

УДК 532.517.4;519.237.8

ПРОГНОЗИРОВАНИЕ РАЗВИТИЯ ШИРИНЫ ЗОНЫ ПЕРЕМЕШИВАНИЯ В ЗАДАЧЕ РЭЛЕЯ–ТЕЙЛОРА С ПОМОЩЬЮ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА

А. С. Нужный, П. А. Кучугов, А. Г. Коржов, В. Б. Розанов

В работе проводился статистический анализ расчетов развития неустойчивости Рэля–Тейлора. Целью данного анализа было оценить возможность прогнозирования интегральных характеристик поздних состояний расчета (в работе рассматривается ширина зоны перемешивания Δ) по их начальным состояниям, не прибегая к численному моделированию. Поля распределения плотности в начальные моменты времени были описаны набором параметров $\{\vec{c}^i\}$, после чего в пространстве параметров описания производилась аппроксимация функции $\Delta(\vec{c})$ по точкам. Точность предсказания аппроксимирующей функции оценивалась на тестовых выборках.

В работе изучаются расчеты развития двумерной неустойчивости Рэля–Тейлора (РТ-неустойчивости). Расчеты представлены в виде полей распределения плотности в дискретные моменты времени. Ранее было показано [1–4], что можно подобрать математическое преобразование, отображающее поля плотности в точки многомерного пространства, в котором процессы обнаруживают устойчивое поведение. Устойчивость заключается в том, что если начальные состояния каких-то процессов близки по евклидову расстоянию в этом пространстве, то и последующие состояния этих процессов так же будут близкими. При этом процессы, далеко отстоящие друг от друга в начальный момент времени, в дальнейшем преимущественно таковыми и остаются.

Это свойство дает основание предполагать, что возможно вероятностное предсказание ряда практически интересных характеристик поздних моментов времени по начальному состоянию без использования прямого расчета по уравнениям гидродинамики.

В данной работе исследуется возможность предсказания ширины зоны перемешивания. Для этого начальные конфигурации полей плотности были описаны набором характеристик $\{c^i\}$. Ширина зоны перемешивания Δ рассматривалась как функция от этих характеристик $\Delta(\vec{c})$. N расчетов дают N примеров соответствия точек в пространстве характеристик $\{\vec{c}_n\}_{n=1}^N$ значениям ширины зоны перемешивания $\{\Delta_n, \Delta_n\}_{n=1}^N$. Множество $\{\vec{c}_n, \Delta_n\}_{n=1}^N$ было аппроксимировано с помощью многослойного персептрона [5, 6]. Последующее тестирование на валидационном множестве данных и сравнение его результатов с тестированием случайного предиктора на том же множестве показало, что персептрон выявляет нетривиальные зависимости между шириной зоны перемешивания в поздние моменты времени и начальным состоянием системы.

Описание расчетных данных. В работе использовались двумерные расчеты развития РТ-неустойчивости, полученные с помощью численного кода NUT [7]. Область моделирования (рис. 1) состоит из двух прямоугольных частей, разделенных горизонтальной плоскостью, перпендикулярно которой приложено ускорение $g = 10^4 g_0$, g_0 – ускорение свободного падения вблизи поверхности земли. В верхней части находится тяжелый газ ксенон, в нижней – легкий гелий. Оба вещества в расчетах считаются несжимаемыми. На верхней и нижней границах области заданы условия непротекания, на боковых стенках – периодические граничные условия.

Термодинамические величины газов подчиняются уравнению состояния идеального газа $p = \rho \frac{R}{\mu} T$, $\epsilon = \frac{1}{\gamma - 1} \frac{R}{\mu} T$. Начальные распределения давления и плотности соответствуют изоэнтропической конфигурации. Размер области моделирования перемешивания составляет 72 мм по горизонтали, 150 мм – по вертикали, расчетная сетка 130×250 точек.

Начальное возмущение поверхности раздела сред задается суперпозицией тригонометрических функций:

$$s = \sum_k a_k \cos \left(\frac{2\pi kx}{L} + \varphi_k \right),$$

где $k = 2, 3, 5, 7, 11, 13$, L – ширина фрейма, амплитуды подчиняются условию $a_k \frac{2\pi k}{L} = \frac{1}{2}$, фазы гармоник φ_k могут принимать значения $\frac{\pi}{4}n$, $n = 0, 1 \dots 7$, и выбираются случайным образом.

Расчеты велись до момента времени 1250 мкс. Выдача конфигураций полей плотности производилась через каждые 50 мкс. Пример поля плотности на заключительный момент времени изображен на рис. 2(а). Расстояние между минимумом и максимумом

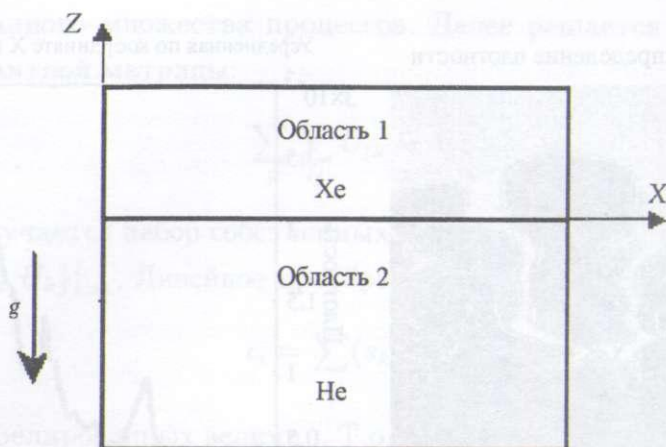


Рис. 1. Область моделирования развития РТ-неустойчивости.

усредненного по X -координате поля плотности будем называть шириной зоны перемешивания Δ рис. 2(б).

Всего в нашем распоряжении было 126 расчетов полей плотности, отличающихся различными начальными возмущениями границы раздела. Поля плотности были зеркально отображены относительно вертикальной оси, что позволило в два раза увеличить количество данных до $N = 252$. Множество расчетов разбивалось на обучающее подмножество, которое использовалось для построения зависимости $\Delta(\vec{c})$, и тестовое.

Предобработка данных. Развитие РТ-неустойчивости целиком определяется начальным возмущением границы раздела, т.к. прочие характеристики от расчета к расчету оставались неизменными. Расчеты велись на сетке, состоящей из 130 узлов в горизонтальном направлении, таким образом, граница раздела представляет собой вектор из $L = 130$ координат (рис. 3), $\vec{s} = \begin{pmatrix} s_1 \\ \dots \\ s_L \end{pmatrix}$. Аппроксимация функции по точкам является обратной задачей, которая, как правило, сводится к оптимизационной, например, к минимизации ошибки аппроксимации. Известно, что при решении оптимизационных задач градиентными методами с увеличением количества свободных параметров увеличивается вероятность “попадания” модели в точку локального экстремума [5]. Учет большего числа компонент входного вектора требует большего количества параметров аппроксимирующей модели, т.о. будем стремиться описать начальные возмущения границ раздела сред по возможности меньшим числом параметров.

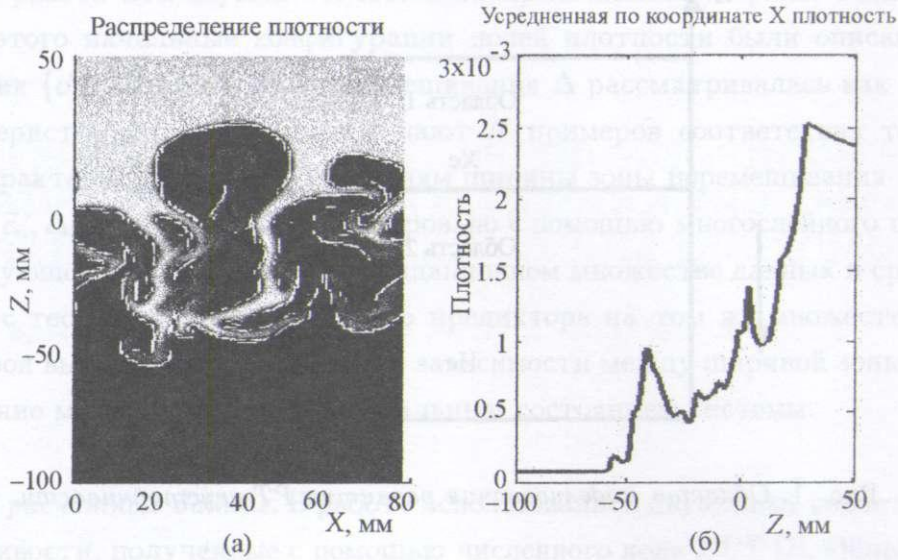


Рис. 2. (а) Поле плотности и (б) усредненное по X-координате поле плотности на момент времени 1250 мкс.

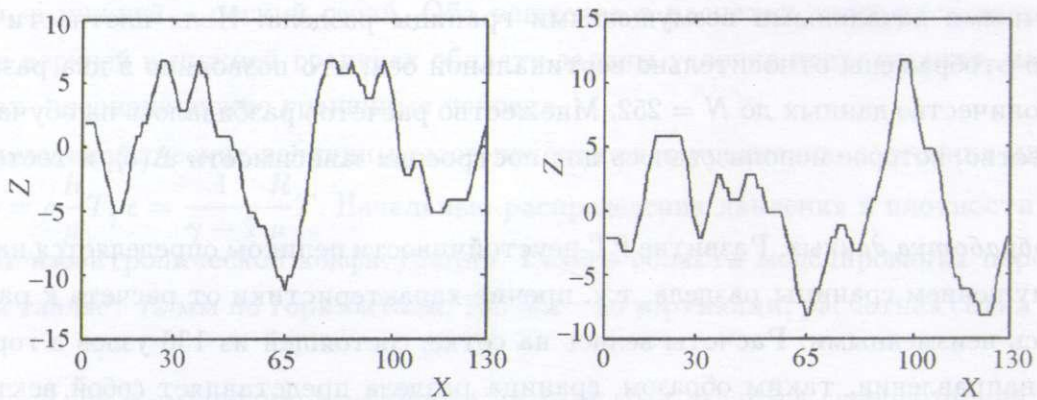


Рис. 3. Примеры границ раздела сред в начальный момент времени.

Оптимальное линейное сжатие, понижающее размерность входного пространства, дает метод главных линейных компонент [5]. Для получения матрицы этого преобразования рассчитаем средний вектор и ковариантную матрицу данных:

$$\bar{s}_i = \frac{1}{N} \sum_{\alpha=1}^N s_i^\alpha, \quad \sum_{ij} \equiv \frac{1}{N-1} \sum_{\alpha=1}^N (s_i^\alpha - \bar{s}_i)(s_j^\alpha - \bar{s}_j),$$

здесь индекс α перебирает все начальные картины распределения плотности процессов,

а N – размер данного множества процессов. Далее решается задача на собственные значения ковариантной матрицы:

$$\sum_j \sum_{ij} U_{jk} = \lambda_k U_{ik}.$$

В результате получается набор собственных значений и соответствующих им собственных векторов $\{\lambda_k, \bar{U}_k\}_{k=1}^L$. Линейное преобразование

$$c_i = \sum_k (s_k - \bar{s}_k) U_{ki}$$

дает набор некоррелированных величин. Т.о. процедура приведения данных в пространство главных компонент требует задания среднего вектора \bar{s} и матрицы поворота \hat{U} . Собственные значения λ_i равны дисперсиям данных по соответствующим осям [5]. Низкодисперсными компонентами можно пренебречь, оставив первые M координат, (где $M \ll L$), при этом основная часть информации будет сохранена. На рис. 4 приведены графики компонент первых восьми собственных векторов.

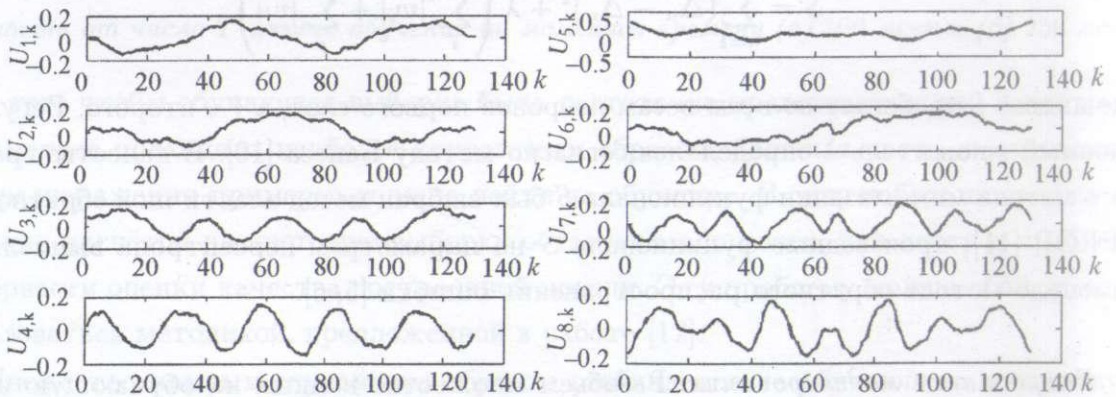


Рис. 4. Собственные вектора матрицы корреляции.

Для выбора достаточного числа компонент M воспользуемся правилом “сломанной трости” (Broken stick model) [8]. Согласно этому правилу собственный вектор остается в списке главных компонент, если соответствующее ему собственное значение удовлетворяет условию:

$$\frac{\lambda_i}{\sum_j \lambda_j} > \frac{1}{L} \sum_{j=i}^L \frac{1}{j}.$$

Аппроксимация ширины зоны перемешивания. В качестве математической модели для аппроксимации ширины зоны перемешивания использовалась двухслойная нейронная сеть – персептрон. Первый слой состоял из $J = 5$ нейронов, второй – из одного. Нейроны, составляющие первый слой, выполняют следующее преобразование над вектором входных параметров: $u^k = f(\vec{c}\vec{w}^k + w_0^k)$. Здесь w_0^k , $\vec{w}^k = (w_1^k, \dots, w_5^k)$ – синаптические веса нейронов, $f(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}$ – функция активации, u^k – выход k -го нейрона первого слоя. Величины $\{u^k\}_{k=1}^J$ играют роль входов для единственного нейрона второго слоя, который имеет линейную функцию активации $\tilde{\Delta} = \vec{u}\vec{v} + v_0$. Величины $v_0, \vec{v} = (v_1, \dots, v_5)$ – его синаптические веса. В качестве выходного значения нейрон второго слоя выдает предсказание ширины зоны перемешивания $\tilde{\Delta}$.

На стадии обучения персептрона веса нейронов изменяются в сторону минимизации ошибки обучения $\sum_{n=1}^N (\tilde{\Delta}_n - \Delta_n)^2$, где N – размер обучающей выборки. Задача поиска экстремума функции многих переменных относится к числу некорректных задач, поэтому требует регуляризации. Согласно общей теории А.Н. Тихонова [9] минимизировался модифицированный функционал вида:

$$S = \sum_{n=1}^N (\tilde{\Delta}_n - \Delta_n)^2 + \lambda \left(\sum_i |w_i| + \sum_j |v_j| \right),$$

где индекс i пробегает по всем весам нейронов первого слоя, а j – второго. Регуляризационный множитель λ определялся согласно методу Байеса [10]. В качестве градиентного метода минимизации функционала S был выбран метод эластичной обратной связи RPROP [11], производные функционала S по параметрам персептрона вычислялись с помощью метода обратного распространения ошибки [5, 6].

Результаты моделирования. Разобьем множество данных на обучающую и тестовую выборки. По начальным возмущениям границ раздела сред $\{\vec{s}_n\}_{n=1}^N$ построим процедуру приведения возмущения в пространство главных компонент (рассчитаем средний вектор \vec{s} и матрицу поворота \hat{U}). Будем обучать персептрон на данных, представленных в пространстве первых $M = 6$ главных компонент. Точность обученного персептрона будем оценивать на тестовой выборке. Ошибку прогнозирования можно

характеризовать величиной $\delta = \sqrt{\frac{\sum_{n=1}^{N_t} (\tilde{\Delta}_n - \Delta_n)^2}{\sum_{n=1}^{N_t} \tilde{\Delta}_n^2}}$, определяемой как отношение между

среднеквадратичными ошибками прогнозирования персептрона и случайного предиктора. Последний в качестве предсказания выдавал случайные числа $\tilde{\Delta}$, распределенные по нормальному закону с параметром $\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^N (\Delta_n - \bar{\Delta}_n)^2}$, где $\bar{\Delta}_n = \frac{1}{N} \sum_{n=1}^N \Delta_n$.

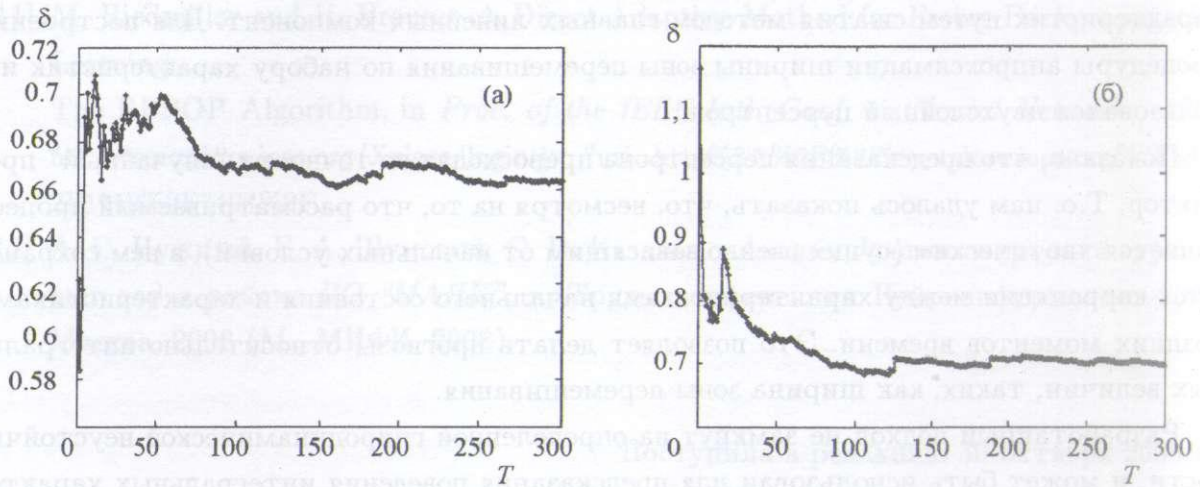


Рис. 5. Среднее отношение δ ошибок предсказания персептрона и случайного предиктора в зависимости от числа T циклов обучения на моменты времени (а) 500 мсек и (б) 250 мсек.

Для того чтобы обучающая выборка была достаточно представительной и настройка параметров проходила наиболее корректным образом, большая часть имеющихся в нашем распоряжении примеров должна пойти на обучение. В силу небольшого количества таких расчётов на тестовую выборку остается мало точек, что ведет к проблеме достоверности оценки качества построенной модели. Эту проблему можно обойти, если воспользоваться методикой, предложенной в работе [12]:

1. Множество данных случайным образом разбивается на обучающую и тестовую выборки.
2. Выполняется процедура построения главных компонент и обучение модели.
3. Проводится тестирование предсказывающей модели, и оцениваются ошибки предсказания персептрона и случайного предиктора.
4. После многократного повторения шагов 1–3 оценивается средняя по числу итераций точность предсказания $\delta = \sqrt{\frac{\sum_{t=1}^T \sum_{n=1}^{N_t} (\tilde{\Delta}_n - \Delta_n)^2}{\sum_{t=1}^T \sum_{n=1}^{N_t} (\tilde{\Delta}_n - \Delta_n)^2}}$, где T – число итераций.

На рис. 5 приведены графики зависимости средней ошибки предсказания от числа циклов T . Из графиков видно, что точность предсказания персептрона превосходит случайный предиктор.

В работе было рассмотрено 126 вариантов развития РТ-неустойчивости. Картины распределения плотности в начальные моменты времени были описаны малым числом характеристик путем сжатия методом главных линейных компонент. Для построения процедуры аппроксимации ширины зоны перемешивания по набору характеристик использовался двухслойный персептрон.

Показано, что предсказания персептрона превосходят по точности "случайный" предиктор. Т.о. нам удалось показать, что, несмотря на то, что рассматриваемый процесс является хаотическим (существенно зависящим от начальных условий), в нем сохраняются корреляции между характеристиками начального состояния и характеристиками поздних моментов времени. Это позволяет делать прогнозы относительно интегральных величин, таких, как ширина зоны перемешивания.

Разработанный подход не замкнут на определенной гидродинамической неустойчивости, и может быть использован для предсказания поведения интегральных характеристик турбулентных течений разных типов.

Работа выполнена при поддержке РФФИ, грант 08-02-01083а.

ЛИТЕРАТУРА

- [1] А. С. Нужный, В. Б. Розанов, Р. В. Степанов, С. А. Шумский, Математическое моделирование, **16**(7), 21 (2004).
- [2] А. С. Нужный, В. Б. Розанов, Р. В. Степанов, С. А. Шумский, Краткие сообщения по физике ФИАН, N 4, 15 (2004).
- [3] А. С. Нужный, В. Б. Розанов, Р. В. Степанов, С. А. Шумский, Краткие сообщения по физике ФИАН, N 2, 41 (2005).
- [4] А. С. Нужный, В. Б. Розанов, Р. В. Степанов, С. А. Шумский, Физика плазмы **31**(3), 342 (2005).
- [5] А. А. Ежов, С. А. Шумский, *Нейрокомпьютинг и его применение в экономике и бизнесе* (М., МИФИ, 1998).
- [6] С. Хайкин, *Нейронные сети: полный курс* (М., Издательский дом Вильямс, 2006).
- [7] В. Ф. Тишкин, В. В. Никишин, И. В. Попов, А. П. Фаворский, Математическое моделирование **7**(5), 15 (1995).

- [8] R. Cangelosi and A. Goriely, *Biology Direct* **2**, 2 (2007).
- [9] А. Н. Тихонов, В. Я. Арсенин, *Методы решения некорректных задач* (М., Наука, 1986).
- [10] С. А. Шумский, *Байесова регуляризация обучения* в: Труды конф. "Современные проблемы нейроинформатики", Москва, 2002 (М., МИФИ, 2002).
- [11] M. Riedmiller and H. Braun, A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm, in *Proc. of the IEEE Intl. Conf. on Neural Networks, 1993*; <http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel3%2F1059%2F7404%2F...>
- [12] А. С. Нужный, Е. А. Шумский, С. В. Казаков, *Анализ динамики уровней грунтовых вод в районе ПО "МАЯК"*, в: Труды конференции Нейроинформатика-2008, Москва, 2008 (М., МИФИ, 2008).

Поступила в редакцию 30 октября 2008 г.