

в) проверить, есть ли полученная лемма w_lem в специализированном словаре;

г) если w_lem есть в специализированном словаре, то поместить полученную лемму слова в текущий терм (term);

д) иначе перейти к шагу 6.

5. Проверить, есть ли полученный терм (term) в частотном словаре:

а) если терм (term) уже есть в частотном словаре, то увеличить частоту встречаемости данного термина на 1;

б) иначе (полученного термина (term) еще нет в частотном словаре) занести терм в частотный словарь с частотой встречаемости, равной 1.

6. Перейти к следующему слову из массива слов ($i = i + 1$).

7. Рассчитать относительную частоту встречаемости для каждого термина из частотного словаря.

8. Отсортировать частотный словарь (либо в алфавитном порядке, либо по частоте встречаемости термов).

9. Сохранить частотный словарь.

10. Завершить работу алгоритма.

Таким образом, в генерируемый частотный словарь попадут только те слова, выражения и формы слов, которые представляют собой термины либо специализированную лексику конкретной предметной области. Это позволит сократить объем частотного словаря, а соответственно, и информационно-терминологического базиса, а также повысить их качество.

В данной работе на основе анализа способов построения частотных словарей авторами предложен модифицированный алгоритм автоматического формирования частотных словарей на основе исходных текстов с применением лемматизации. При использовании данного алгоритма можно обеспечить построение информационно-терминологического базиса

меньшего объема и с сопоставимым уровнем качества. Таким образом, удастся достичь снижения трудоемкости изучения ИТБ и упрощения процесса его формирования.

Библиографические ссылки

1. Огнерубов С. С. Формирование информационно-терминологического базиса в мультилингвистических системах обучения : дис. ... канд. техн. наук. Красноярск, 2008.

2. Ковалев И. В., Карасева М. В., Лесков В. О. Алгоритмизация процедур включения связанных лексем в структуру информационно-терминологического базиса // Программные продукты и системы. 2009. № 4. С. 28–32.

3. Атанов Г. А., Пустынникова И. Н. Структурирование понятий предметной области с помощью методов представления знаний // Искусственный интеллект. 1997. № 2. С. 29–40.

References

1. Ognerubov S. S. *Formirovaniye informatsionno-terminologicheskogo bazisa v mul'tilingvisticheskikh sistemakh obucheniya*. Diss. kand. tekhn. Nauk [Formation of the informational and terminological basis in multilingual training systems. Thesis for candidate's degree]. Krasnoyarsk, 2008.

2. Kovalev I. V., karaseva M. V., Leskov V. O. *Programmnie produkti i sistemi*. 2009, no. 4, p. 28–32.

3. Atanov G. A., Pustinnikova I. N. *Iskusstvennii intellect*. 1997 no 2, p. 29–40.

© Ковалев И. В., Середин А. И., Карасева М. В., Зеленков П. В., Храпунова В. В., 2013

УДК 62.501

О НЕПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ СИСТЕМЫ УПРАВЛЕНИЯ КОНВЕРТЕРНОЙ ПЛАВКОЙ

М. Е. Корнет

Сибирский государственный аэрокосмический университет имени академика М. Ф. Решетнева
Российская Федерация, Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: maria4business@mail.ru

Рассматривается задача непараметрической идентификации дискретно-непрерывного процесса кислородно-конвертерной плавки стали. Отмечаются особенности идентификации в условиях параметрической и непараметрической неопределенности. Подчеркивается целесообразность использования непараметрических методов в условиях малой априорной информации. Построен ряд непараметрических моделей процесса кислородно-конвертерной плавки стали по важнейшим выходным переменным, определяющим химический состав металла на повалке: углероду (C), сере (S), фосфору (P) и марганцу (Mn). Приводятся некоторые результаты моделирования процесса кислородно-конвертерной плавки стали, которые показывают, что процесс ведется в рамках технологического регламента, однако этого не достаточно для качественного управления процессом плавки.

Ключевые слова: идентификация дискретно-непрерывных процессов, непараметрическая регрессия, непараметрическая идентификация кислородно-конвертерной плавки.

ABOUT THE NONPARAMETRIC IDENTIFICATION OF THE OXYGEN-CONVERTER STEELMAKING CONTROL SYSTEM

M. E. Kornet

Siberian State Aerospace University named after academician M. F. Reshetnev
31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660014, Russian Federation
E-mail: maria4business@mail.ru

The paper considers the problem of the nonparametric identification of the discrete and continuous process of the basic oxygen steel melting. The identification peculiarities in conditions of parametric and nonparametric uncertainty are pointed out. The appropriateness of the nonparametric methods use in conditions of small information given a priori is emphasized. Some non-parametrical models of the process of basic oxygen steel melting are built according to the key output variables determining the metal chemical composition at turndown: carbon (C), sulfur (S), phosphorus (P) and manganese (Mn). Some results of the modeling process of the basic oxygen steel melting showed that the process is conducted in the framework of the technological order, however, it is not enough for the qualified control by the melting process.

Keywords: discrete-continuous process identification, nonparametric regression, nonparametric identification for oxygen-converter steelmaking.

Проблема моделирования, идентификации надолго останется одной из центральных проблем теории управления. Постановка задачи идентификации во многом зависит от класса исследуемого процесса (статические, динамические, линейные, нелинейные и др.) и от объема априорной информации. Эта проблема продолжает оставаться актуальной для различных реальных процессов дискретно-непрерывного типа. В докладе приводятся некоторые результаты моделирования процесса выплавки стали.

Постановка задачи. Приведем общепринятую схему исследуемого процесса, принятую в теории идентификации (рис. 1).

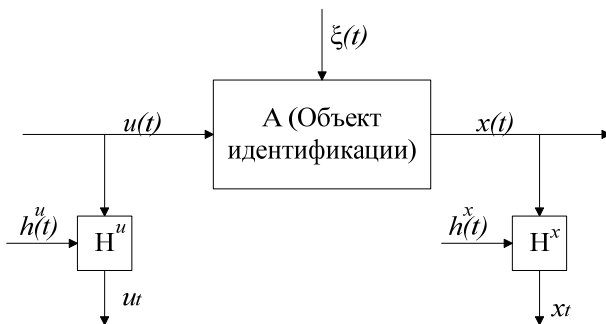


Рис. 1. Общая схема исследуемого процесса и контроля переменных

На рис. 1 приняты обозначения: A – неизвестный с точностью до параметров оператор объекта; $x(t) \in \Omega(x) \subset R^1$ – вектор выходных переменных процесса; $u(t) = (u_1(t), u_2(t), \dots, u_m(t)) \in \Omega(u) \subset R^m$ – вектор входных воздействий; $\xi(t)$ – векторная случайная помеха; (t) – непрерывное время; H^u , H^x – каналы связи, соответствующие различным переменным, включающие в себя средства контроля; $h^u(t)$, $h^x(t)$ – случайные помехи измерений соответствующих пе-

ременных процесса с нулевыми математическими ожиданиями и ограниченной дисперсией. Контроль $u(t)$, $x(t)$ осуществляется через интервал времени Δt . Таким образом, мы имеем выборку входных-выходных переменных $\{x_i, u_i, i = \overline{1, s}\}$, где s – объем выборки.

Итак, стоит задача идентификации, т. е. необходимо восстановить зависимости между входными-выходными характеристиками процесса. В зависимости от уровня априорной информации об объекте различают идентификацию в узком и широком смыслах.

Идентификация в узком и широком смыслах. При моделировании разнообразных дискретно-непрерывных процессов в настоящее время доминирует теория идентификации в «узком» смысле [1]. Ее содержание состоит в том, что на первом этапе на основании имеющейся априорной информации определяется параметрический класс операторов A^α , например:

$$\tilde{x}_\alpha(t) = A^\alpha(u(t), \alpha), \quad (1)$$

где A^α – параметрическая структура модели; α – вектор параметров. На втором этапе осуществляется оценка параметров α на основе имеющейся выборки $\{x_i, u_i, i = \overline{1, s}\}$, где s – объем выборки. Оценка параметров может осуществляться с помощью многочисленных рекуррентных процедур, в частности, методом стохастических аппроксимаций либо методом наименьших квадратов. Успех решения задачи идентификации в этом случае существенно зависит от того, насколько «удачно» определен оператор (1). В настоящее время теория параметрической идентификации является наиболее развитой [1].

Идентификация в широком смысле предполагает отсутствие этапа выбора параметрического класса оператора [2]. Часто оказывается значительно проще определить класс операторов на основе сведений качественного характера, например, линейности про-

цесса или его нелинейности, однозначности либо неоднозначности и др. В этом случае задача идентификации состоит в оценивании такого оператора на основе выборки $\{x_i, u_i, i = \overline{1, s}\}$:

$$\tilde{x}_s(t) = A_s(u(t), \vec{x}_s, \vec{u}_s), \quad (2)$$

где $\vec{x}_s = (x_1, x_2, \dots, x_s)$, $\vec{u}_s = (u_1, u_2, \dots, u_s)$ – временные векторы. Оценка оператора A_s может быть осуществлена средствами непараметрической статистики [2; 3]. Примечательным здесь является то, что при этом исключается этап выбора параметрической структуры. Тем самым можно утверждать, что идентификация в широком смысле является более адекватной для некоторых реальных задач.

Непараметрическая оценка функции регрессии по наблюдениям. Пусть даны наблюдения $\{x_i, u_i, i = \overline{1, s}\}$ случайных величин x, u , распределенных с неизвестными плотностями вероятности $p(x, u)$, $p(u) > 0 \forall u \in \Omega(u)$. Для восстановления $\tilde{x} = M\{x | u\}$ используются непараметрические оценки [2; 3]:

$$x_s(u) = \sum_{i=1}^s x_i \prod_{j=1}^m \Phi(c_s^{-1}(u^j - u_i^j)) / \sum_{i=1}^s \prod_{j=1}^m \Phi(c_s^{-1}(u^j - u_i^j)), \quad (3)$$

где $\Phi(c_s^{-1}(u^j - u_i^j))$, $i = \overline{1, s}$, $j = \overline{1, m}$ – ядерная колоколообразная функция и коэффициент размытости ядра c_s удовлетворяют некоторым условиям сходимости [2; 3].

В данной работе в качестве колоколообразной функции $\Phi(c_s^{-1}(u^j - u_i^j))$ было использовано треугольное ядро вида

$$\Phi(c_s^{-1}(u^j - u_i^j)) = \begin{cases} 1 - |c_s^{-1}(u^j - u_i^j)|, & \text{если } |c_s^{-1}(u^j - u_i^j)| \leq 1, \\ 0, & \text{если } |c_s^{-1}(u^j - u_i^j)| > 1. \end{cases} \quad (4)$$

Параметр размытости c_s определяется путем решения задачи минимизации квадратичного показателя соответствия выхода объекта и выхода модели, основанного на методе скользящего экзамена, когда в модели (3) исключается i -я переменная, предъявляемая для экзамена:

$$R(c_s) = \sum_{k=1}^s (x_k - x_s(u_k, c_s))^2 = \min_{c_s}, \quad k \neq i. \quad (5)$$

Если же входная переменная u представляет собой вектор, то для каждой компоненты u необходимо найти свой параметр размытости c_s . Однако если предварительно привести выборку значений вектора u к одному интервалу, используя операции центрирования и нормирования, то параметр размытости можно принять скалярной величиной.

Примеры непараметрических зависимостей для конвертерной плавки стали. В работе рассматрива-

ется процесс конвертерной выплавки стали [4], который может быть представлен в виде схемы (рис. 2), на которой показаны входные-выходные переменные процесса.

Введем обозначения для входных переменных процесса. Векторная переменная «Расход материала» (т) включает в себя: u_1 – расход чугуна, u_2 – расход лома, u_3 – расход извести, u_4 – расход электродного боя, u_5 – расход флюса ФОМИ, u_6 – расход ФМ-1, u_7 – расход агломерата офлюсованного, u_8 – расход угля ССО. К входным переменным относятся u_9 – кислород на продувку, м³ и u_{10} – кислород на прогрев, м³. Химический состав заливаемого чугуна (%) определяется содержанием следующих элементов: u_{11} – кремний Si, u_{12} – марганец Mn, u_{13} – сера S, u_{14} – фосфор P. Входную переменную «Температура чугуна» (С°) обозначим u_{15} .

К выходным переменным процесса относится векторная переменная «Химический состав металла на повалке» (%), которая включает в себя элементы: x_1 – углерод C, x_2 – марганец Mn, x_3 – сера S, x_4 – фосфор P, x_5 – температура металла на первой повалке, С°, x_6 – температура металла на второй повалке, С°; выходная векторная переменная «Химический состав конечного шлака на повалке» (%), которая включает в себя: x_7 – CaO, x_8 – SiO₂, x_9 – FeO, x_{10} – MgO, x_{11} – Al₂O₃, x_{12} – S, x_{13} – MnO, x_{14} – P₂O₅, x_{15} – TiO₂, x_{16} – V₂O₅.

Важнейшей выходной переменной является векторная переменная «Химический состав металла на повалке» (%), которая включает в себя четыре переменные: $x_1(u)$ (углерод C), $x_2(u)$ (марганец Mn), $x_3(u)$ (S сера), $x_4(u)$ (фосфор P). Для вычислительного эксперимента было построено 4 непараметрических модели вида

$$x_s^l(u) = \sum_{i=1}^{35} x_i \prod_{j=1}^{15} \Phi(c_s^{-1}(u^j - u_i^j)) / \sum_{i=1}^{35} \prod_{j=1}^{15} \Phi(c_s^{-1}(u^j - u_i^j)), \quad (6)$$

где $l = \overline{1, 4}$, объем выборки в данном случае был равен $s = 35$.

Для оценки полученных моделей была использована квадратичная ошибка:

$$R_s = \frac{1}{s} \sum_{i=1}^s (x_i - x_i^s)^2, \quad (7)$$

где x_i – измеренное значение выходной переменной; x_i^s – полученная оценка; R_s – квадратичная ошибка.

Непараметрические модели были построены в режиме «скользящего среднего» для выходных переменных x_1, x_2, x_3, x_4 . Полученные результаты представлены в таблице где показаны: квадратичная ошибка моделирования R_s (7), оценка дисперсии D_s выхода объекта x_i , а также ошибка прогнозирования $W = R_s / D_s$.

Как видно из таблицы, ошибка прогнозирования достаточно велика. Это говорит о том, что полученный прогноз является достаточно грубым.

На рис. 3 показан характер зависимости концентрации углерода C (%) в металле на повалке от содержания лома (т.) в расходных материалах. Содержание лома в данном случае является величиной нормированной. Графики представляют собой срезы при

14 фиксированных входных переменных $u_1, u_3, u_4, \dots, u_{15}$, а переменная u_2 (содержание лома) изменяется в границах $[-2; 2,5]$. Кривая 1 соответствует случаю, когда все фиксированные переменные принимали свое среднее значение. Кривая 2 соответствует сдвигу фиксированной переменной u_1 (чугун) на 0,3 от своего среднего значения, кривая 3 – сдвиг на величину $-0,3$ от среднего значения.

Из приведенного рисунка следует, что характер зависимости нелинейный и достаточно точно описывает канал «расход материала (чугун, лом, известь, электродный бой, флюс, ФМ-1, агломерат офлюсованный, уголь ССО), кислород на продувку, кислород на прогрев, химический состав заливаемого чугуна (кремний Si, марганец Mn, сера S, фосфор P), температура чугуна – углерод (C) в металле на повалке». А это означает, что использование подобных моделей может оказаться полезным для управления процессом конвертерной выплавки стали.

На рис. 4 показан характер зависимости концентрации фосфора P (%) от содержания извести (т) в расходных материалах. Содержание извести здесь

является величиной нормированной. Как и в предыдущем случае, графики представляют собой срезы при 14 фиксированных входных переменных и одной изменяющейся переменной u_2 (содержание извести), которая изменяется в границах $[-2; 2,5]$. Кривая 1 соответствует случаю, когда все фиксированные переменные принимали свое среднее значение. Кривая 2 соответствует сдвигу фиксированной переменной u_2 (лом) на 0,3 от своего среднего значения, кривая 3 – сдвиг на величину $-0,3$ от среднего значения.

Рис. 4 показывает нелинейный характер рассматриваемой зависимости, причем здесь нелинейность уже более сильная, чем в предыдущем случае (см. рис. 3). Зависимость описывает канал «расход материала (чугун, лом, известь, электродный бой, флюс, ФМ-1, агломерат офлюсованный, уголь ССО), кислород на продувку, кислород на прогрев, химический состав заливаемого чугуна (кремний Si, марганец Mn, сера S, фосфор P), температура чугуна – фосфор (P) в металле на повалке». Использование подобных моделей также может оказаться полезным для управления процессом конвертерной выплавки стали.

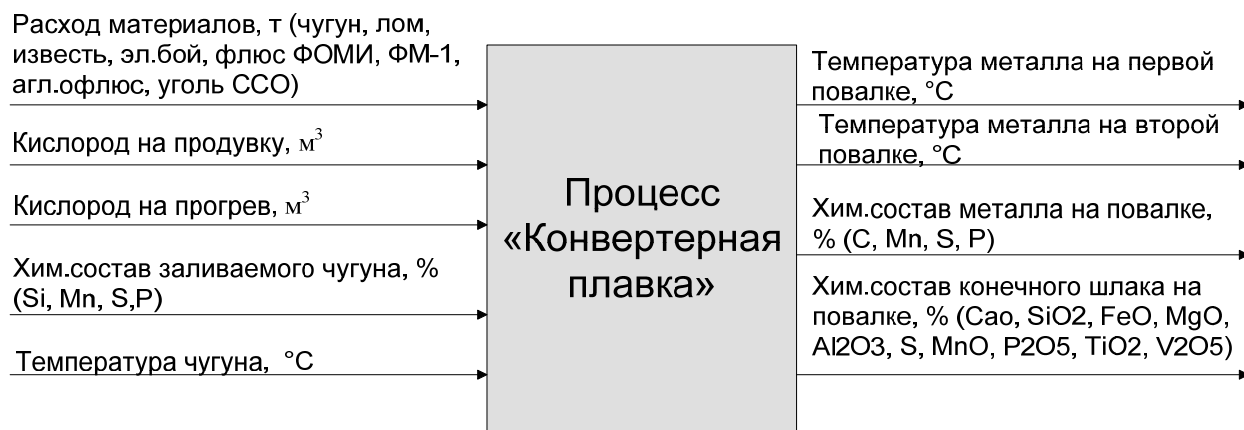


Рис. 2. Входные и выходные переменные процесса конвертерной плавки

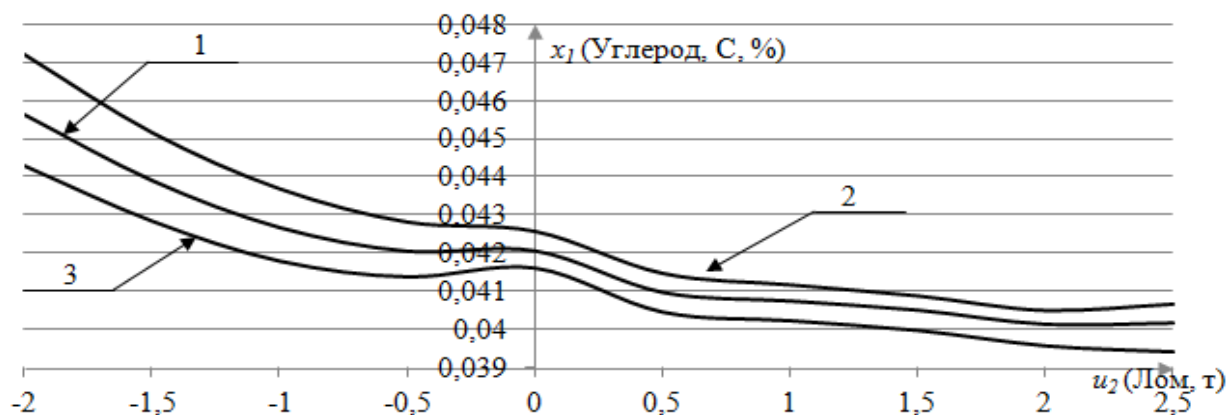


Рис. 3. Характер зависимости концентрации углерода C на повалке от содержания лома (т) в расходных материалах лома

Результаты вычислительного эксперимента

Элемент	Показатель		
	Ошибка R_i	Оценка дисперсии D_s выходной переменной x_i	Отношение W
Углерод, С (x_1)	0,000 197 78	0,000 234	0,845
Марганец, Mn (x_2)	0,000 430 135 8	0,000 540 4	0,85
Сера, S (x_3)	0,000 007 55	0,000 011 83	0,64
Фосфор, P (x_4)	0,000 040 77	0,000 041 137	0,99

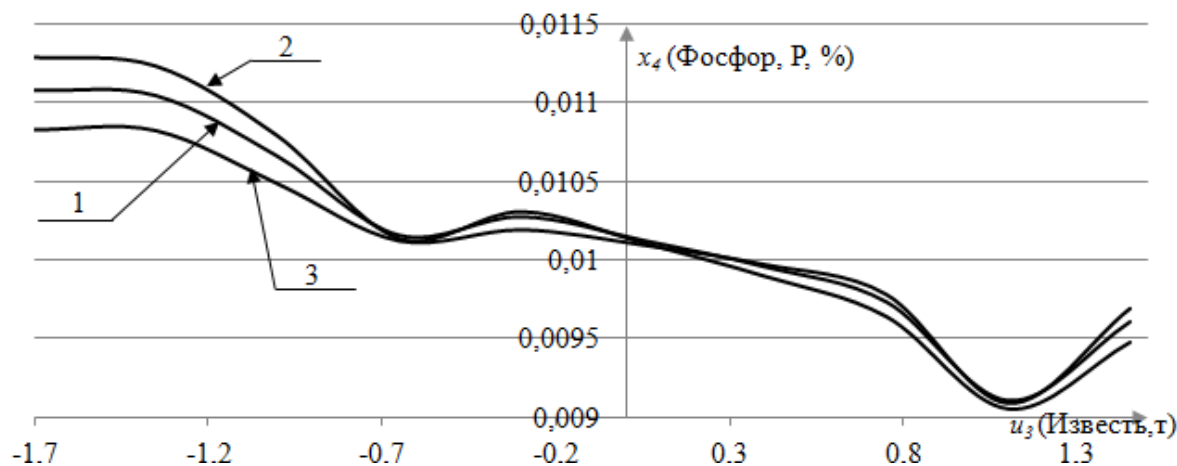


Рис. 4. Характер зависимости концентрации фосфора на повалке от содержания в расходных материалах извести

В статье рассмотрена задача идентификации дискретно-непрерывных процессов в условиях параметрической и непараметрической неопределенности. Анализируется вопрос идентификации в узком и широком смыслах. Приводятся непараметрические модели для многомерных дискретно-непрерывных процессов.

Построены непараметрические модели для следующих важнейших выходных переменных процесса, определяющих содержание основных химических элементов в металле на повалке: углерод (С), марганец (Mn), сера (S), фосфор (P). Приводятся результаты предварительных расчетов. Вычислительные эксперименты показали, что управление процессом кислородно-конвертерной плавки ведется неудовлетворительно, но, тем не менее, соответствует технологическому регламенту. Иными словами, в рамках технологического регламента управляемый процесс представляет собой «облако», т. е. переменные процесса могут принимать произвольные значения из всей области определения переменных, регламентируемой технологической картой. На нынешних предприятиях технологический регламент чаще всего достаточно широк, и если процесс ведется в рамках технологического регламента, но не рациональным образом, то ожидать конечного продукта высокого качества не следует. Полученные непараметрические модели требуют дальнейшего исследования.

Библиографические ссылки

1. Цыпкин Я. З. Адаптация и обучение в автоматических системах. М. : Наука, 1968. 400 с.
2. Надарая Э. А. Непараметрические оценки плотности вероятности и кривой регрессии. Тбилиси : Изд. Тбил. ун-та, 1983. 194 с.
3. Медведев А. В. Непараметрические системы адаптации. Новосибирск : Наука, 1983. 174 с.
4. Воскобойников В. Г. Общая металлургия. М. : Металлургия, 1998. 768 с.

References

1. Tsypkin J. Z. *Adaptacia i obuchenie v avtomaticheskikh sistemah* (The adaptation and learning in automatic systems). Moscow, Nauka, 1968, 400 p.
2. Nadaraya E. A. *Neparametricheskie ocenki plotnosti veroyatnosti i krivoj regressii* (Non-parametric estimation of the probability density and the regression curve). Tbilisi, izd. Tbil. un-ta, 1983, 194 p.
3. Medvedev A. V. *Neparametricheskie sistemy adaptacii* (Nonparametric adaptation systems). Novosibirsk, Nauka, 1983, 174 p.
4. Voskoboynikov V. G. *Obschaya metallurgia* (Overall Metallurgy). Moscow, Metallurgia, 1997. 768 p.

© Корнет М. Е., 2013

УДК 532.516

ПРЕОБРАЗОВАНИЯ ЭКВИВАЛЕНТНОСТИ УРАВНЕНИЙ ДВИЖЕНИЯ ДВУМЕРНОГО СЛОЯ ИДЕАЛЬНОЙ ЖИДКОСТИ*

Д. А. Краснова

Институт вычислительного моделирования СО РАН
Российская Федерация, 660036, Красноярск, Академгородок, 50/44. E-mail: krasnova-d@mail.ru

Исследованы групповые свойства уравнений движения двумерного слоя идеальной жидкости относительно функции, описывающей толщину слоя жидкости под свободной границей. Уравнения записаны в модифицированных переменных, что позволило зафиксировать границы области по новой переменной ξ . Поставлена задача группового анализа, найден продолженный оператор на первые производные, используя критерий инвариантности, построены определяющие уравнения и найдены преобразования эквивалентности для системы уравнений (1)–(4). Преобразования эквивалентности – это такие преобразования, которые сохраняют структуру исходной системы уравнений. Доказано, что преобразования эквивалентности для системы (5)–(8) имеют структуру бесконечномерной группы преобразований. Данная задача имеет прикладное значение для нахождения точных решений систем дифференциальных уравнений вида (1)–(4).

Ключевые слова: групповой анализ, идеальная жидкость.

TRANSFORMATIONS OF THE EQUIVALENCE OF THE MOVEMENT EQUATIONS OF A TWO-DIMENSIONAL LAYER OF IDEAL LIQUID

D. A. Krasnova

Institute computational modeling SB RAS
50/44, Academgorodok, Krasnoyarsk, 660036, Russian Federation. E-mail: krasnova-d@mail.ru

Group properties of the equations of the movement of a two-dimensional layer of ideal liquid concerning the function describing thickness of a layer of liquid under free border are investigated. The equations are written down in the modified variables that allowed to record area borders on a variable. The task of the group analysis is set, the continued operator on the first derivatives is found, using the criterion of invariancy, the defining equations are constructed and equivalence transformations for system of the equations (1)–(4) are found. Transformations of equivalence are such transformations which keep structure of initial system of the equations. It is proved that the equivalence transformations for the system (5)–(8) have the structure of the infinite-dimensional group of transformations. This task has applied the value for finding of exact decisions of systems of the differential equations of a look (1)–(4).

Keywords: Group analysis, ideal liquid.

Описание системы уравнений. Многие исследователи нелинейных уравнений давно используют групповой анализ дифференциальных уравнений. Так как принципы инвариантности закладываются при выводе уравнений, применение группового анализа особенно эффективно и плодотворно в области фундаментальной механики и физики [1–4].

В последнее время особо актуальным стало исследование задач математического моделирования разнообразных физических процессов. Главной целью группового анализа является построение точных решений дифференциальных уравнений. На пути построения этих решений важной частью задачи является нахождение преобразований эквивалентности для заданной системы уравнений. В данной работе получены преобразования эквивалентности для системы уравнений, описывающей двумерное движение иде-

альной жидкости, записанных в модифицированных переменных.

Рассмотрим систему уравнений движения двумерного слоя идеальной жидкости по наклонной плоскости. Уравнения запишем в декартовой системе координат так, чтобы ось z была ортогональна к подложке, а ось x направлена в сторону действия скатывающей силы. Жидкость занимает область $\Omega = \{(x, z) : -\infty < x < +\infty, 0 < z < H(x, t)\}$, где t – время; H – толщина слоя жидкости. В уравнения движения жидкости входят компоненты вектора скорости (u, w) , p – давление (скатывающие силы заменой переменных можно включить в давление). Слой жидкости имеет твердую подложку при $z = 0$ и свободную границу при $z = H(x, t)$ (см. рисунок).

* Работа выполнена при финансовой поддержке РФФИ, грант 11–01–00238.