{b}}+M \le \overline{\boldsymbol{b}}-M$ и интервальный вектор (12) — правильный, т. е. его концы не «перехлёстывают» друг за друга.

Но из свойств распознающего функционала следует, что для интервальной системы линейных алгебраических уравнений (11) максимум распознающего функционала равен нулю, т. е. $\max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, A, \tilde{\boldsymbol{b}}) = 0$. В самом деле, величины $\operatorname{rad}\boldsymbol{b}_i$ входят слагаемыми во все выражения, стоящие в (5) под знаком минимума по i. Поэтому если одновременно увеличить или уменьшить все $\operatorname{rad}\boldsymbol{b}_i$ на одну и ту же величину при неизменных серединах $\operatorname{mid}\boldsymbol{b}_i$ компонент правой части, то и общее значение распознающего функционала увеличится или уменьшится на ту же самую величину. Иными словами, если взять константу $C \ge 0$ и интервальный m-вектор

$$\boldsymbol{e} = \big([-1,1],\ldots,[-1,1]\big)^{\top},$$

то для системы $Ax = \boldsymbol{b} + C\boldsymbol{e}$, в которой все компоненты правой части одновременно расширены на $[-C, C]$, имеем

$$\operatorname{Tol}(x, A, \boldsymbol{b} + C\boldsymbol{e}) = \operatorname{Tol}(x, A, \boldsymbol{b}) + C. \tag{13}$$

Значит,

$$\max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, A, \boldsymbol{b} + C\boldsymbol{e}) = \max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, A, \boldsymbol{b}) + C \tag{14}$$

(и это также верно для систем (3)–(4) с произвольными матрицами $\boldsymbol{A}$, а не только точечными). Совершенно аналогичным образом на допусковое множество решений и распознающий функционал действует и равномерное сужение вектора правой части. Если мы сузили все компоненты на одну и ту же величину M, то и максимум распознающего функционала новой интервальной системы уменьшится на M.

Итак, в силу свойств распознающего функционала допусковое множество решений $\Xi_{tol}(A, \tilde{\boldsymbol{b}})$ для системы (11) имеет пустую внутренность (такие множества часто называют «нетелесными»), и мы будем условно считать этот факт равносильным свойству «иметь нулевые размеры». Естественно, что это упрощающее допущение, так как в действительности допусковое множество решений при нулевом максимуме распознающего функционала может быть отнюдь не одноточечным. Тем не менее, мы принимаем это упрощение, в пользу которого говорит также то обстоятельство, что ситуация с нулевым максимумом распознающего функционала неустойчива: соответствующее допусковое множество решений может стать пустым при сколь угодно малом возмущении данных.

Ещё один факт, который касается вспомогательной системы (11) с суженной правой частью и следует из (13)–(14), состоит в том, что определённая ранее точка $\hat{x}$ является аргументом максимума распознающего функционала также для системы (11):

$$\hat{x} = \arg\max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, A, \tilde{\boldsymbol{b}})$$

По этой причине точка $\hat{b} = A\hat{x}$ лежит в $\tilde{\boldsymbol{b}}$.

Из сказанного вытекает, что множество решений интервальной линейной системы $Ax = \boldsymbol{b}$ получается из множества решений системы $Ax = \tilde{\boldsymbol{b}}$, которое имеет «пренебрежимые размеры» и для которого $\max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, \boldsymbol{A}, \tilde{\boldsymbol{b}}) = 0$, с помощью расширения вектора $\tilde{\boldsymbol{b}}$ в каждой компоненте одновременно на $[-M, M]$, где

$$M = \max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, \boldsymbol{A}, \boldsymbol{b}).$$

Как следствие, в неравенстве (10) мы можем взять Δb таким, что

$$|\Delta b_i| = \left| b'_i - \hat{b}_i \right| \le M, \qquad i = 1, 2, \ldots, m, \tag{15}$$

причём оценка сверху может достигаться, и потому

$$\|\Delta b\|_\infty = \max_{x\in\mathbb{R}^n} \operatorname{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) \tag{16}$$

в чебышёвской векторной норме $\|y\|_\infty := \max_{1\le i\le m} |y_i|$.

Строго говоря, интервальный вектор $\tilde{\boldsymbol{b}} \ni \hat{b}$ может иметь ненулевые размеры, и тогда $\|\Delta b\|_\infty$ как будто способна превзойти M. Но если мы хотим сделать неравенство (10) наиболее точным и достижимым, то его правая часть должна зануляться при $\max \mathrm{Tol} = 0$, и это предопределяет выбор (15)–(16).

Чтобы подставить найденное значение $\|\Delta b\|_\infty$ в (10), где фигурируют евклидовы нормы, воспользуемся неравенством эквивалентности векторных норм. Как известно, для любого вектора y (см. [5])

$$\|y\|_\infty \le \|y\|_2 \le \sqrt{n}\, \|y\|_\infty. \tag{17}$$

В целом получаем

$$\|x' - \hat{x}\|_2 \;\lessapprox\; \sqrt{n}\ \mathrm{cond}_2 A \cdot \Big\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,\Big\|_2 \cdot \frac{\max \mathrm{Tol}}{\|b\|_2}. \tag{18}$$

2.3 Общие интервальные системы

Рассмотрим, наконец, общие интервальные системы линейных алгебраических уравнений $\boldsymbol{A}x = \boldsymbol{b}$, с интервальной матрицей коэффициентов. Пусть, как и ранее,

$$\hat{x} \;=\; \arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,(x, \boldsymbol{A}, \boldsymbol{b})$$

— аргумент безусловного максимума распознающего функционала для системы $\boldsymbol{A}x = \boldsymbol{b}$. Пусть также x' — какая-то точка из непустого допускового множества решений $\Xi_{tol}(\boldsymbol{A}, \boldsymbol{b})$. Как оценить $\|x' - \hat{x}\|_2$?

В силу известных свойств допускового множества решений (см., к примеру, [4]) его можно представить как

$$\Xi_{tol}(\boldsymbol{A}, \boldsymbol{b}) \;=\; \bigcap_{A\in\boldsymbol{A}} \{\, x \in \mathbb{R}^n \mid Ax \in \boldsymbol{b} \,\} \;=\; \bigcap_{A\in\boldsymbol{A}} \Xi_{tol}(A, \boldsymbol{b}),$$

т. е. как пересечение множеств решений отдельных ИСЛАУ, имеющих точечные матрицы A из $\boldsymbol{A}$. Для каждой интервальной линейной системы $Ax = \boldsymbol{b}$ с $A \in \boldsymbol{A}$ точки x' и $\hat{x}$ лежат во множестве решений $\Xi_{tol}(A, \boldsymbol{b})$, и мы имеем оценку (18). Следовательно, для допускового множества решений системы $\boldsymbol{A}x = \boldsymbol{b}$, которое является пересечением множеств решений отдельных систем $Ax = \boldsymbol{b}$ с $A \in \boldsymbol{A}$, должно быть справедливо

$$\|x' - \hat{x}\|_2 \;\lessapprox\; \min_{A\in\boldsymbol{A}} \left\{ \sqrt{n}\ \mathrm{cond}_2 A \cdot \Big\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,\Big\|_2 \cdot \frac{\max \mathrm{Tol}}{\|b\|_2} \right\}.$$

Правую часть в полученном неравенстве можно оценить приближённо, если внести минимум по $A \in \boldsymbol{A}$ под фигурные скобки, приближённо оценив отдельные множители, — число обусловленности $\mathrm{cond}\, A$, норму аргумента максимума $\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,\|$ и максимум распознающего функционала $\max \mathrm{Tol}$.

Полагая

$$\min_{A\in\boldsymbol{A}} \Big\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,(x, A, \boldsymbol{b})\Big\| \;\approx\; \Big\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,(x, \boldsymbol{A}, \boldsymbol{b})\Big\|,$$

$$\min_{A\in\boldsymbol{A}} \max_{x\in\mathbb{R}^n} \mathrm{Tol}\,(x, A, \boldsymbol{b}) \;\approx\; \max_{x\in\mathbb{R}^n} \mathrm{Tol}\,(x, \boldsymbol{A}, \boldsymbol{b}),$$

будем иметь

$$\|x' - \hat{x}\|_2 \;\lessapprox\; \sqrt{n}\ \min_{A\in\boldsymbol{A}} \mathrm{cond}_2 A \cdot \Big\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,\Big\| \cdot \frac{\max \mathrm{Tol}}{\|b\|}. \tag{19}$$

Эта же оценка в силу неравенства эквивалентности норм (17) верна и для чебышёвской нормы:

$$\max_{x'\in\Xi_{tol}(\boldsymbol{A},\boldsymbol{b})} \|x' - \hat{x}\|_\infty \;\lessapprox\; \sqrt{n}\,\Big(\min_{A\in\boldsymbol{A}} \mathrm{cond}_2(A)\Big) \cdot \Big\|\arg\max_{x\in\mathbb{R}^n} \mathrm{Tol}\,\Big\|_2 \cdot \frac{\max_{x\in\mathbb{R}^n} \mathrm{Tol}}{\|b\|_2},$$

что оправдывает оценку IVE (6).

3 Обсуждение

В целом предложенная в этой работе мера вариабельности IVE находится довольно просто, так как после вычисления псевдорешения мы уже знаем $\max \mathrm{Tol}$ и $\arg\max \mathrm{Tol}$. Наиболее сложным при вычислении IVE является нахождение минимума чисел обусловленности точечных матриц из заданной интервальной матрицы. Основываясь на квазивогнутости спектрального числа обусловленности, можно организовать перебор угловых точечных матриц, что в общем случае, конечно же, непрактично и неэффективно.

Если интервальная матрица «достаточно узка» и не сильно отличается от точечной матрицы, то приближённо можно положить $\min_{A\in\boldsymbol{A}} \mathrm{cond}_2(A) \approx \mathrm{cond}_2(\mathrm{mid}\,\boldsymbol{A})$. Но в общем случае этот рецепт может работать плохо, так как левая и правая части выписанного приближённого равенства будут отличаться весьма сильно.

Для более ответственных задач можно организовать поиск минимума числа обусловленности каким-либо методом прямого поиска по множеству всех вершин интервальной матрицы. Нахождение градиента целевой функции в этой задаче затруднено её сложным характером, но вот вычисление значений целевой функции выполняется достаточно эффективно. Как следствие, наиболее подходящими инструментами решения этой задачи будут методы нулевого порядка, которые используют только значения целевой функции, т. е. числа обусловленности матрицы.

Можно предложить, к примеру, следующий эвристический алгоритм оценивания $\min \mathrm{cond}_2(A)$. Задавшись некоторым натуральным числом N, организуем N шагов случайного блуждания по множеству угловых матриц интервальной матрицы $\boldsymbol{A}$. Будем вычислять число обусловленности полученной случайной угловой матрицы $\check{A}$ и её противоположной относительно диагонали бруса $\boldsymbol{A}$, следуя идее так называемых диагональных методов глобальной оптимизации [8]. Иными словами, мы вычисляем $\mathrm{cond}_2(\check{A})$ и ещё число обусловленности матрицы, у которой элементами взяты противоположные к $\check{A}$ концы интервалов $\boldsymbol{A}$, что делается для усиления «репрезентативности» нашего случайного блуждания. Наименьшее из значений, полученных после таких N шагов, берётся в качестве оценки для $\min \mathrm{cond}_2(A)$. Ясно, что конкретное значение числа шагов N нужно брать в зависимости от желаемой точности оценивания $\min \mathrm{cond}_2(A)$ и тех ресурсов, которые мы готовы на него потратить. Если нужно получить оценку для $\min \mathrm{cond}_2(A)$ как можно быстрее, то и N может быть невелико. Кроме того, при небольших N возможно также применение к результату какой-то корректирующей процедуры (например, умножение на поправочный коэффициент из открытого интервала $]0,1[$).

Список литературы

[1] Шарый С.П. Интервальная регуляризация для решения систем линейных алгебраических уравнений // Труды Международной конференции «Марчуковские научные чтения – 2017». ИВМиМГ СО РАН, Новосибирск, Россия, 25 июня – 14 июля 2017 г., Новосибирск. С. 975–982.

[2] Shary, S.P. Interval regularization for imprecise linear algebraic equations. Deposited in arXiv.org on 27 Sep 2018, No. arXiv:1810.01481. 21 p.

[3] Kearfott, R.B., Nakao, M., Neumaier, A., Rump, S., Shary, S.P., van Hentenryck, P. Standardized notation in interval analysis // Вычислительные Технологии. 2010. Т. 15, №1. С. 7–13.

[4] Шарый С.П. Конечномерный интервальный анализ. Новосибирск: ИВТ СО РАН & XYZ, 2019. 630 с. Электронная версия книги доступна на `http://www.nsc.ru/interval/?page=Library/InteBooks`

[5] Голуб Дж., ван Лоун Ч. Матричные вычисления. М.: Мир, 1999. 548 с.

[6] Уоткинс Д. Основы матричных вычислений. М.: «Бином. Лаборатория знаний», 2009. 664 с.

[7] Годунов С.К. Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997. 388 с.

[8] Сергеев Я.Д., Квасов Д.Е. Диагональные методы глобальной оптимизации. М.: ФИЗМАТЛИТ, 2008. 352 с.

Шарый Сергей Петрович — д.ф.-м.н., вед. науч. сотр. Института вычислительных технологий СО РАН;
e-mail: shary@ict.nsc.ru.
Дата поступления — 1 июня 2019 г.

FORMAL SPECIFICATION AND VERIFICATION OF THE STANDARD MATHEMATICAL FUNCTIONS

N. V. Shilov[1], I. S. Anureev[2], E. V. Bodin[2], D. A. Kondratiev[2], A. V. Promsky[2], S. O. Shilova, B. L. Faifel[3]

[1] *Innopolis University, 420500, Innopolis, Russia*
[2] *A.P. Ershov Institute of Informatics Systems, 630090, Novosibirsk, Russia*
[3] *Yu. Gagarin State Technical University of Saratov, 410054, Saratov, Russia*

UDC 004.052
DOI: 10.24411/9999-016A-2019-10088

Research project "Platform-independent approach to formal specification and verification of standard mathematical functions" is aimed onto a development of an incremental combined approach to the specification and verification of the standard mathematical functions like `sqrt`, `cos`, `sin`, etc. Platform-independence means that we attempt to design a relatively simple axiomatization of the computer arithmetic in terms of real, rational, and integer arithmetic (i.e. the fields $\mathbb{R}$ and $\mathbb{Q}$ of real and rational numbers, the ring $\mathbb{Z}$ of integers) but dont specify neither base of the computer arithmetic, nor a format of numbers representation. Incrementality means that we start with the most straightforward specification of the simplest easy to verify algorithm in real numbers and finish with a realistic specification and a verification of an algorithm in computer arithmetic. We call our approach combined because we start with a manual (pen-and-paper) verification of some selected algorithm in real numbers, then use these algorithm and verification as a draft and proof-outlines for the algorithm in computer arithmetic and its manual verification, and finish with a computer-aided validation of our manual proofs with some proof-assistant system (to avoid appeals to "obviousness" that are very common in human-carried proofs). In the paper we present first steps towards a platform-independent incremental combined approach to specification and verification of the standard functions `cos` and `sin` that implement mathematical trigonometric functions cos and sin.

Keywords: fix-point numbers, floating-point numbers, computer/machine arithmetic, formal verification, partial and total correctness, Hoare triples, Floyd verification method of inductive assertions, irrational numbers, periodic real functions, Taylor expansion/series.

Introduction

One who has a look at verification research and practice may observe that there exist *verification in large (scale)* and *verification in small (scale)*: verification in large deals (usually) behavioral properties of large-scale complex critical systems like the *Curiosity* Mars mission [11], while verification in small addresses (usually) functional properties of small programs like computing the standard trigonometry functions [9]. Verification of behavioral properties of a safety/mission/avalability-critical system doesn't guaranty safety/liveness/fairness of the system but may detect some bugs that may cause a very expensive and/or fatal system failure (like launch failure from launch-site "Vostochny" November 28, 2017, [16]). At the same time verification in small also is of the high importance: a *tiny* bug/mistake/error in a *small* but *frequently/massively* used function/program may cause a huge money losses; it is true in particular for the standard computer functions (available in the standard libraries) [8]. Of course this division of the verification research onto two streams — in large and in small scale — is just a split not a break because all verification research work altogether towards incorporation of the formal verification into the software development cycle — at compilation/linking stages maybe [10].

Our paper deals with *verification in small*, in particular, it looks like that it is about the same topic as [9] i.e. formal verification of two standard computer functions `cos` and `sin` that implement well-known trigonometry

The work has been supported by the Russian foundation of basic research (no. 17-01-00789 *Platform-independent approach to formal specification and verification of standard mathematical functions*).

ISBN 978-5-901548-42-4

mathematical real functions $\cos, \sin : \mathbb{R} \to \mathbb{R}$. But there are serious differences between [9] and our paper. Firstly, the cited paper is platform-dependent (Intel®IA-64 architecture), its approach is neither incremental nor combined; next it provides neither definition of the both cos and sin functions, nor specification of their computer partners `cos` and `sin`; finally, because of use of HOL-light, all algorithms in the cited paper are functional but not imperative. In contrast, in our paper we present platform-independent and incremental approach, based on provided formal definition for mathematical functions, discuss several variants of formal specifications for their computer partners, use Hoare total correctness assertions [1, 7] for logical specification of imperative algorithms that implements the computer functions, and finish with manual (pen-and-paper) verification (using Floyd-Hoare approach [1, 7]) of the computer functions for argument value in the rage $[-1, 1]$ (in radian measure). (Thus we postpone computer-aide validation of our proof for the future while the paper [9] have done computer-aided formal verification.)

Our present paper is a next one in a series of our papers devoted to the development of a platform-independent incremental combined approach to specification and verification of the standard mathematical functions [17–20]. Position papers [17, 18] have stated our concern regarding a need of

- better specification and incremental combined platform-independent verification of standard functions,
- introduction and standardization of a certification process for the standard functions,
- inclusion of an incremental combined platform-independent verification into this certification.

A work-in-progress electronic preprint [19] has presented a human-oriented specification and pen-and-paper verification of a computer square root function that implements Newton-Raphson method by *non-adaptive* `for`-loop (with a pre-computed number of iterations) and uses a look-up table for initial approximations. The specification in [19] has been presented as a total correctness assertion with use of precise arithmetic and the mathematical square root $\sqrt{\dots}$, algorithms has been presented by imperative pseudo-code with explicit distinction between precise and machine arithmetic, manual verification has been done in Floyd-Hoare style and adjustment (matching) of runs of algorithms with precise arithmetics and with machine arithmetics. It is possible to say that the primary contribution of the paper [19] was an axiomatisation of properties of a machine (fix-point as well as floating-point) arithmetic that are sufficient to carry out the verification.

A journal (Russian) paper [20] is based on an improved axiomatization from [19]. In the cited paper an *adaptive* imperative algorithm implementing the same Newton-Raphson method for a square root function has been specified by total correctness assertions and verified manually using Floyd-Hoare approach in both fix-point and floating-point arithmetics; the post-condition of the total correctness assertion states that the final overall error is not greater that $2ulp$ where ulp is *Unit in the Last Place* — the unit of the last meaningful digit. The paper [20] has reported two steps towards computer-aided validation and verification of the used adaptive algorithm:

- an implementation of a fix-point data type according to the axiomatization can be found at `https://bitbucket.org/ainoneko/lib_verify/src/`,
- ACL2 proofs of
  - the consistency of the computer arithmetics axiomatization,
  - the existence of a look-up table with initial approximations for $\sqrt{\dots}$

  can be found at `https://github.com/apple2-66/c-light/tree/master/experiments/square-root`.

1 Computing trigonometric functions in Fix-point Arithmetic

First we axiomatized a platform-independent fix-point arithmetic in the electronic preprint [19] and then improved the initial axiomatization in the journal paper [20]. In the present paper we follow the later version, but explicitly admit that there may be several different fix-point data types simultaneously.

A fix-point data-type (with Gaussian rounding) $\mathbb{D}$ satisfies the following axioms.

- The set of values $Val_{\mathbb{D}}$ is a finite set of rational numbers $\mathbb{Q}$ (and reals $\mathbb{R}$) such that
  - it contains the least $\inf_{\mathbb{D}} < 0$ and the largest $\sup_{\mathbb{D}} > 0$ elements,
  - altogether with

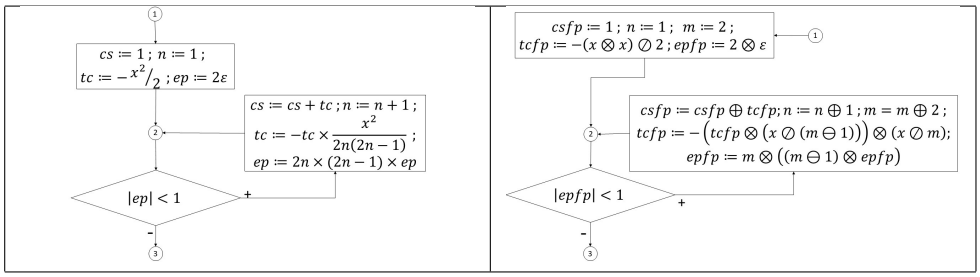


Figure 1: Flowcharts of the algorithms *CosCodeInZerOne* (left) and *CosCodeInFixPoint* (right)

- all rational numbers in $[\inf_{\mathbb{D}}, \sup_{\mathbb{D}}]$ with a step $\delta_{\mathbb{D}} > 0$,
- all integers $Int_{\mathbb{D}}$ in the range $[-\inf_{\mathbb{D}}, \sup_{\mathbb{D}}]$.

- Admissible operations include machine addition $\oplus$, subtraction $\ominus$, multiplication $\otimes$, division $\oslash$, integer rounding up $\lceil\ \rceil$ and down $\lfloor\ \rfloor$.

  Machine addition and subtraction. If the exact result of the standard mathematical addition (subtraction) of two fix-point values falls within the interval $[\inf_{\mathbb{D}}, \sup_{\mathbb{D}}]$, then machine addition (subtraction respectively) of these arguments equals to the result of the mathematical operation (and notation $+$ and $-$ is used in this case).

  Machine multiplication and division. These operations return values that are nearest in $Val_{\mathbb{D}}$ to the exact result of the corresponding standard mathematical operation: for any $x, y \in Val_{\mathbb{D}}$
  - if $x \times y \in Val_{\mathbb{D}}$ then $x \otimes y = x \times y$;
  - if $x/y \in Val_{\mathbb{D}}$ then $x \oslash y = x/y$;
  - if $x \times y \in [\inf_{\mathbb{D}}, \sup_{\mathbb{D}}]$ then $|x \otimes y - x \times y| \le \delta_{\mathbb{D}}/2$;
  - if $x/y \in [\inf_{\mathbb{D}}, \sup_{\mathbb{D}}]$ then $|x \oslash y - x/y| \le \delta_{\mathbb{D}}/2$;

  Integer rounding up and down are defined for all values in $Val_{\mathbb{D}}$.

- Admissible binary relations include all standard equalities and inequalities (within $[\inf_{\mathbb{D}}, \sup_{\mathbb{D}}]$) denoted in the standard way $=, \neq, \le, \ge, <, >$.

We study the trigonometric functions for argument values in the range $[-1, 1]$. The algorithm to compute cosine in this range is presented in the left part of the Fig. 1; the corresponding specification follows:

$$[(0 < \varepsilon < 1)\ \&\ (-1 \le x \le 1)]\ \ CosCodeInZerOne\ \ [|cs - \cos x| \le \varepsilon]\,. \tag{1}$$

Correctness of the specification (1) is easy to prove using the following invariant:

$$\begin{cases} 0 < \varepsilon < 1 \\ -1 \le x \le 1 \\ sign = (-1)^n \\ ep = (-1)^n \times (2n)! \times \varepsilon \\ tc = \frac{x^{2n}}{(2n)!} \\ cs = \sum_{m=0}^{m=(n-1)} (-1)^m \frac{x^{2m}}{(2m)!} \end{cases} \tag{2}$$

The above algorithm *CosCodeInZerOne* can be converted into algorithm *CosCodeInFixPoint* (with fix-point arithmetic) presented in the right part of the Fig. 1; the corresponding specification should be modified as follows:

$$\begin{bmatrix} \delta_{\mathbb{D}} < 1\ \& \\ 0 < \varepsilon < 1\ \&\ \varepsilon \in Val_{\mathbb{D}}\ \& \\ \exists N \in Int_{\mathbb{D}} :\ ((2N)! \times \varepsilon \ge 1)\ \& \\ -1 \le x \le 1\ \&\ x \in Val_{\mathbb{D}} \end{bmatrix} \\ CosCodeInFixPoint \\ \begin{bmatrix} |csfp - \cos x| \le \left(\varepsilon + \frac{3n\delta_{\mathbb{D}}}{2(1-\delta_{\mathbb{D}})}\right)\ \& \\ n = \min\{N\ :\ ((2N)! \times \varepsilon \ge 1)\} \end{bmatrix} \tag{3}$$

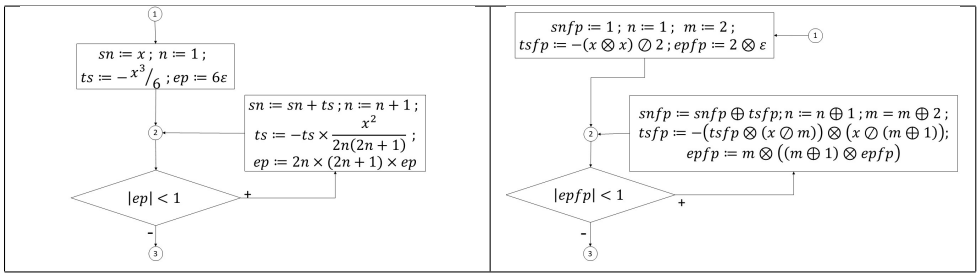


Figure 2: Flowcharts of the algorithms *SinCodeInZerOne* (left) and *SinCodeInFixPoint* (right)

```
(defun my-cos (x &optional (eps 1E-8))
  (let ((a 1) (s 1) (k 0))
    (loop
      (when (<= (abs a) eps) (return s))
      (setq a (- (/ (* a x x) (+ k 1) (+ k 2)))
                s (+ s a)
                k (+ k 2)))))
```

Figure 3: Lisp-function to compute approximations for cos in unbounded rational arithmetic with accuracy 10^{-8}

The algorithm that computes sine values for arguments in the range $[-1, 1]$ is presented in the left part of the Fig. 2; the corresponding specification follows:

$$[(0 < \varepsilon < 1)\ \&\ (-1 \le x \le 1)]\ \ SinCodeInZerOne\ \ [|sn - \sin x| \le \varepsilon] \tag{4}$$

Again, similarly to the above, the algorithm *SinCodeInZerOne* can be converted into algorithm *SinCodeInFixPoint* (with fix-point arithmetic) presented in the right part of the Fig. 2; the corresponding specification follows:

$$\left[\begin{array}{l} \delta_{\mathbb{D}} < 1\ \& \\ 0 < \varepsilon < 1\ \&\ \varepsilon \in Val_{\mathbb{D}}\ \& \\ \exists N \in Int_{\mathbb{D}} :\ ((2N+1)! \times \varepsilon \ge 1)\ \& \\ -1 \le x \le 1\ \&\ x \in Val_{\mathbb{D}} \end{array}\right] \quad SinCodeInFixPoint \quad \left[\begin{array}{l} |snfp - \sin x| \le \left(\varepsilon + \frac{3n\delta_{\mathbb{D}}}{2(1-\delta_{\mathbb{D}})}\right)\ \& \\ n = \min\{N\ :\ ((2N+1)! \times \varepsilon \ge 1)\} \end{array}\right] \tag{5}$$

Conclusion

In this paper we concentrate on design, specification and (a preliminary manual) verification of two trigonometric functions cos and sin in platform-independent fix-point arithmetic for small argument values in the range $[-1, 1]$ and use Taylor expansions as the definitions of the functions. Let us enumerate below some problems that need further theoretical and experimental research.

First, we should try to implement our verified algorithms on the virtual computer (for our fix-point arithmetic) available at `https://bitbucket.org/ainoneko/lib_verify/src/` and then test these implementations against

- selected algebraic values for these functions (for example, $\sin\frac{\pi}{6} = \frac{1}{2}$, $\sin\frac{\pi}{4} = \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$, $\cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$, etc.) in lines with test approach suggested and explained in [12,13];
- automatically generated test data computed using Taylor expansions in any language that supports unbounded integer arithmetic; for example, a Lisp-function in Fig. 3 computes approximations for cos in unbounded rational arithmetics with accuracy 10^{-8}.

Next we should complete a pen-and-paper proof of the specification (3) and a proof of the specification (5). Then we should validate/implement both proofs using some proof-assistance since manual proofs accompanied by

Table 1: Result of evaluation of `(my-cos 50)`

nominator	denominator
24370613165454113267560338608221954	252554104938731843322256481149588 16
98255428138520309455467035800407203	946608988211936130235611855567635 90
92481216493267961919792183534114282	788966318443878980153006888502210 53
43256901695353743984506265611950655	371046957284699682594602061094908 15
23779221083103374016633819981723287	736175504358202660509266505949702 81
8060581913569126766599	3572299506856327202849

computer-aided proofs is the core idea of the combined approach to verification. Currently in our studies of the square-root function [19, 20]) we are using proof-assistance ACL2 for proof-validation/implementation, but may change our choice later.

Finally we should move from computation, specification and verification of approximations of the trigonometric functions for small argument values in fix-point arithmetic to relatively big argument values in floating-point arithmetics. Computing of the values of the trigonometric functions for big argument values may be reduced to small argument values either using periodicity, or (for example) Chebychev polynomials, the trigonometric addition, the double-angle, and the half-angle formulas. (Remark that in the first case we need to compute approximate values of the constant π with high precision.)

We would like to finish the paper with a remark that the test-based approach from [12, 13] may be used for argument range larger than $[-1, 1]$; automated testing against valid approximations computed using unbounded rational arithmetic also may help; for example table 1 presents rational approximation of $\cos 50$ computed as `(my-cos 50)` using unbounded rational arithmetic (this rational value is "equal" to a float-point value 0.9649660286).

Please refer our pre-print [21] for further details.

References

[1] Apt K.R., de Boer F.S., Olderog E.-R. Verification of Sequential and Concurrent Programs. Springer-Verlag, 2009.

[2] C reference. `https://en.cppreference.com/w/c`. (Visited December 27, 2018.)

[3] C reference: cos, cosf, cosl. `https://en.cppreference.com/w/c/numeric/math/cos`. (Visited December 27, 2018.)

[4] C reference: Fundamental types. `https://en.cppreference.com/w/cpp/language/types`. (Visited December 27, 2018.)

[5] C reference: sin, sinf, sinl. `https://en.cppreference.com/w/c/numeric/math/sin`. (Visited December 27, 2018.)

[6] Dorn W.S., McCracken D.D. Numerical Methods with Fortran IV Case Studies. John Wiley & Sons, 1972.

[7] Gries D. The Science of Programming. Springer-Verlag, 1981.

[8] Grohoski G. Verifying Oracle's SPARC Processors with ACL2. Slides of the Invited talk for 14th International Workshop on the ACL2 Theorem Prover and Its Applications. `http://www.cs.utexas.edu/users/moore/acl2/workshop-2017/slides-accepted/grohoski-ACL2_talk.pdf`. (Visited December 27, 2018.)

[9] Harrison J. Formal Verification of Floating Point Trigonometric Functions. Lecture Notes in Computer Science. 2000. Vol. 1954. P.217–233.

[10] Hoare C.A.R. The Verifying Compiler: A Grand Challenge for Computing Research. Lecture Notes in Computer Science. 2003. Vol. 2890. P. 1–12.

[11] Holzmann G.J. Mars Code. Commun. ACM, 2014, Vol. 57(2), p. 64–73.

[12] Kuliamin V. Standardization and Testing of Mathematical Functions. Programming and Computer Software. 2007. Vol. 33(3). P. 154–173.

[13] Kuliamin V.V. Standardization and Testing of Mathematical Functions in floating point numbers. Lecture Notes in Computer Science. 2010. Vol. 5947. P. 257–268.

[14] Pi(number). Encyclopedia of Mathematics. http://www.encyclopediaofmath.org/index.php?title=Pi(number)&oldid=43586. (Visited December 27, 2018.)

[15] Radian. Encyclopedia of Mathematics. http://www.encyclopediaofmath.org/index.php?title=Radian&oldid=31576. (Visited December 27, 2018.)

[16] "Roskosmos" nazval prichinu neudachnogo zapuska s kosmodroma Vostochnyy (Roskosmos named the cause of the launch failure from launch-site "Vostochny"). https://www.rbc.ru/politics/12/12/2017/5a2ebcd59a79479d29667115. (In Russian. Visited December 27, 2018.)

[17] Shilov N.V. On the need to specify and verify standard functions. The Bulletin of the Novosibirsk Computing Center (Series: Computer Science), 2015, Vol. 38, p. 105–119.

[18] Shilov N.V., Promsky A.V. On specification andd verification of standard mathematical functions. Humanities and Science University Journal, 2016, Vol. 19, p. 57–68.

[19] Shilov N.V. Anureev I.S. Berdyshev M., Kondratyev D., Promsky A.V. Towards platform-independent verification of the standard mathematical functions: the square root function. https://arxiv.org/abs/1801.00969 [arXiv:abs/1801.00969]. (Visited December 27, 2018.)

[20] Shilov N. V., Kondratyev D. A., Anureev I. S., Bodin E. V, Promsky A. V. Platform-independent Specification and Verification of the Standard Mathematical Square Root Function. Modeling and Analysis of Information Systems, 2018, Vol. 25(6), p. 637–666.

[21] Shilov N.V., Faifel B.L., Shilova S.O., Promsky A.V. Towards platform-independent verification of the standard trigonometry functions. https://arxiv.org/abs/1901.03414 [arXiv:abs/1901.03414]. (Visited April 27, 2019.)

[22] Trigonometric functions. V.I. Bityutskov (originator), Encyclopedia of Mathematics. http://www.encyclopediaofmath.org/index.php?title=Trigonometric_functions&oldid=14919. (Visited December 27, 2018.)

[23] Trigonometric functions: Computation. From Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Trigonometric_functions#Computation. (Visited December 27, 2018.)

[24] Weisstein E. W. Trigonometric Addition Formulas. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Double-AngleFormulas.html.

[25] Weisstein E. W. Double-Angle Formulas. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/TrigonometricAdditionFormulas.html.

[26] Weisstein E. W. Half-Angle Formulas. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Half-AngleFormulas.html.

[27] Weisstein E. W. Leibniz Criterion. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/LeibnizCriterion.html.

Shilov Nikolay Vyacheslavovich — PhD, Assistant Professor of Innopolis University;
e-mail: shiloviis@mail.ru;
Anureev Igor Sergeevich — PhD., Senior Researcher of A.P. Ershov Institute of Informatics Systems SB RAS
e-mail: anureev@iis.nsk.su;
Bodin Evgeniy Viktorovich — Researcher of A.P. Ershov Institute of Informatics Systems SB RAS
e-mail: bodin@iis.nsk.su;
Kondratiev D. A. — PhD student of A.P. Ershov Institute of Informatics Systems SB RAS;
e-mail: apple-66@mail.ru;
Promsky Aleksey Vladimirovich — PhD., Science Secretary of A.P. Ershov Institute of Informatics Systems SB RAS;
e-mail: promsky@iis.nsk.su;
Shilova Svetlana Olegovna — Retired from A.P. Ershov Institute of Informatics Systems SB RAS;
e-mail: shilov61@inbox.ru;

Faifel Boris Leonidovich — PhD., Assistant Professor of Yu. Gagarin State Technical University of Saratov;
e-mail: catstail@yandex.ru.
Received — June 1, 2019

ЧИСЛЕННЫЙ АНАЛИЗ СТОХАСТИЧЕСКОГО ОСЦИЛЛЯТОРА ВАН-ДЕР-ПОЛЯ

М. А. Якунин

Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск

УДК 519.676
DOI: 10.24411/9999-016A-2019-10089

Исследуется влияние пуассоновского шума на поведение стохастического осциллятора Ван-дер-Поля. Путем статистического моделирования траекторий решения стохастического дифференциального уравнения (СДУ), задающего осциллятор, численно исследовано влияние параметров СДУ на характер колебаний первого и второго моментов решения СДУ и влияние размеров шага интегрирования численной схемы и ансамбля моделируемых траекторий решения СДУ на точность оценок моментов решения при большой величине скачков пуассоновской составляющей.

Ключевые слова: стохастическое дифференциальное уравнение, осциллятор Ван-дер-Поля, пуассоновская составляющая.

Введение

Исследования стохастических осцилляторов с полиномиальной нелинейностью, в частности, осцилляторов Ван-дер-Поля и Дуффинга, при наличии нормального белого шума или пуассоновского шума ранее проводились в работах [1–3]. Настоящая работа является продолжением работ [4–6] по исследованию влияния пуассоновского шума на поведение осциллятора Ван-дер-Поля и на оценки функционалов от численного решения СДУ, задающего осциллятор, с помощью обобщенного явного метода Эйлера. Статистическое моделирование траекторий решения СДУ Ван-дер-Поля с пуассоновской составляющей осуществляется по схеме, приведенной в [4]. Численно исследовано влияние параметров СДУ на характер колебаний первого и второго моментов решения СДУ и влияние размеров шага интегрирования численной схемы и ансамбля моделируемых траекторий решения СДУ на точность оценок моментов решения в случае большой постоянной величины скачков пуассоновской составляющей, приводящей к появлению высокочастотных осцилляций на фоне основных колебаний оценок моментных функций.

1 Осциллятор Ван-дер-Поля

Рассматривается осциллятор Ван-дер-Поля с пуассоновским шумом, задающийся СДУ второго порядка

$$\frac{d^2z(t)}{dt^2} - \mu(1 - bz^2(t))\frac{dz(t)}{dt} + \omega^2 z(t) = \mathrm{P}(t)z(t), \quad 0 \le t \le T, \tag{1}$$

где $\mu \ge 0$ — параметр нелинейности, параметр ω совместно с μ определяет частоту колебаний, $b > 0$ — масштабирующий коэффициент, $\mathrm{P}(t)$ — пуассоновский белый шум вида

$$\mathrm{P}(t) = \sum_{k=1}^{N_T} \xi_k \delta(t - \tau_k). \tag{2}$$

В (2) N_T — случайное число скачков на интервале $(0, T]$ (процесс Пуассона), ξ_k — независимые одинаково распределенные случайные величины скачков в случайные моменты времени τ_k, ξ_k независимы с τ_k, $\delta(t)$

Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2019-0002).

ISBN 978-5-901548-42-4

— дельта-функция. При $N_T = 0$ формально будем полагать $\mathrm{P}(t) \equiv 0$. Вводя новые переменные $y_1 = z$, $y_2 = dz/dt$, СДУ (1) можно переписать в виде системы СДУ

$$\begin{aligned} dy_1 &= y_2 dt, \quad y_1(0) = y_{10}, \\ dy_2 &= \big(\mu y_2(1 - by_1^2) - \omega^2 y_1\big)dt + y_1 \int\limits_{\mathbf{R}} x\nu(dx, dt), \quad y_2(0) = y_{20}, \end{aligned} \tag{3}$$

где ν — однородная пуассоновская мера в пространстве $\mathbf{R} \times [0, T]$ с характеристической мерой, определяемой функцией

$$\varphi(x) = \lambda f_\xi(x),$$

λ — интенсивность скачков, $f_\xi(x)$ — плотность распределения случайной величины скачков ξ с математическим ожиданием m_ξ и дисперсией σ_ξ^2. В случае пуассоновского процесса со скачками постоянной величины формально будем задавать $\varphi(x) = \lambda\delta(x - m_\xi)$. Второе уравнение системы (3) можно представить с использованием центрированной меры $\widetilde{\nu}(dx, dt) = \nu(dx, dt) - \varphi(x)dxdt$ в виде

$$dy_2 = \big(\mu y_2(1 - by_1^2) - \widetilde{\omega}^2 y_1\big)dt + y_1 \int\limits_{\mathbf{R}} x\widetilde{\nu}(dx, dt),$$

где $\widetilde{\omega}^2 = \omega^2 - \lambda m_\xi$. Численные исследования решения системы СДУ (3) показали [4], что при нормальной плотности распределения $f_\xi(x)$ разные значения параметров пуассоновской составляющей λ, m_ξ, σ_ξ обеспечивают близкие по скорости затухания колебаний оценки математического ожидания первой компоненты решения системы СДУ в случае постоянства величин $\lambda(m_\xi^2 + \sigma_\xi^2)$, $\widetilde{\omega}$ и μ. В случае постоянных скачков пуассоновской составляющей ($m_\xi > 0$, $\sigma_\xi = 0$) появляются высокочастотные осцилляции ("биения"), накладывающиеся на основные колебания оценок моментных функций [6].

2 Численные эксперименты

В численной схеме решения СДУ (3) на основе обобщенного явного метода Эйлера используется адаптированная к скачкам временная сетка, представляющая собой суперпозицию заданной равномерной временной сетки $\{t_n\}$, $n = 0, ..., N_h$, $t_0 = 0$, $t_{N_h} = T$ с постоянным шагом $h = T/N_h$ и множества моментов скачков моделируемой траектории решения [4]. Величины скачков моделируются постоянными ($m_\xi > 0$) либо случайными нормальными с математическим ожиданием m_ξ и дисперсией σ_ξ^2. Оценки математического ожидания $m_i(t) = \mathbf{E}y_i(t)$ и второго момента $s_{ii}(t) = \mathbf{E}y_i^2(t)$, $i = 1, 2$, каждой компоненты численного решения случайного процесса $\{y_1(t), y_2(t)\}$ на основе ансамбля из N_{tr} смоделированных траекторий решения вычисляются стандартным способом [4]. Ниже исследуется только первая компонента решения, в приведенных ниже примерах задавались значения $y_{10} = 3$, $y_{20} = 0$, интервал интегрирования $T = 200$.

Пример 1. Исследуется зависимость точности оценок математического ожидания $m_1(t)$ и второго момента $s_{11}(t)$ решения системы СДУ Ван-дер-Поля (3) от размеров шага интегрирования численной схемы и ансамбля моделируемых траекторий решения в случае большой постоянной величины скачков пуассоновской составляющей. Значения параметров СДУ: $\mu = 10$, $b = 0.3$, $\omega = \sqrt{1.2}$, $\lambda = 0.1$, $m_\xi = 2$, $\sigma_\xi = 0$. Шаг интегрирования h уменьшался от значения 10^{-3} ($N_h = 2 \cdot 10^5$) до 10^{-5} ($N_h = 2 \cdot 10^7$), число моделируемых траекторий решения увеличивалось от $N_{tr} = 10^5$ до $N_{tr} = 10^8$ при $h = 10^{-3}$ и от $N_{tr} = 10^5$ до $N_{tr} = 10^6$ при $h = 10^{-5}$. Как видно из рис. 1, увеличение размера шага интегрирования ведет, как и при отсутствии "биений" [6], к завышению периода колебаний оценки математического ожидания решения. При малых размерах шага $h \le 10^{-4}$ графики оценок математического ожидания визуально совпадают между собой. Рис. 2 показывает, что для оценки второго момента решения, имеющей медленно затухающие "биения", увеличение размера шага интегрирования ведет к некоторому завышению периода "биений", меньшей их амплитуде и, кроме того, как и в случае их отсутствия [5], к завышению стационарного значения, на которое с ростом t постепенно выходит эта оценка. Визуальное совпадение графиков оценок второго момента, в том числе "биений", наблюдается при $h \le 10^{-5}$.

Увеличение размера ансамбля моделируемых траекторий решения N_{tr} при $h = 10^{-3}$ с 10^5 до 10^8, при $h = 10^{-4}$ с 10^5 до 10^7 и при $h = 10^{-5}$ с 10^5 до 10^6 не приводит в каждом из этих трех случаев к визуальному изменению графиков оценок моментных функций для первого и второго моментов решения СДУ. Сглаживания "биений" при увеличении N_{tr} не наблюдается.

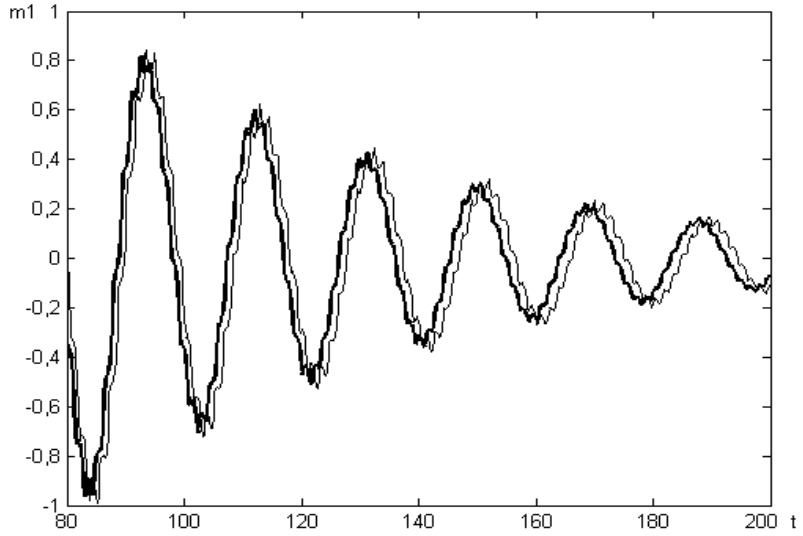


Рис. 1: Оценки математического ожидания решения СДУ Ван-дер-Поля, $\mu = 10$, при $h = 10^{-3}$, $N_{tr} = 10^8$ (тонкая линия) и $h = 10^{-5}$, $N_{tr} = 10^6$ (жирная линия)

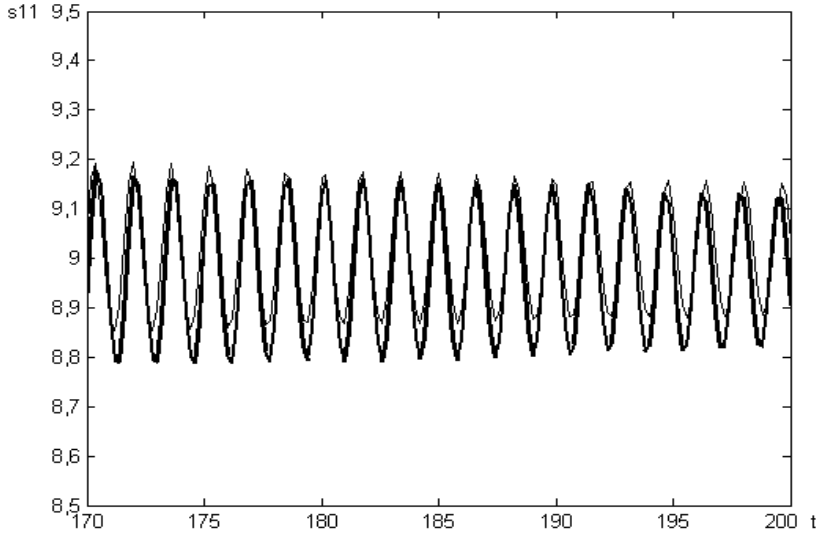


Рис. 2: Оценки второго момента решения СДУ Ван-дер-Поля, $\mu = 10$, при $h = 10^{-3}$, $N_{tr} = 10^8$ (тонкая линия) и $h = 10^{-5}$, $N_{tr} = 10^6$ (жирная линия)

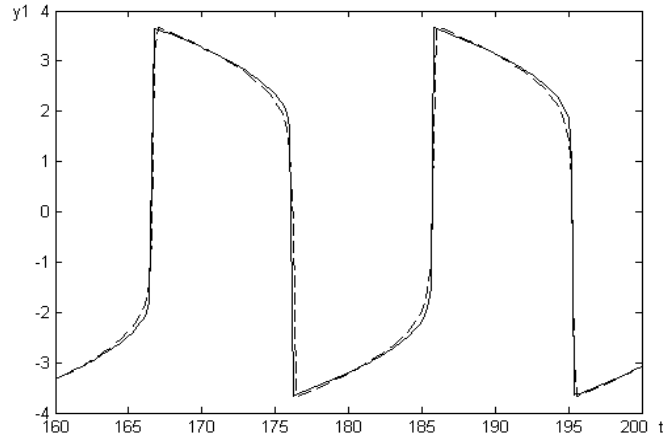


Рис. 3: Решения ОДУ Ван-дер-Поля при $\mu = 50$, $\omega = \sqrt{4.48}$ (сплошная линия) и $\mu = 10$, $\omega = 1$ (штриховая линия)

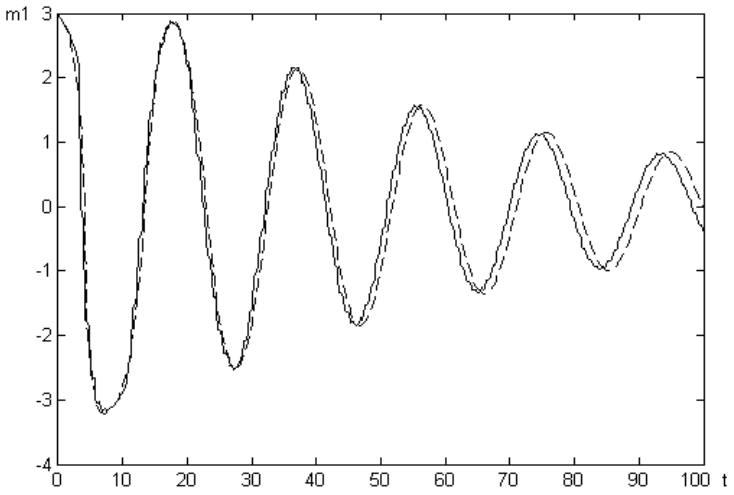


Рис. 4: Оценки математического ожидания решения СДУ Ван-дер-Поля, $\mu = 50$, при $m_\xi = 5$, $\sigma_\xi = 0$ (сплошная линия) и $m_\xi = 0$, $\sigma_\xi = 5$ (штриховая линия)

Пример 2. Исследуются оценки математического ожидания $m_1(t)$ и второго момента $s_{11}(t)$ решения системы СДУ Ван-дер-Поля (3) при различных значениях параметра нелинейности μ в случае большой величины скачков пуассоновской составляющей. Сопутствующей системой ОДУ для СДУ Ван-дер-Поля является система (3) без пуассоновской составляющей. При больших μ ОДУ Ван-дер-Поля описывает релаксационные колебания с периодом, пропорциональным μ, поэтому увеличение μ при постоянной величине ω требует соответствующего увеличения интервала интегрирования $[0, T]$, если требуется оставить неизменным число периодов колебаний на интервале интегрирования. С другой стороны, можно оставить неизменными интервал интегрирования и число колебаний, а увеличивать величину ω, и тогда потребуется при численном решении СДУ уменьшать размер шага интегрирования h. На рис. 3 приведены графики численных решений двух ОДУ Ван-дер-Поля: одного при $\mu = 10$, $\omega = 1$, $b = 0.3$, соответствующего рассмотренному в примере 1 СДУ, и второго с близким периодом колебаний решения при $\mu = 50$, $\omega = \sqrt{4.48}$, $b = 0.3$. В обоих случаях шаг интегрирования $h = 2 \cdot 10^{-8}$.

На рис. 4 приведен график оценки математического ожидания решения СДУ Ван-дер-Поля при $\mu = 50$. Значения параметров СДУ подобраны таким образом, чтобы обеспечить скорость затухания ее колебаний, близкую к случаю оценки при $\mu = 10$, рассмотренному в примере 1: $\lambda = 0.28$, $m_\xi = 5$, $\omega^2 = 4.48 + \lambda m_\xi = 5.88$,

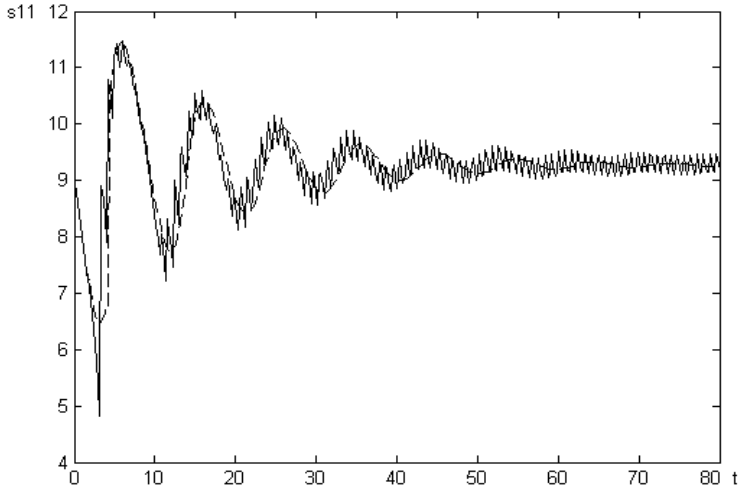


Рис. 5: Оценки второго момента решения СДУ Ван-дер-Поля, $\mu = 50$, при $m_\xi = 5$, $\sigma_\xi = 0$ (сплошная линия) и $m_\xi = 0$, $\sigma_\xi = 5$ (штриховая линия)

$\sigma_\xi = 0$, $b = 0.3$, значения параметров численной схемы $N_{tr} = 10^6$, $h = 10^{-5}$. Графики оценок математического ожидания решения СДУ (рис. 4) и второго момента (рис. 5) показывают меньшую амплитуду "биений" при большем значении μ, а в остальном $m_1(t)$ и $s_{11}(t)$ ведут себя сходно при $\mu = 10$ и $\mu = 50$. Для оценок в случае нормальной случайной величины скачков при $\sigma_\xi = 5$, $m_\xi = 0$, $\omega^2 = 4.48$, как видно из рис. 4, 5, "биений" не наблюдается.

Заметим, что вычисления оценок первого и второго моментов решения СДУ Ван-дер-Поля для варианта "биений" при заданных выше значениях параметров СДУ $\mu = 50$, $m_\xi = 5$, $\sigma_\xi = 0$, но при других значениях параметров численной схемы: $N_{tr} = 10^5$, $h = 10^{-6}$, показали визуальное совпадение графиков этих оценок с соответствующими графиками, приведенными на рис. 4, 5 (сплошные линии).

Пример 3. Подавление колебаний осциллятора Ван-дер-Поля пуассоновской составляющей происходит в том случае, когда значения ее параметров в СДУ (3) удовлетворяют условию $\lambda m_\xi = \omega^2$, т.е. $\widetilde{\omega} = 0$. Рис. 6 демонстрирует подавление колебаний и одновременное возникновение "биений" оценки математического ожидания $m_1(t)$ решения СДУ при $m_\xi = 2$, $\lambda = 0.125$, $\omega = 0.5$. Период "биений" в данном примере $T_f \approx 1/\lambda$. Для сравнения на рис. 6 приведен также вариант, когда таких "биений" не наблюдается при $m_\xi = 0.25$, $\lambda = 8$, $\omega = \sqrt{2}$. Для обоих вариантов выполняется условие постоянства величины λm_ξ^2, что соответствует, как указывалось выше, близким скоростям затухания основных колебаний оценок при различных значениях параметров СДУ. Для обоих вариантов $\mu = 10$, $b = 0.3$, $\sigma_\xi = 0$, значения параметров численной схемы $N_{tr} = 10^6$, $h = 10^{-5}$. Заметим, что в обоих вариантах не происходит выход оценки второго момента решения СДУ на стационарное значение на рассматриваемом интервале интегрирования T.

Заключение

Проведенные численные исследования влияния пуассоновской составляющей на решение СДУ Ван-дер-Поля показали, что при использованных в настоящей работе размерах шага интегрирования численной схемы и ансамбля моделируемых траекторий решения СДУ большие постоянные величины скачков $m_\xi > 0$ пуассоновской составляющей приводят к появлению высокочастотных осцилляций, накладывающихся на основные колебания оценок первого и второго моментов решения СДУ. Амплитуда таких "биений" уменьшается с увеличением значения параметра нелинейности μ в случае постоянства частоты и скорости затухания основных колебаний оценок моментных функций. При нормальной случайной величине скачков с нулевым математическим ожиданием и большой дисперсией σ_ξ^2 "биения" отсутствуют.

При использовании для численного решения СДУ Ван-дер-Поля обобщенного явного метода Эйлера величину шага интегрирования численной схемы решения, необходимую для достижения требуемой точности вычислений, можно приближенно оценить, численно решая несколько раз сопутствующую систему ОДУ

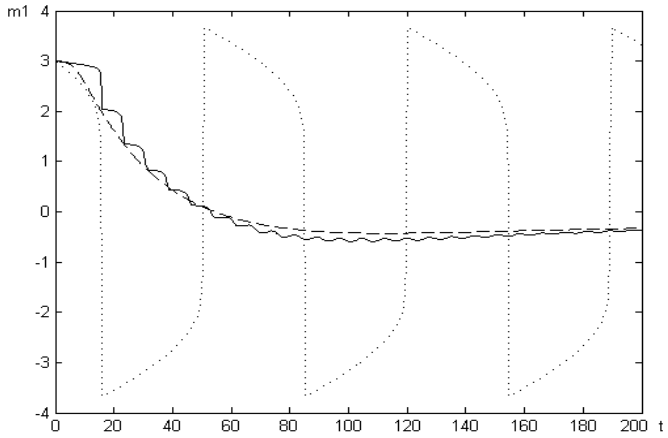


Рис. 6: Оценки математического ожидания решения СДУ Ван-дер-Поля в случае подавления колебаний ($\omega^2 = \lambda m_\xi$) при $m_\xi = 2$, $\lambda = 0.125$ (сплошная линия) и $m_\xi = 0.25$, $\lambda = 8$ (штриховая линия). Точечная линия – решение сопутствующей системы ОДУ Ван-дер-Поля при $\omega = 0.5$

Ван-дер-Поля явным методом Эйлера при различных постоянных значениях h и оценивая по этим результатам уменьшение периода релаксационных колебаний решения, т.е. повышение точности вычислений, при уменьшении h. Статистическую погрешность оценки математического ожидания решения СДУ при больших t можно приближенно оценить стандартным способом [4], используя оценку второго момента решения СДУ.

Список литературы

[1] Guo S.S., Er G.K. The probabilistic solution of stochastic oscillators with even nonlinearity under poisson excitation // Central European Journal of Physics. 2012. Vol. 10, N 3. P. 702–707.

[2] Er G.K., Zhu H.T., Iu V.P., Kou K.P. PDF solution of nonlinear oscillators subject to multiplicative Poisson pulse excitation on displacement // Nonlinear Dynamics. 2009. Vol. 55, iss. 4. P. 337–348.

[3] Ruoshi Yuan, Xinan Wang, Yian Ma, Bo Yuan, and Ping Ao. Exploring a noisy van der Pol type oscillator with a stochastic approach // Physical Review E87. 2013, 062109.

[4] Артемьев С.С., Якунин М.А. Параметрический анализ осциллирующих решений СДУ с винеровской и пуассоновской составляющими методом Монте-Карло // Сиб. журн. индустриальной математики. 2017. Т. 20, № 2. С. 3–14.

[5] Якунин М.А. Анализ осциллятора Ван-дер-Поля с пуассоновской составляющей методом статистического моделирования // Труды Международной конференции "Вычислительная математика и математическая геофизика", посвященной 90-летию со дня рождения академика А.С. Алексеева. ИВМиМГ СО РАН. Новосибирск. 8 – 12 октября 2018 г. С. 419-424. [Электрон. ресурс].

[6] Якунин М.А. Анализ стохастических осцилляторов методом статистического моделирования // Труды Международной конференции "Марчуковские научные чтения – 2017". ИВМиМГ СО РАН. Новосибирск. 25 июня – 14 июля 2017 г. С. 1033-1038. [Электрон. ресурс]. http://conf.nsc.ru/cam17/ru/proceedings.

Якунин Михаил Александрович — к.ф.-м.н., ст.науч.сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: yma@osmf.sscc.ru.
Дата поступления — 30 мая 2019 г.

ЧИСЛЕННЫЙ МЕТОД ПРОГНОЗИРОВАНИЯ ТЕМПЕРАТУРЫ С ПОМОЩЬЮ УРАВНЕНИЯ ВОЛЬТЕРРА

Н. М. Япарова, Т. П. Гаврилова

Южно-Уральский государственный университет (НИУ), 454080, Челябинск

УДК 519.6, 517.9
DOI: 10.24411/9999-016A-2019-10090

В работе рассматривается обратная граничная задача для уравнения параболического типа с начальными и граничными условиями, известными на части границы, связанная с задачей определения температуры на одной из границ и внутри линейного объекта по результатам измерения температурных функций, проводимых вблизи другой границы. С помощью преобразования Лапласа задача сводится к интегральному уравнению, характеризующему прямую зависимость неизвестной граничной функции от исходных данных. В работе предложен численный метод решения интегрального уравнения, основанный на методе регуляризации с апостериорным выбором параметров регуляризации. Решение этого уравнения послужило основой для прогнозирования температуры внутри всего линейного объекта. С целью проверки эффективности предложенного подхода и получения экспериментальных оценок погрешностей был осуществлен вычислительный эксперимент, результаты которого представлены в работе и свидетельствуют о достаточной точности полученных численных решений.
Ключевые слова: параболическое уравнение, обратная граничная задача, преобразование Лапласа, численные методы, вычислительный эксперимент.

Введение

Развитие современных технологий в металлургии, машиностроении, химической промышленности неразрывно связано с исследованием сущности процессов теплообмена. Математические модели теплопереноса представлены обратными задачами для уравнений параболического типа с начальными условиями и граничными условиями, известными на части границы. Исследованию методов решения обратных граничных задач, посвящены работы многих исследователей [1]– [7].

В работе предложена интегральная модель теплопереноса, полученная с помощью прямого и обратного преобразований Лапласа, и численный метод определения температуры на одной из границ и внутри линейного объекта по результатам изменения температурных функций на другой границе и вблизи нее.

В данной работе для построения численного решения разработан метод, основанный на схеме Лаврентьева с апостериорным выбором параметров регуляризации. Устойчивость вычислительной схемы достигается за счет согласования параметров регуляризации с шагами дискретизации. С целью получения экспериментальных оценок погрешностей и оценки эффективности предложенного подхода к решению обратной граничной задачи проведен вычислительный эксперимент, результаты которого представлены в работе.

1 Постановка задачи

Рассмотрим обратную задачу линейного теплопереноса

$$u_t = au_{xx}, \quad x \in (0, l), \quad t \geq 0, \tag{1}$$

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации прикладных научных исследований в рамках базовой части Государственного задания «Разработка, исследование и реализация алгоритмов обработки данных динамических измерений пространственно-распределенных объектов», техническое задание 8.9692.2017/8.9 от 17.02.2017.

ISBN 978-5-901548-42-4

$$u(0,t)=\varphi(t),\quad u(x_0,t)=g(t),\quad t\geq 0, \tag{2}$$

$$u(x,0)=C,\quad x\in[0,l], \tag{3}$$

где x_0 — точка, расположенная вблизи границы линейного объекта. В данной задаче требуется найти граничное значение функции

$$u(l,t)=\psi(t), \tag{4}$$

и затем, используя $\psi(t)$, найти температурные значения во внутренних точках линейного объекта.

Основываясь на физических предпосылках математической модели, полагаем, что $g(t),\varphi(t)\in C^{2+\eta}[0,T]$ при всех $T>0$ и $\eta\in(0,1)$. Также существуют постоянные $C_0,C_1,C_2\geq 0$ и $\beta_0,\beta_1,\beta_2\geq 0$ такие, что $|u(x,t)|\leq C_0e^{\beta_0(x+t)}$, $|\varphi(t)|\leq C_1e^{\beta_1 t}$ и $|\psi(t)|\leq C_2e^{\beta_2 t}$ выполнены для $x\in[0,l]$ и при всех $t\in[0,\infty)$. При этом функции $\varphi(t)$, $\psi(t)$ и $g(t)$ удовлетворяют условиям Дирихле для любого $t\in[0,T]$ при всех $T>0$. Существование и единственность решения $u(x,t)\in H^{2,1}((0,l)\times(0,T))\cap C([0,l]\times[0,T])$ задачи (1)–(4) обоснованы в [1], [5].

Существенная особенность рассматриваемой задачи заключается в том, что результаты измерений, полученные в процессе эксперимента, содержат отклонения от истинных значений. Представим эту ситуацию следующим образом. Вместо g_0 и φ_0 известны некоторые приближенные значения g_δ и φ_δ и уровень погрешности $\delta>0$ такой, что $\max\{\|\varphi_\delta-\varphi_0\|,\|\mathrm{g}_\delta-\mathrm{g}_0\|\}\leq\delta$. В данной задаче требуется по φ_δ, g_δ, δ найти граничное значение функции

$$u_\delta^\alpha(l,t)=\psi_\delta^\alpha(t)$$

и, основываясь на полученных результатах, определить температуру $u_\delta^\alpha(x,t)$ вдоль всего линейного объекта.

2 Интегральная модель теплообмена и численный метод решения задачи теплообмена

Для получения интегрального уравнения, связывающего исходные данные и $u(l,t)=\psi(t)$, найдем решение следующей прямой задачи, полагая, что искомая функция $\psi(t)$ нам известна:

$$u_t=au_{xx},\quad x\in(0,l),\quad t\geq 0, \tag{5}$$

$$u(0,t)=\varphi(t),\quad u(l,t)=\psi(t),\quad t\geq 0, \tag{6}$$

$$u(x,0)=C,\quad x\in[0,l]. \tag{7}$$

Исходя из свойств функций и следуя результатам, представленным в работе [8], применим прямое преобразование Лапласа в задаче (5)–(7), получим следующую формулу для представления функции $u(x,t)$:

$$\begin{aligned}u(x,t)=&\frac{2\pi a}{l^2}\sum_{n=1}^{\infty}n\sin\left(\frac{\pi nx}{l}\right)e^{\frac{-n^2\pi^2a}{l^2}t}\int_0^t\varphi(\tau)e^{\frac{n^2\pi^2a}{l^2}\tau}\,d\tau+\\+&\frac{2\pi a}{l^2}\sum_{n=1}^{\infty}(-1)^{n+1}n\sin\left(\frac{\pi nx}{l}\right)e^{\frac{-n^2\pi^2a}{l^2}t}\int_0^t\psi(\tau)e^{\frac{n^2\pi^2a}{l^2}\tau}\,d\tau+\\+&\frac{4C}{\pi}\sum_{n=1}^{\infty}\frac{1}{2n-1}\sin\frac{(2n-1)\pi x}{l}\cdot e^{\frac{-(2n-1)^2\pi^2a}{l^2}t}.\end{aligned} \tag{8}$$

Так как решение задачи (1)–(4) требуется построить в ограниченной области $[0,l]\times[0,T]$, а для построения численного решения применяют, как правило, сетку с конечным числом узлов, то, используя идею, представленную в [4], аппроксимируем точное решение задачи (5)–(7) конечным рядом. Учитывая, что функция $u(x_0,t)=g(t)$ известна, получаем, что решение задачи (1)–(4) сводится к решению следующего интегрального уравнения:

$$A\psi=\int_0^t K_N(t-\tau)\psi(\tau)\,d\tau=h(t), \tag{9}$$

где

$$K_N(t-\tau)=\frac{2\pi a}{l^2}\sum_{n=1}^{N}(-1)^{n+1}n\sin\left(\frac{\pi nx_0}{l}\right)e^{\frac{-n^2\pi^2a}{l^2}(t-\tau)},$$

$$L_N(t-\tau) = \frac{2\pi a}{l^2}\sum_{n=1}^{N} n \sin\left(\frac{\pi n x_0}{l}\right) e^{\frac{-n^2\pi^2 a}{l^2}(t-\tau)},$$

$$h(t) = g(t) - f(t) - \int_0^t L_N(t-\tau)\varphi(\tau)d\tau,$$

$$f(t) = \frac{4C}{\pi}\sum_{n=1}^{N}\frac{1}{2n-1}\sin\frac{(2n-1)\pi x_0}{l} e^{\frac{-(2n-1)^2\pi^2 a}{l^2}t}.$$

Численные решения уравнения (9) получаем, используя вычислительную схему, основанную на методе регуляризации, идея которого предложена в [8]. Согласно этому подходу, регуляризованное решение поставленной задачи определяется из уравнения

$$A\psi + \alpha\psi = h_\delta(t). \tag{10}$$

При этом должны соблюдаться условия $\|A\psi - h_\delta(t)\| \le \delta$, где

$$h_\delta(t) = g_\delta(t) - \int_0^t L_N(t-\tau)\varphi_\delta(\tau)\,d\tau.$$

Параметр α выбираем апостериорно, используя итерационный процесс и идею, предложенную в [9]. Для этого аппроксимируем интегралы, входящие в уравнение (9), суммами, используя метод правых прямоугольников. Из полученной при этом системы алгебраических уравнений находим $\psi_\delta^\alpha(t)$. После построения $\psi_\delta^\alpha(t)$ вычисляем абсолютную погрешность Δ_ψ, определяемую формулой

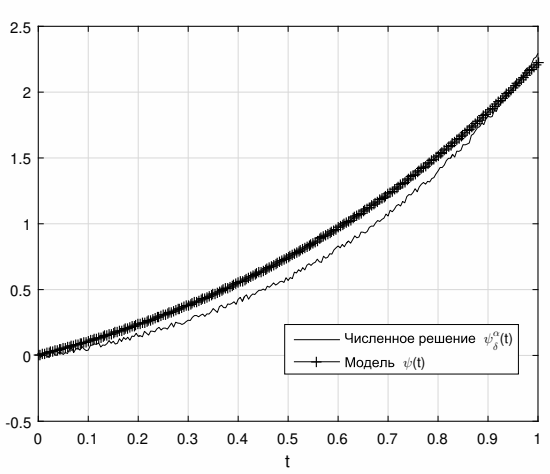
$$\Delta_\psi = \|\psi_\delta^\alpha - \psi\|. \tag{11}$$

Значение параметра α определяем, исходя из условия минимальности (11).

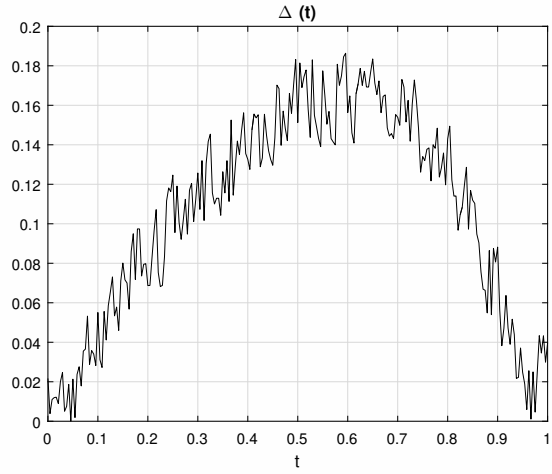
3 Результаты вычислительного эксперимента

На рисунках и в таблице представлены результаты вычислительного эксперимента для тестовых функций $\varphi(t) = te^{0,8t},\ \psi(t) = 0,5(e^{1,8t} - 1)$.

На рис. 1 представлены графики тестовой функций $\psi(t)$ и приближенного численного решения ψ_δ^α уравнения (10). Также представлен график функции погрешности, вычисляемой по формуле $\Delta_\psi(t) = |\psi_\delta^\alpha - \psi(t)|$.

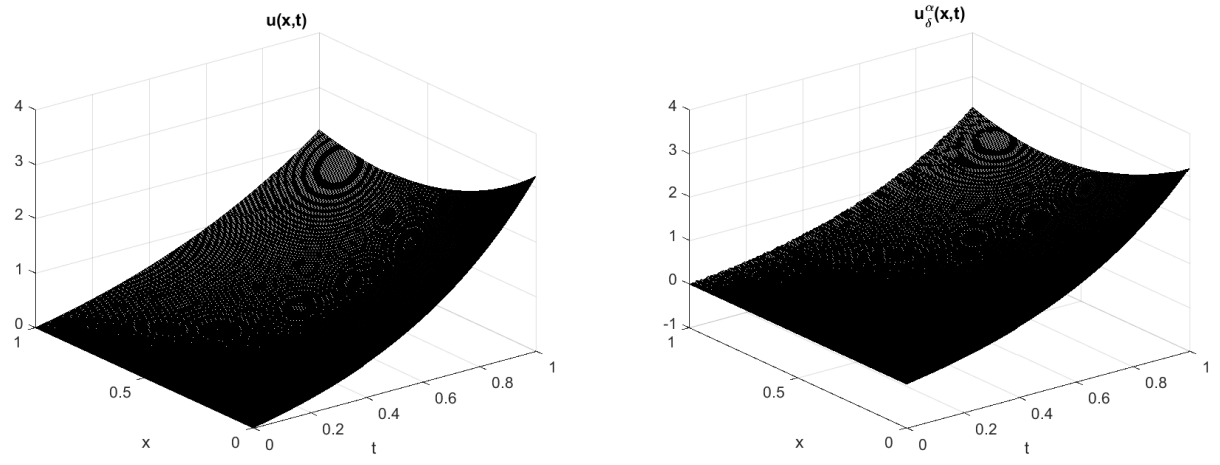


а



б

Рис. 1: *а* — Результаты численного решения интегрального уравнения (11); *б* график функции погрешности $\Delta_\psi(t)$



а *б*

Рис. 2: Результаты численных решений прямой задачи (5)–(7): *а* — для тестовой функции $\psi(t)$; *б* — на основе $\psi_\delta^\alpha(t)$, найденного из уравнения (11)

Таблица 1: Экспериментальные оценки погрешностей

Погрешность исходных данных δ	Абсолютная погрешность Δ_ψ	Относительная погрешность ε_ψ
0,1	0,1961	0,0923
0,05	0,1894	0,0885
0,03	0,1802	0,0821

На рис. 2 построена поверхность $u(x,t)$, соответствующая решению прямой задачи (5)–(7) и поверхность $u_\delta^\alpha(x,t)$, соответствующая приближенному решению обратной задачи.

На основании полученных результатов эксперимента можно сделать следующие выводы. Предложенный метод решения задачи линейного теплообмена позволяет получать регуляризованные решения с удовлетворительной точностью. При этом уровень погрешности исходных данных оказывает влияние на величину уклонения приближенного решения $\psi_\delta^\alpha(t)$ от тестовых функций $\psi(t)$.

Список литературы

[1] Алифанов О.М. Обратные задачи теплообмена. М.: Машиностроение, 1988. 280 с.

[2] Васин В.В., Агеев А.Л. Некорректные задачи с априорной информацией. Екатеринбург: Наука, 1993. 264 с.

[3] Кабанихин С.И. Обратные и некорректные задачи. Новосибирск: Сиб. научн. изд-во, 2009. 457 с.

[4] Лаврентьев М.М., Романов В.Г., Шишатский С.П. Некорректные задачи математической физики и анализа. М.: Наука, 1980. 287 с.

[5] Апарцин А.С., Бакушинский А.Б. Приближенное решение интегральных уравнений Вольтерра I рода методом квадратурных сумм // Дифференциальные и интегральные уравнения. Иркутск: Иркут. гос. ун-т. 1972. Вып. 1. С. 248—258.

[6] Васин В.В. Регулярный алгоритм аппроксимации негладких решений для интегральных уравнений Фредгольма первого рода // Выч. технологии. 2010. Т.15. № 2. С. 15–23.

[7] Марчук Г.И. Методы вычислительной математики. СПб.: Лань, 2009. 608 с.

[8] Лаврентьев М.М. Условно-корректные задачи для дифференциальных уравнений. Новосибирск: НГУ, 1973. 71 с.

[9] Леонов А.С. О квазиоптимальном выборе параметра регуляризации в методе Лаврентьева // Сиб. матем. журн. 1993. Т.34, № 4. С. 695–703.

Япарова Наталья Михайловна — к.ф.-м.н., доцент, зав. кафедры вычислительной математики и высокопроизводительных вычислений, Южно-Уральский государственный университет (национальный); e-mail: iaparovanm@susu.ru;

Гаврилова Татьяна Петровна — старший преподаватель кафедры вычислительной математики и высокопроизводительных вычислений, Южно-Уральский государственный университет (национальный); e-mail: gavrilovatp@susu.ru.

Дата поступления — 31 мая 2019 г.

ЧИСЛЕННАЯ РЕКОНСТРУКЦИЯ ПОЛЕЙ АЭРОЗОЛЬНЫХ ВЫПАДЕНИЙ ПРИМЕСЕЙ ПО ДАННЫМ НАЗЕМНОГО И СПУТНИКОВОГО МОНИТОРИНГА

Т. В. Ярославцева[1], Р. А. Амикишиева[2,3], В. Ф. Рапута[2]

[1] *Новосибирский институт гигиены, 630108, Новосибирск*

[2] *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

[3] *Новосибирский государственный университет, 630090, Новосибирск*

УДК 551.511.42.001.572
DOI: 10.24411/9999-016A-2019-10091

Проводится численный анализ результатов спутниковых наблюдений загрязнения снежного покрова в окрестностях промышленных предприятий и автотрасс. Основу исследования составляют модели реконструкции полей аэрозольных выпадений примесей в локальном и региональном масштабе от точечных, линейных и площадных источников атмосферных примеси. Апробация предложенного подхода выполнена с использованием зимнего спутникового снимка окрестностей г. Искитима и пгт Линёво и модельных описаний полей загрязнения в окрестностях Новосибирского электродного завода, цеха по сухому производству цемента и автотрассы. Оцифровка полей загрязнения проводится в тонах серого цвета с последующим поиском аппроксимирующих соотношений с данными маршрутных наблюдений. Очаги загрязнения от этих источников были проиндексированы с помощью дискретной шкалы оттенков серого цвета по направлениям доминирующих выносов примесей. В зависимости от типов источников было установлено наличие функциональных связей между данными наземного и спутникового мониторинга. Показано, что в ближней зоне загрязнения для случая точечных и линейных источников характерна степенная зависимость. Для процессов загрязнения регионального масштаба предпочтительно использовать логарифмическую связь.

Ключевые слова: загрязнение, снежный покров, мониторинг, спутниковые наблюдения, тона серого цвета, модель реконструкции, функциональная связь.

Введение

Наземный мониторинг процессов загрязнения атмосферы является трудоёмкой задачей, требует привлечения значительных ресурсов. Для снижения затрат при проведении экспериментальных исследований следует использовать методы численного моделирования в рамках постановок прямых и обратных задач переноса примесей, математические методы планирования эксперимента [1-5]. В зависимости от типов источников, свойств примесей, масштабов протекающих процессов загрязнения целесообразно использовать различные приближения к решениям уравнений переноса примесей. Весьма эффективными могут быть асимптотические описания решений в приземном слое атмосферы для сравнительно небольших расстояний от источников примесей [4, 6].

При описании полей концентраций примесей в атмосфере регионального масштаба также допускаются существенные упрощения. Из анализа экспериментальных и теоретических исследований вытекает, что в зимнее время, начиная с расстояний порядка 5–7 км от источника, распределение концентрации примеси выравнивается по высоте слоя перемешивания [1, 7]. Для таких удалений влияние ряда параметров становится незначительным. К ним следует отнести высоту источника, скорость оседания аэрозольных частиц и т.д.

Работа выполнена при финансовой поддержке РФФИ и Правительства Новосибирской области в рамках научного проекта № 19-47-540008, в рамках Госзадания (№ 0315-2019-0004), программы президиума РАН № 51 (№ 0315-2018-0016).

ISBN 978-5-901548-42-4

Использование спутниковой информации создаёт возможности для снижения затрат на проведение наземного мониторинга и получения более полной картины загрязнения территорий [8–11]. В этом случае необходимым этапом исследований является поиск количественных связей между различными видами мониторинга.

Цель работы заключается в разработке методов численного анализа данных наземного и спутникового мониторинга процессов загрязнения снежного покрова в окрестностях источников выбросов примесей.

1 Объекты и методы исследований

Материалами для исследований послужили зимние снимки территорий в зонах интенсивного антропогенного загрязнения. К таковым следует отнести территории, расположенные в районе г. Искитим и пгт Линёво. На снимке с ИСЗ «Landsat-8» за 25 марта 2019 года (рис. 1), полученном в Сибирском центре ФГБУ «Научно-исследовательский центр космической гидрометеорологии «ПЛАНЕТА» (http://www.rcpod.ru), отчётливо зафиксированы зоны высокого загрязнения от источников Новосибирского электродного завода (НЭЗ) и Чернореченского цементного заводов (ЧЦЗ), цеха по производству цемента сухим способом, угольных разрезов, технологических трасс.

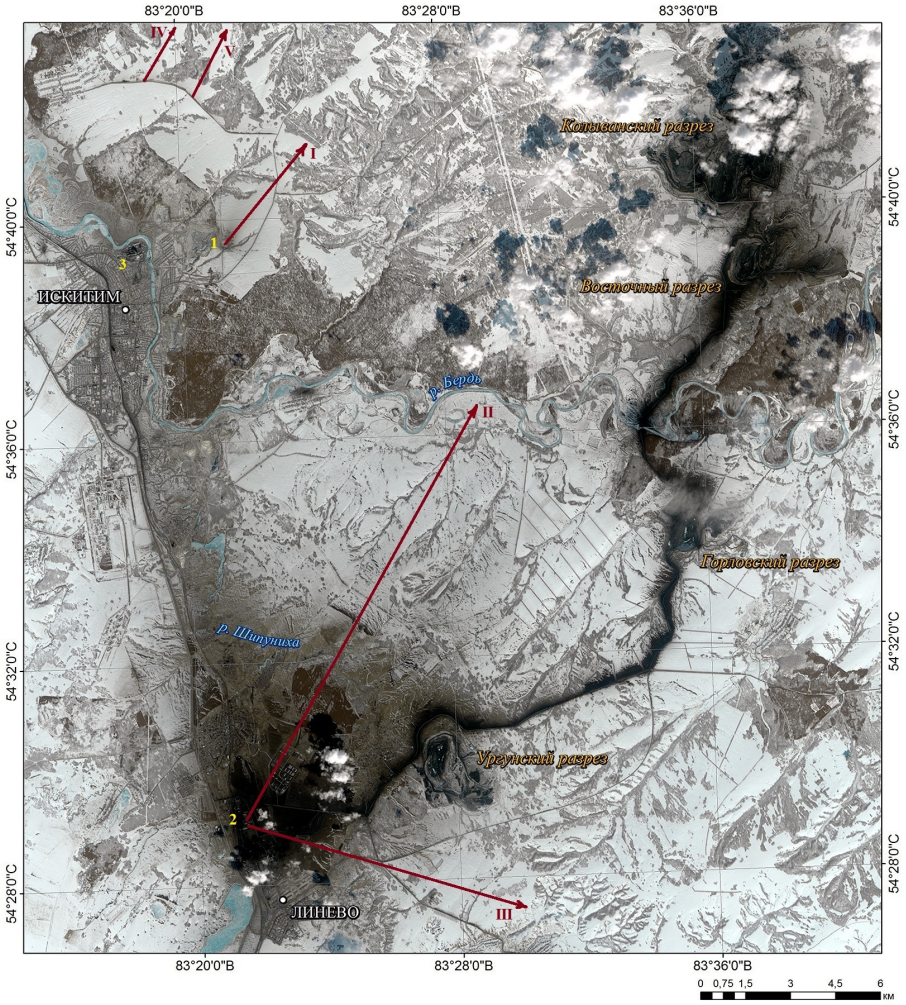


Рис. 1: Спутниковый снимок территорий в районе г. Искитима и пгт Линёво от 25 марта 2019 г. с ИСЗ «Landsat-8». 1 — цех по производству цемента сухим способом, 2 — высотные трубы обжигового цеха Новосибирского электродного завода, 3 — промплощадка Чернореченского цементного завода

1.1 Модели оценивания полей выпадений примеси

Наземный мониторинг полей загрязнений требует значительных затрат. В связи с этим представляется удобным использование для оптимизации экспериментальных исследований моделей и методов реконструкции полей аэрозольных выпадений примесей.

В случае точечного источника поле выпадений примеси в радиальном направлении от него достаточно адекватно описывается следующей зависимостью [4, 12]

$$q(r,\vec{\theta}) = \frac{\theta_1}{r^{\theta_2}} \cdot e^{\frac{-r_{max}}{r}} \tag{1}$$

где r — расстояние от источника, $\vec{\theta} = (\theta_1, \theta_2)$, θ_1 и θ_2 представляют собой комплексы физических величин, характеризующих мощность эмиссии и дисперсный состав примеси соответственно [4], r_{max} — расстояние, на котором достигается максимальная приземная концентрация для лёгкой примеси.

Для линейного источника поле выпадений примеси на определённом удалении от него можно представить в более удобном виде

$$q(r,\theta) = \frac{\theta}{r} \cdot e^{\frac{-r_{max}}{r}} \tag{2}$$

На значительном удалении от точечного источника поле описывается следующим асимптотическим соотношением [1, 12]

$$q(r,\varphi,S) = \frac{S \cdot g(\varphi)}{r} \tag{3}$$

где r, φ — полярные координаты расчетной точки с началом в месте расположения источника, $g(\varphi)$ — вероятность противоположного φ направления ветра на высотах пограничного слоя атмосферы, S — параметр, оцениваемый по данным наблюдений.

2 Результаты и обсуждение

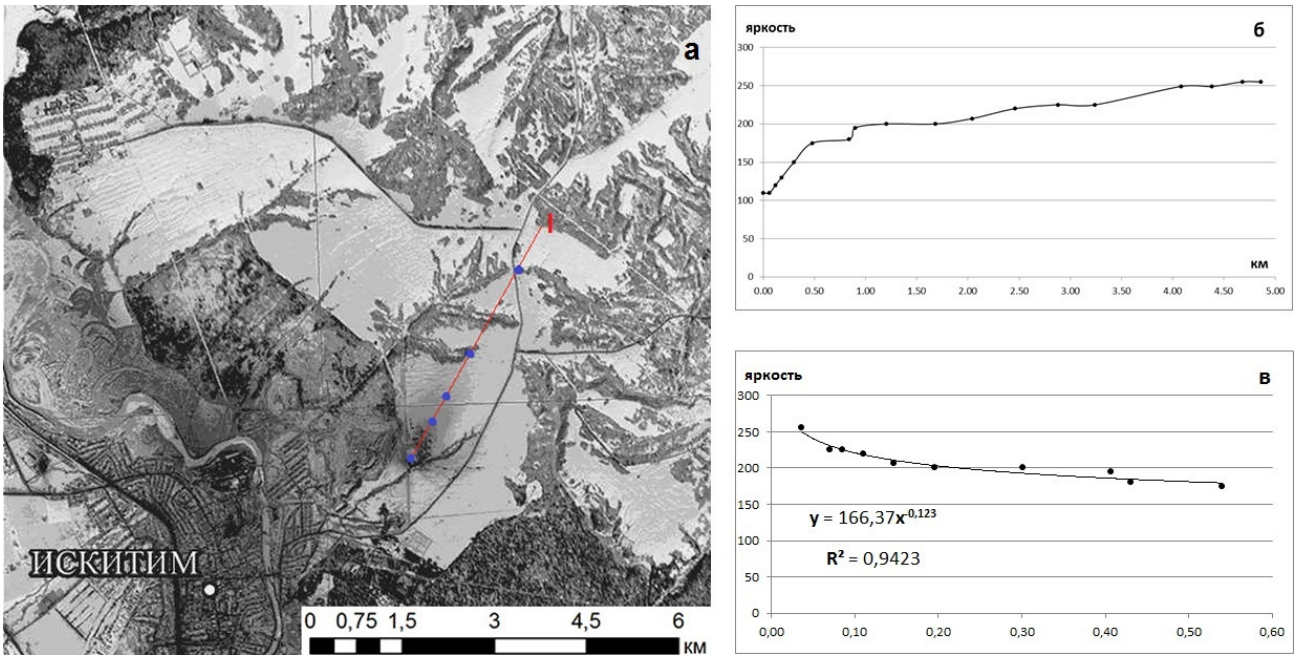


Рис. 2: Поля загрязнения снежного покрова севернее г. Искитима (25.03.2019 г.) (а). Изменения тонов серого цвета снежного покрова (б) в направлении I и функциональная связь (в)

Визуальная информация, представленная на рис. 1, с использованием соотношений (1)–(3) позволила выполнить количественное исследование загрязнения снежного покрова по направлениям выноса примеси

в окрестностях ряда доминирующих источников, включая трубы НЭЗ, ЧЦЗ, а также автотрассы Искитим–Тальменка. На рис. 2 представлен фрагмент рис. 1 с оцифровкой в тонах серого цвета поля загрязнения, создаваемого источниками цеха по сухому производству цемента.

Изменения яркостных характеристик, приведённых на рис. 2б, по мере удаления от источников выбросов пыли, меняется монотонным образом. Их максимум достигается на расстоянии более 4 км. Функциональная связь между изменениями тонов серого цвета и относительной концентрацией примеси вполне удовлетворительно описывается степенной зависимостью.

На рис. 3 представлена динамика изменения тонов в направлениях II и III от источников атмосферных выбросов НЭЗ.

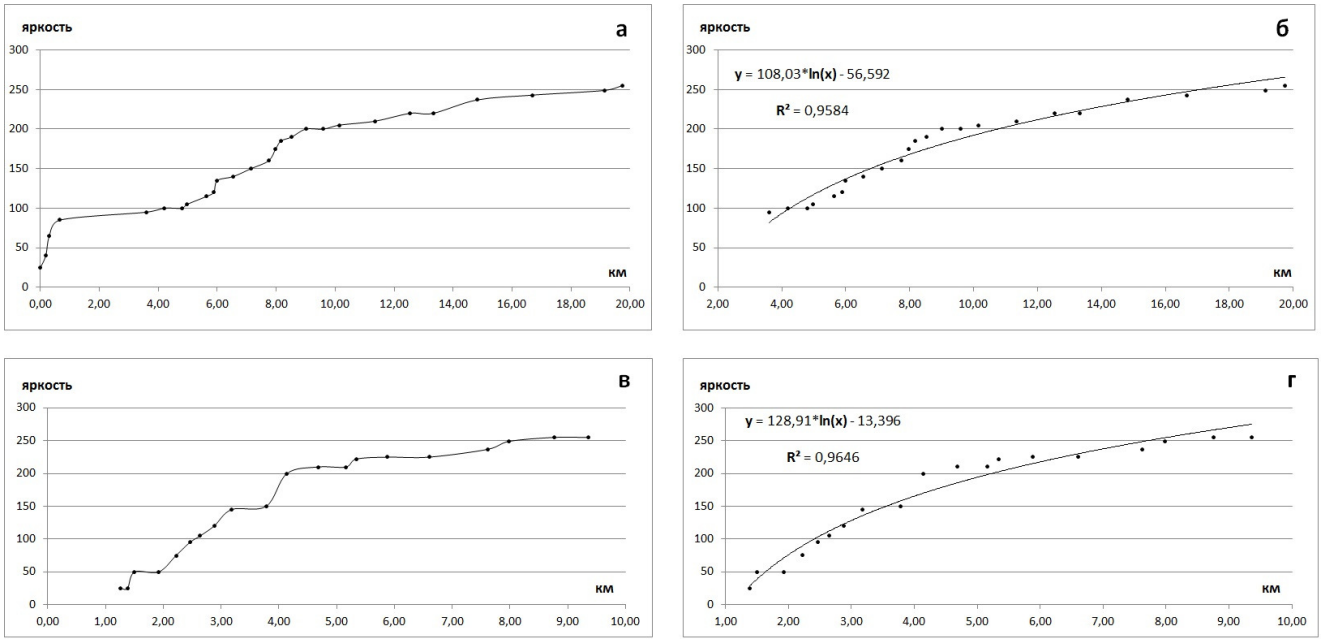


Рис. 3: Изменения тонов серого цвета снежного покрова в направлениях II и III выпадений примесей от высотных труб НЭЗ (а, в). Соответствующие функциональные связи между тонами и относительными концентрациями (б, г)

Анализ информации на рис. 3 с использованием региональной модели (3) позволил выявить наличие функциональных связей между изменениями тонов серого цвета по мере удаления от источников и динамикой уменьшения концентраций примесей. Вполне удовлетворительное их описание может быть достигнуто с помощью логарифмической аппроксимации рассматриваемых рядов данных.

На рис. 4а, 4в изображена динамика изменения тонов серого цвета по мере удаления от автотрассы Искитим–Тальменка в направлениях IV и V. На рис. 4в, 4г с использованием модели реконструкции (2) представлены функциональные связи.

Анализ результатов, приведённых на рис. 4, показывает, что выход на максимальные яркостные характеристики происходит уже на сравнительно небольших расстояниях от автотрассы. Это обусловлено сравнительно небольшой эффективной высотой источника.

Заключение

Предложены экономичные по числу опорных точек модели реконструкции локального и регионального загрязнения территорий источниками примесей. Проведена апробация разработанных моделей на данных спутниковых наблюдений аэрозольного загрязнения снежного покрова в окрестностях ряда крупных промышленных предприятий Новосибирской области. Выявлены количественные закономерности между полями выпадения взвешенных веществ и интенсивностью изменения тонов серого цвета на космических снимках в радиальных направлениях от точечных и линейных источников. Установлено, что для загрязнений регионального масштаба между этими полями характерна функциональная связь, выражаемая логарифмической

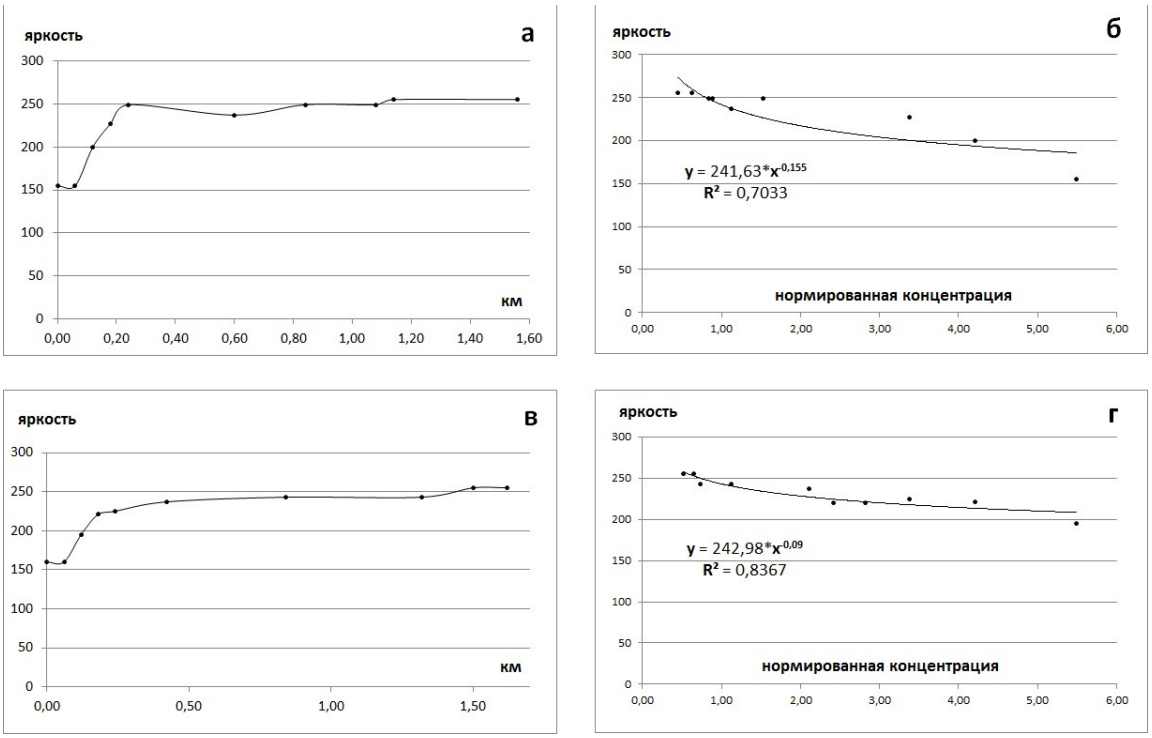


Рис. 4: Изменения тонов серого цвета в направлениях IV (а) и V (в) выпадений примесей от автотрассы Искитим–Тальменка. Функциональные связи в этих направлениях (в) и (г)

зависимостью.

Полученные зависимости позволяют существенно снизить затраты на проведение наземных мониторинговых исследований загрязнения территорий в окрестностях промышленных предприятий. По ограниченному числу опорных точек наблюдений может быть проведена пространственная реконструкция полей выпадений, выполнена оценка суммарных выбросов примесей.

Список литературы

[1] Бызова Н.Л., Гаргер Е.К., Иванов В.Н. Экспериментальные исследования атмосферной диффузии и расчеты рассеяния примеси. Л.: Гидрометеоиздат, 1991. 279 с.

[2] Obolkin V.A., Potemkin V.L., Makukhin V.L., Chipanina Y.V., Marinayte I.I. Low-level atmospheric jets as main mechanism of long-range transport of power plant plumes in the Lake Baikal Region // Int. J. Environ. Studies. 2014. Vol. 71 (3). P. 391–397.

[3] Атмосферная турбулентность и моделирование распространения примесей / Под ред. Ф.Т.М. Ньистада и Х. Ван Допа. Л.: Гидрометеоиздат, 1985. 351 с.

[4] Рапута В.Ф., Шлычков В.А., Леженин А.А., Романов А.Н., Ярославцева Т.В. Численный анализ данных аэрозольных выпадений примесей от высотного источника // Оптика атмосферы и океана. 2014. Т. 27. № 8. С. 713–718.

[5] Успенский А.Б., Федоров В.В. Вычислительные аспекты метода наименьших квадратов при анализе и планировании регрессионных экспериментов. М.: Изд-во МГУ, 1975. 168 с.

[6] Берлянд М. Е. Прогноз и регулирование загрязнения атмосферы. Л.: Гидрометеоиздат, 1985. 272 с.

[7] Fallah-Shorshani M., Shekarrizfard M., Hatzopoulou M. Evaluation of regional and local atmospheric dispersion models for the analysis of traffic-related air pollution in urban areas // Atmos. Environ. 2017. Vol. 167. P. 270–282.

[8] Балтер Б.М., Балтер Д.Б., Егоров В.В., Стальная М.В. Использование данных ИСЗ Landsat для определения концентрации загрязнителей в шлейфах от продувки газовых скважин на основании модели источника // Исследование Земли из космоса. 2014. № 2. С. 55–66.

[9] Кондратьев К.Я., Григорьев Ал.А., Покровский О.М., Шалина Е.В. Космическое дистанционное зондирование атмосферного аэрозоля. Л.: Гидрометеоиздат, 1983. 216 с.

[10] Прокачева В.Г., Усачев В.Ф. Снежный покров как индикатор кумулятивного загрязнения в сфере влияния городов и дорог // Метеорология и гидрология. — 2013. — № 3. — С. 94–106.

[11] Balter B.M., Faminskaya M.V. Irregularly emitting air pollution sources: acute health risk assessment using aermod and the Monte Carlo approach to emission rate // Air Quality, Atmosphere and Health. 2017. Vol. 10 (4). P. 401–409.

[12] Lezhenin A.A., Raputa V.F., Yaroslavtseva T.V. Numerical Analysis of Atmospheric Circulation and Pollution Transfer in the Environs of Norilsk Industrial Region // Atmospheric and Oceanic Optics. 2016. Vol. 29 (6). P. 565–569.

Ярославцева Татьяна Владимировна — к.т.н., ст. науч.сотр.
Новосибирского НИИ гигиены Роспотребнадзора;
e-mail: tani-ta@list.ru;
Амикишиева Руслана Александровна — магистрант первого курса ФИТ
Новосибирского государственного университета;
e-mail: ruslana215w@mail.ru.
Рапута Владимир Федотович — д.ф.-м.н., гл. науч. сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: raputa@sscc.ru.
Дата поступления — 30 апреля 2019 г.

Содержание

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СИСТЕМЫ ТЕЛЕУПРАВЛЕНИЯ ЗУР ЗРК МАЛОЙ ДАЛЬНОСТИ
Т. А. Аверина, И. М. Косачев, К. Н. Чугай ... **3**

ПРИМЕНЕНИЕ МЕТОДА МАКСИМАЛЬНОГО СЕЧЕНИЯ В ЗАДАЧАХ ОЦЕНИВАНИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ
Т. А. Аверина, К. А. Рыбаков ... **9**

МОДЕЛИРОВАНИЕ ТУРБУЛЕНТНОГО ТЕЧЕНИЯ НА ОСНОВЕ ГЕНЕРАЦИИ СТОХАСТИЧЕСКОГО ПОЛЯ СКОРОСТЕЙ
А. В. Александров, Л. В. Дородницын, А. П. Дубень ... **14**

ДЕКОМПОЗИЦИЯ И РАСПАРАЛЛЕЛИВАНИЕ ВЫЧИСЛИТЕЛЬНОЁМКИХ ФУНКЦИЙ ОБРАБОТКИ ДАННЫХ МИКРОСЕЙСМИЧЕСКОГО МОНИТОРИНГА
К. С. Алсынбаев ... **20**

О РЕШЕНИИ ОБРАТНЫХ ЗАДАЧ ЧИСЛЕННОЙ ОПТИМИЗАЦИИ КОМПОЗИТНЫХ КОНСТРУКЦИЙ
Е. В. Амелина, С. К. Голушко ... **27**

АСПЕКТЫ ВЫЧИСЛЕНИЯ ДЕТЕРМИНАНТОВ МАТРИЦ И ИХ РАСПРЕДЕЛЕНИЙ ПО ОПРЕДЕЛИТЕЛЯМ
Н. А. Антипин ... **34**

МОДЕЛИРОВАНИЕ ДИНАМИКИ ГАЗОСОДЕРЖАНИЯ НЕФТИ В ПРОЦЕССЕ РАЗРАБОТКИ МЕСТОРОЖДЕНИЙ
А. В. Базовкин ... **38**

УЧЕТ ИНФОРМАЦИИ В ДИСКРЕТНО-НЕПРЕРЫВНЫХ СИСТЕМАХ УРАВНЕНИЙ ТИПА УРЫСОНА
В. А. Белозуб, М. Г. Козлова ... **45**

ОБ АЛГОРИТМАХ ОТЫСКАНИЯ НАИМЕНЬШЕГО РАССТОЯНИЯ МЕЖДУ СИСТЕМАМИ ТОЧЕК В ПРОСТРАНСТВЕ
И. А. Блатов, Е. В. Китаева ... **52**

ОПРЕДЕЛЕНИЕ ИЗМЕНЕНИЯ ПРОСТРАНСТВЕННО-ВРЕМЕННОГО СОСТОЯНИЯ ТЕХНОГЕННЫХ СИСТЕМ МЕТОДАМИ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ
Т. Ю. Бугакова ... **58**

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНОГО «ЯДЕРНОГО» АЛГОРИТМА И МЕТОДА ПОЛИГОНА ЧАСТОТ
Т. Е. Булгакова, А. В. Войтишек ... **65**

АЛГОРИТМЫ МЕТОДА МОНТЕ-КАРЛО ДЛЯ МОДЕЛИРОВАНИЯ АВТОТРАНСПОРТНОГО ПОТОКА СО СЛУЧАЙНЫМИ ПАРАМЕТРАМИ
А. В. Бурмистров, М. А. Коротченко ... **72**

ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ ДЛЯ РЕШЕНИЯ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ, ОПИСЫВАЮЩИХ ПРОЦЕСС ОКИСЛЕНИЯ ИЗОПРОПИЛБЕНЗОЛА
М. К. Вовденко, С. А. Габитов, И. М. Губайдуллин ... **80**

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ БИОМАРКЕРОВ ДЕГЕНЕРАТИВНЫХ ЗАБОЛЕВАНИЙ
О. Ф. Воропаева, К. С. Гаврилова, С. Д. Сенотрусова ... **86**

ОБ ОДНОМ АЛГОРИТМЕ СТАТИСТИЧЕСКОЙ РЕГУЛЯРИЗАЦИИ
Ю. В. Гласко **92**

КОМБИНАТОРНАЯ СТРУКТУРА 5-МЕРНОЙ МОДЕЛИ КОЛЬЦЕВОЙ ГЕННОЙ СЕТИ
В. П. Голубятников, В. С. Градов **99**

О ЦИКЛАХ В БЛОЧНО-ЛИНЕЙНЫХ МОДЕЛЯХ КОЛЬЦЕВЫХ ГЕННЫХ СЕТЕЙ
В. П. Голубятников, Л. С. Минушкина **106**

АЛГОРИТМЫ ПОСТРОЕНИЯ КВАЗИСТРУКТУРИРОВАННЫХ СЕТОЧНЫХ МОДЕЛЕЙ ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА НА ОСНОВЕ КОНЦЕПЦИИ БСМ
В. С. Горшунов, Л. А. Голубева, В. П. Ильин **112**

ПОИСК ЦИКЛОВ В ОДНОЙ n-МЕРНОЙ МОДЕЛИ КОЛЬЦЕВОЙ ГЕННОЙ СЕТИ
В. С. Градов **116**

ФРАКТАЛЬНЫЕ ХАРАКТЕРИСТИКИ ГЕНОМОВ КАК ИНСТРУМЕНТ ВНУТРИ- И МЕЖВИДОВОЙ КЛАССИФИКАЦИИ ОРГАНИЗМОВ
В. Д. Гусев, Л. А. Мирошниченко, Ю. П. Джиоев **122**

РЕШЕНИЕ ЗАДАЧИ ТЕПЛООБМЕНА КРЫЛА САМОЛЁТА МЕТОДОМ МОНТЕ-КАРЛО
С. А. Гусев, В. Н. Николаев **127**

ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЗАДАЧ СО СЛУЧАЙНЫМИ ВХОДНЫМИ ДАННЫМИ
Б. С. Добронец, О. А. Попова **133**

ИСПОЛЬЗОВАНИЕ КОНТРОЛЛЕРА LEAP MOTION ДЛЯ ПРИКЛАДНЫХ СИСТЕМ РАСПОЗНАВАНИЯ ЯЗЫКА ЖЕСТОВ
Д. Г. Еникеев, С. А. Мустафина **140**

АДАПТИВНЫЕ ФОРМУЛЫ ЧИСЛЕННОГО ДИФФЕРЕНЦИРОВАНИЯ ПРИ НАЛИЧИИ ПОГРАНИЧНОГО СЛОЯ
А. И. Задорин, В. П. Ильин **144**

О МОДЕЛЯХ СЕЙСМОГЕНЕРИРУЮЩИХ ГЕОЛОГИЧЕСКИХ СТРУКТУР
М. В. Зарецкая, О. М. Бабешко, В. В. Лозовой **151**

АНАЛИЗ ВЛИЯНИЯ СЛУЧАЙНЫХ ШУМОВ НА СОЛИТОННОЕ РЕШЕНИЕ УРАВНЕНИЯ SIN-ГОРДОНА
А. А. Иванов, Д. Д. Смирнов **156**

BOUNDARY ELEMENT METHOD IN 3D PROBLEMS OF MATHEMATICAL PHYSICS
V. Ivanov **162**

MICROCHANNEL AMPLIFIER BEHAVIOR IN STRONG MAGNETIC FIELD
V. Ivanov **169**

ОБЗОР И АНАЛИЗ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПРОЦЕССА КАТАЛИТИЧЕСКОГО КРЕКИНГА
Г. И. Исламова, И. М. Губайдуллин **175**

DIRECT PROBLEM FOR THE HYPERBOLIC FIRST-ORDER SYSTEM OF ACOUSTIC WAVES PROPAGATION
S. I. Kabanikhin, D. V. Klyuchinskiy, I. M. Kulikov, N. S. Novikov, M. A. Shishlenin **181**

МЕТОД ОПОРНЫХ ВЕКТОРОВ РЕШЕНИЯ ЗАДАЧ ОБ ИСТОЧНИКЕ ДЛЯ ПАРАБОЛИЧЕСКИХ УРАВНЕНИЙ
С. И. Кабанихин, О. И. Криворотько, Е. А. Ануфриенко **188**

АЛГОРИТМ ВЫДЕЛЕНИЯ УГЛОВЫХ СТРУТУР НА ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ МАСШТАБИРУЕМОГО ИЕРАРХИЧЕСКОГО ДЕТЕКТОРА
И. Г. Казанцев, Б. О. Мухаметжанова, К. Т. Искаков **196**

НЕКЛАССИЧЕСКИЕ ЗАДАЧИ В МОДЕЛЯХ ГЛОБАЛЬНОЙ ЭЛЕКТРИЧЕСКОЙ ЦЕПИ
А. В. Калинин, А. А. Тюхтина, С. Р. Лаврова **203**

MATHEMATICAL AND NUMERICAL MODELING OF OIL POLLUTION WASTE PROCESSING
G. B. Kalmenova, G. T. Balakaeva **210**

ЧИСЛЕННОЕ РЕШЕНИЕ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ ТРЕХМЕРНЫХ ЗАДАЧ ДИРИХЛЕ ДЛЯ УРАВНЕНИЯ ГЕЛЬМГОЛЬЦА С ИСПОЛЬЗОВАНИЕМ МОЗАИЧНО-СКЕЛЕТОННОГО МЕТОДА
А. А. Каширин, С. И. Смагин, М. Ю. Тимофеенко **215**

ОЦЕНКА СИСТЕМАТИЧЕСКОЙ ОШИБКИ ДАННЫХ СПУТНИКОВОГО ЗОНДИРОВАНИЯ О ТЕМПЕРАТУРЕ ПОВЕРХНОСТИ ЗЕМЛИ
Н. В. Киланова, Д. Л. Чубаров, Е. Г. Климова **222**

КОМБИНАТОРНАЯ СТРУКТУРА МОДЕЛЕЙ КОЛЬЦЕВЫХ ГЕННЫХ СЕТЕЙ
Н. Е. Кириллова **229**

ПРИМЕНЕНИ АНСАМБЛЕВОГО ПИ-АЛГОРИТМА В ЗАДАЧЕ ОБРАТНОГО МОДЕЛИРОВАНИЯ
Е. Г. Климова **233**

ИСПОЛЬЗОВАНИЕ ЧЕТЫРЕХМЕРНЫХ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ АВТОМАТИЗАЦИИ ПОСТРОЕНИЯ МОДЕЛЕЙ МЕСТНОСТИ
А. А. Колесников, П. М. Кикин **240**

MICOSI: ОТКРЫТЫЙ ПАКЕТ ДЛЯ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ МИТОЗА В СТАДИЯХ ПРОМЕТАФАЗЫ, МЕТАФАЗЫ И АНАФАЗЫ
М. А. Кривов, П. С. Иванов **246**

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТУРБУЛЕНТНЫХ ТЕЧЕНИЙ В УСТОЙЧИВО СТРАТИФИЦИРОВАННОЙ СРЕДЕ: ЦИРКУЛЯЦИЯ НАД ГОРОДСКИМ ОСТРОВОМ ТЕПЛА
Л. И. Курбацкая **253**

ВЕСОВЫЕ ОЦЕНКИ МЕТОДА МОНТЕ-КАРЛО ДЛЯ НЕИНВАЗИВНОГО ИЗМЕРЕНИЯ ОПТИЧЕСКИХ ПАРАМЕТРОВ БИОЛОГИЧЕСКИХ ТКАНЕЙ
А. В. Лаппа, А. Е. Анчугова, Д. Ю. Шакаева **261**

ЭКСПЕРИМЕНТАЛЬНЫЕ И ЧИСЛЕННЫЕ ИССЛЕДОВАНИЯ ПРОЦЕССОВ РАСПРОСТРАНЕНИЯ ПРИМЕСЕЙ ОТ ВЫСОТНЫХ ИСТОЧНИКОВ
А. А. Леженин, В. Ф. Рапута, Р. А. Амикишиева **271**

КРИТЕРИИ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ ПРИ АНАЛИЗЕ БОЛЬШИХ ВЫБОРОК: ПРОБЛЕМЫ И ИХ РЕШЕНИЕ
Б. Ю. Лемешко, С. Б. Лемешко, И. В. Веретельникова, П. Ю. Блинов **277**

СРАВНЕНИЕ РАЗНОСТНЫХ СХЕМ РАЗЛИЧНЫХ ПОРЯДКОВ ТОЧНОСТИ НА АДАПТИВНЫХ СЕТКАХ ДЛЯ РАСЧЕТА ПОГРАНИЧНЫХ И ВНУТРЕННИХ СЛОЕВ
В. Д. Лисейкин, В. И. Паасонен **284**

РЕШЕНИЕ МЕТОДОМ МОНТЕ–КАРЛО НЕЛИНЕЙНОГО УРАВНЕНИЯ ШРЕДИНГЕРА
В. Л. Лукинов **291**

ЧИСЛЕННО УСТОЙЧИВЫЙ ВЕРОЯТНОСТНЫЙ КЛАССИФИКАТОР ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ
В. Л. Лукинов **299**

ОБРАТНАЯ ЗАДАЧА С ФИНАЛЬНЫМ ПЕРЕОПРЕДЕЛЕНИЕМ ДЛЯ СВЕРХУСТОЙЧИВОЙ ГИПЕРБОЛИЧЕСКОЙ СИСТЕМЫ
Н. А. Люлько **305**

АЛГОРИТМЫ ПОСТРОЕНИЯ НАЧАЛЬНОГО ПЛАНА В ЗАДАЧАХ ТРАНСПОРТНОЙ ЛОГИСТИКИ
О. А. Ляхов **312**

МОДЕЛИРОВАНИЕ ПОЛЕЙ ДАВЛЕНИЯ В ТРЕЩИНОВАТЫХ КОЛЛЕКТОРАХ
А. А. Мазитов, Ю. О. Бобренева, И. М. Губайдуллин **318**

ПОДХОД К МОДЕЛИРОВАНИЮ ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМ НА БАЗЕ РАСПРЕДЕЛЕННЫХ СЕТЕЙ ИНТЕЛЛЕКТУАЛЬНЫХ ГЕОСЕНСОРОВ
А. А. Майоров, А. В. Матерухин, О. Г. Гвоздев **323**

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ РАЗМЕЩЕНИЯ СТОКОВ В БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЯХ С НЕНАДЁЖНЫМИ УЗЛАМИ ДЛЯ ПОВЫШЕНИЯ ВЕРОЯТНОСТИ УСПЕШНОГО МОНИТОРИНГА
Д. А. Мигов, К. А. Волжанкина **328**

АЛГОРИТМЫ ВЕРОЯТНОСТНОГО АНАЛИЗА СЕЙСМИЧЕСКОЙ ОПАСНОСТИ
В. А. Миронов, С. А. Перетокин, К. В. Симонов, М. А. Курако **333**

DYNAMIC APPROACH TO CLASSICAL MOMENT PROBLEM
A. S. Mikhaylov, V. S. Mikahaylov **339**

АЛГОРИТМЫ МЕТОДА МОНТЕ–КАРЛО ДЛЯ ИССЛЕДОВАНИЯ ВРЕМЕННОЙ АСИМПТОТИКИ ПОТОКА ЧАСТИЦ С РАЗМНОЖЕНИЕМ В СЛУЧАЙНОЙ СРЕДЕ
Г. А. Михайлов, Г. З. Лотова **346**

МОДЕЛИРОВАНИЕ СТРУКТУРЫ И СТАБИЛЬНОСТИ БИМЕТАЛЛИЧЕСКИХ НАНОВОЛОКОН МЕТОДОМ МОНТЕ-КАРЛО
В. С. Мясниченко, Н. Ю. Сдобняков, Л. Кирилов, Р. Михов **351**

МЕТОДЫ ПРЕДОБУСЛАВЛИВАНИЯ, ОСНОВАННЫЕ НА ГРАФОВЫХ ПРЕОБРАЗОВАНИЯХ
Г. А. Омарова, Д. В. Перевозкин **358**

КЛЕТОЧНО-АВТОМАТНОЕ МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННЫХ ПРОЦЕССОВ НА ТРИАНГУЛЯЦИОННЫХ СЕТКАХ
А. В. Павлова, С. Е. Рубцов, П. Р. Родионов **362**

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ ТРЕХСЛОЙНЫХ РАЗНОСТНЫХ СХЕМ ДЛЯ УРАВНЕНИЯ ШРЕДИНГЕРА С СИММЕТРИЧНЫМИ МНОГОТОЧЕЧНЫМИ ОПЕРАТОРАМИ
В. И. Паасонен **366**

КОМПАКТНЫЕ СХЕМЫ НА РАСШИРЕННЫХ АСИММЕТРИЧНЫХ ШАБЛОНАХ ДЛЯ УРАВНЕНИЯ ШРЕДИНГЕРА
В. И. Паасонен, М. П. Федорук **373**

ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ СХОДИМОСТИ ДВУХСТАДИЙНОЙ ДИСКРЕТНО-АНАЛИТИЧЕСКОЙ ЧИСЛЕННОЙ СХЕМЫ ДЛЯ УРАВНЕНИЙ ТИПА ПРОДУКЦИИ-ДЕСТРУКЦИИ
А. В. Пененко, А. Б. Салимова **380**

ПОСТРОЕНИЕ МОДЕЛИ ИЗМЕНЕНИЯ ВО ВРЕМЕНИ СОДЕРЖАНИЯ ТЕМАТИЧЕСКИХ КЛАСТЕРОВ В КОЛЛЕКЦИЯХ НАУЧНЫХ ТЕКСТОВ
И. С. Пименов, Н. В. Саломатина **385**

ЗАДАЧА ПРОДОЛЖЕНИЯ РЕШЕНИЯ ТРЕХМЕРНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ С ДАННЫМИ НА ЧАСТИ ГРАНИЦЫ
А. Ю. Приходько, М. А. Шишленин **393**

ЗАДАЧА ОЦЕНИВАНИЯ ПОЛЯ КОНЦЕНТРАЦИИ ПРИМЕСИ ОТ ПЛОЩАДНОГО ИСТОЧНИКА
В. Ф. Рапута, Т. В. Ярославцева, Р. А. Амикишиева **400**

ПОСТРОЕНИЕ РЕГУЛЯРИЗИРУЮЩЕГО ОПЕРАТОРА ПРИ ОБРАТНОМ АНАЛИЗЕ ОШИБОК
А. Н. Рогалев **406**

ИСПОЛЬЗОВАНИЕ ЕСТЕСТВЕННОГО ЯЗЫКА ПРИ СОЗДАНИИ ИМИТАЦИОННЫХ МОДЕЛЕЙ
С. В. Рудометов, В. В. Окольнишников **413**

МОДЕЛИРОВАНИЕ РАБОТЫ ОЧИСТНОГО ЗАБОЯ УГОЛЬНОЙ ШАХТЫ
С. В. Рудометов, В. В. Окольнишников, А. А. Ордин **418**

СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ ОПЕРАТОРОВ ДИФФЕРЕНЦИРОВАНИЯ И ИНТЕГРИРОВАНИЯ ОТНОСИТЕЛЬНО ОРТОГОНАЛЬНЫХ ФИНИТНЫХ ФУНКЦИЙ. I. СИММЕТРИЧНЫЕ СПЛАЙНЫ
К. А. Рыбаков, В. В. Рыбин **423**

СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ ОПЕРАТОРОВ ДИФФЕРЕНЦИРОВАНИЯ И ИНТЕГРИРОВАНИЯ ОТНОСИТЕЛЬНО ОРТОГОНАЛЬНЫХ ФИНИТНЫХ ФУНКЦИЙ. II. НЕСИММЕТРИЧНЫЕ СПЛАЙНЫ
К. А. Рыбаков, В. В. Рыбин **431**

АНАЛИТИЧЕСКАЯ И ЧИСЛЕННАЯ РЕГУЛЯРИЗАЦИЯ РЕШЕНИЯ ЗАДАЧИ ИНТЕГРАЛЬНОЙ ГЕОМЕТРИИ НА СЕМЕЙСТВАХ ПАРАБОЛ И ЛОМАНЫХ
А. К. Сеидуллаев, Г. М. Джайков **438**

МАТРИЧНЫЙ АЛГОРИТМ ВЫЧИСЛЕНИЯ ЦЕПНЫХ ДРОБЕЙ
В. А. Селезнев, Л. В. Пехтерева **444**

АЛГОРИТМЫ АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ МЕДИЦИНСКИХ ДАННЫХ
К. В. Симонов, А. Г. Зотин, Ю. А. Хамад, М. А. Курако, Т. В. Черепанова **448**

ОБУСЛОВЛЕННОСТЬ ОБРАТНЫХ ЗАДАЧ И ИТЕРАЦИОННЫЕ МЕТОДЫ ИХ РЕШЕНИЯ
С. Б. Сорокин **454**

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ВОЛНОВЫХ ПРОЦЕССОВ В ГЕОЛОГИЧЕСКИХ СРЕДАХ С ГАЗОВЫМИ КАРМАНАМИ В ЗОНЕ АРКТИЧЕСКОГО ШЕЛЬФА С ПОМОЩЬЮ СЕТОЧНО-ХАРАКТЕРИСТИЧЕСКОГО МЕТОДА
П. В. Стогний, Н. И. Хохлов, И. Б. Петров **461**

ТРЁХМЕРНОЕ МОДЕЛИРОВАНИЕ СИГНАЛОВ ЭЛЕКТРОМАГНИТНОГО ЗОНДА С ТОРОИДАЛЬНЫМИ КАТУШКАМИ
И. В. Суродина **465**

“БУДУЩЕЕ ЗЕМЛИ”: ГЛОБАЛЬНЫЕ ПРОБЛЕМЫ И РАДИАЦИОННОЕ ПОЛЕ, РАДИАЦИОННЫЙ ФОРСИНГ И ДИСТАНЦИОННОЕ ЗОНДИРОВАНИЕ ЗЕМЛИ
Т. А. Сушкевич, С. А. Стрелков, С. В. Максакова **472**

О РЕШЕНИИ ОБРАТНОЙ ГРАНИЧНОЙ ЗАДАЧИ ТЕПЛОПРОВОДНОСТИ ДЛЯ ШАРА, СОСТОЯЩЕГО ИЗ КОМПОЗИТНЫХ МАТЕРИАЛОВ
В. П. Танана, Б. А. Марков **480**

ВОССТАНОВЛЕНИЕ КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ ЗЕМНОЙ ПОВЕРХНОСТИ ПО ДАННЫМ ПРИБОРОВ MODIS, САНГУР, ГСА И МСУ-100
М. В. Тарасенков, В. В. Белов, М. В. Энгель, А. В. Зимовая, Ю.В. Гриднев **486**

ВАРИАЦИОННЫЕ МЕТОДЫ ПОСТРОЕНИЯ РАЗНОСТНЫХ УРАВНЕНИИ И ИТЕРАЦИОННЫХ АЛГОРИТМОВ ДЛЯ ПРИБЛИЖЕННОГО РЕШЕНИЯ ЗАДАЧИ ДИРИХЛЕ
Н. М. Темирбеков, Ж. Р. Жаксылыкова **493**

ДВУХСЕТОЧНЫЙ МЕТОД ДЛЯ НЕЛИНЕЙНОГО СИНГУЛЯРНО ВОЗМУЩЕННОГО УРАВНЕНИЯ ВТОРОГО ПОРЯДКА С ДВУМЯ ПАРАМЕТРАМИ
С. В. Тиховская **500**

ВЛИЯНИЕ САМООРГАНИЗАЦИИ ПОВЕРХНОСТНЫХ ЗАРЯДОВ НА КВАНТОВЫЙ МИКРОКОНТАКТ
О. А. Ткаченко, Д. Г. Бакшеев, О. П. Сушков, В. А. Ткаченко **508**

МОДЕЛИРОВАНИЕ САМООРГАНИЗАЦИИ ЛОКАЛИЗОВАННЫХ ЗАРЯДОВ НА ГРАНИЦЕ ПОЛУПРОВОДНИКА С ПОДЗАТВОРНЫМ ДИЭЛЕКТРИКОМ
О. А. Ткаченко, Д. Г. Бакшеев, В. А. Ткаченко, О. П. Сушков **515**

ЛУЧ СВЕТА В ТУРБУЛЕНТНОЙ СРЕДЕ
В. В. Учайкин **522**

ДРОБНО-ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ В СТАТИСТИКЕ АФТЕРШОКОВ
В. В. Учайкин, Е. В. Кожемякина **529**

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПЕРЕНОСА ОПТИЧЕСКИХ СИГНАЛОВ ЧЕРЕЗ ВЗВОЛНОВАННУЮ ПОВЕРХНОСТЬ
О. С. Ухинова, Б. А. Каргин **537**

МЕТОД КОЛЛОКАЦИИ И НАИМЕНЬШИХ КВАДРАТОВ И ЕГО ПРИЛОЖЕНИЕ К РАСЧЕТУ НАПРЯЖЕННО-ДЕФОРМИРОВАННОГО СОСТОЯНИЯ ПЛАСТИН
В. П. Шапеев, С. К. Голушко, В. А. Беляев, Л. С. Брындин **543**

ОЦЕНКА ВАРИАБЕЛЬНОСТИ РЕШЕНИЯ, ПОЛУЧЕННОГО С ПОМОЩЬЮ ИНТЕРВАЛЬНОЙ РЕГУЛЯРИЗАЦИИ
С. П. Шарый **550**

FORMAL SPECIFICATION AND VERIFICATION OF THE STANDARD MATHEMATICAL FUNCTIONS
N. V. Shilov, I. S. Anureev, E. V. Bodin, D. A. Kondratiev, A. V. Promsky, S. O. Shilova, B. L. Faifel . **557**

ЧИСЛЕННЫЙ АНАЛИЗ СТОХАСТИЧЕСКОГО ОСЦИЛЛЯТОРА ВАН-ДЕР-ПОЛЯ
М. А. Якунин **564**

ЧИСЛЕННЫЙ МЕТОД ПРОГНОЗИРОВАНИЯ ТЕМПЕРАТУРЫ С ПОМОЩЬЮ УРАВНЕНИЯ ВОЛЬТЕРРА
Н. М. Япарова, Т. П. Гаврилова **570**

ЧИСЛЕННАЯ РЕКОНСТРУКЦИЯ ПОЛЕЙ АЭРОЗОЛЬНЫХ ВЫПАДЕНИЙ ПРИМЕСЕЙ ПО ДАННЫМ НАЗЕМНОГО И СПУТНИКОВОГО МОНИТОРИНГА
Т. В. Ярославцева, Р. А. Амикишиева, В. Ф. Рапута **575**