T. M. Pham, Yu. P. Virchenko, Exhaustive study of the noise-induced phase transition in a stochastic model of self-catalyzed reactions, *TMF*, 2016, Volume 188, Number 2, 318–336

DOI: https://doi.org/10.4213/tmf9094

ТЕОРЕТИЧЕСКАЯ
И МАТЕМАТИЧЕСКАЯ
ФИЗИКА
Том 188, № 2
август, 2016

 Т. М. Фам*, **Ю. П. Вирченко***

ПОЛНОЕ ИССЛЕДОВАНИЕ ИНДУЦИРОВАННОГО ШУМОМ ФАЗОВОГО ПЕРЕХОДА В СТОХАСТИЧЕСКОЙ МОДЕЛИ АВТОКАТАЛИТИЧЕСКИХ РЕАКЦИЙ

Проведено полное исследование стационарной плотности распределения в пространстве относительных концентраций для трехпараметрической стохастической модели Хорстхемке–Лефевера бинарной автокаталитической циклической химической реакции, учитывающей возмущения, вызванные тепловыми флуктуациями реагентов. Эта модель представляет собой стационарный диффузионный случайный процесс, порождаемый стохастическим уравнением с дифференциалом Стратоновича, маргинальная плотность распределения которого допускает бифуркационную перестройку от унимодальной к бимодальной при увеличении интенсивности шума, что физически интерпретируется как динамический фазовый переход, индуцированный флуктуациями в системе.

Ключевые слова: бимодальное распределение, бифуркация, критическая поверхность, стехиометрические коэффициенты, стохастическое дифференциальное уравнение, диффузионный марковский процесс, уравнение Фоккера–Планка, уравнения химической кинетики, фазовая диаграмма, фазовый переход, индуцированный шумом, флуктуации.

DOI: 10.4929/tmf9094

1. ВВЕДЕНИЕ

При теоретическом изучении различных явлений в естественных науках возникают математические модели, которые связаны со стохастическими динамическими системами. Их формулировка и исследование основаны на понятии стохастического дифференциального уравнения и привлечении общей теории таких уравнений. Одной из таких стохастических моделей является так называемая *генетическая модель*, введенная в книге [1] как модель, иллюстрирующая эволюцию со временем в некоторых биологических процессах. В работе [2] было предложено применение этой модели для описания кинетики бинарных циклических химических реакций при наличии катализаторов (см. также статью [3], где проведен более детальный

*Национальный исследовательский университет “Белгородский государственный университет”, Белгород, Россия. E-mail: virch@bsu.edu.ru

вывод уравнений модели на основе химической кинетики). Там же был дан анализ стационарного решения модели в частном симметричном случае, результаты которого приведены в монографии [4]. В ней также была проанализирована связь между моделью авторов и моделью, рассмотренной в книге [1].

В динамике, описываемой генетической моделью, проявляется так называемый *индуцированный шумом фазовый переход* при изменении ее свободных параметров. С математической точки зрения он представляет собой бифуркационную перестройку стационарной плотности распределения случайной величины $\tilde{x}(t)$ – значения в момент времени t случайного процесса, который определяется моделью. Причем такая перестройка отсутствует в детерминированном пределе модели при равной нулю интенсивности шума – параметра, характеризующего влияние стохастического слагаемого в соответствующем стохастическом дифференциальном уравнении. Именно это обусловило интерес к исследованию генетической модели. Дополнительным обстоятельством, привлекающим внимание к изучению этой модели, является экспериментальное подтверждение наличия указанного фазового перехода [5].

Бифуркация, свойственная генетической модели, представляет собой частный случай фазовых переходов под воздействием шума. Начало их интенсивному математическому исследованию было положено в 70-х годах прошлого столетия, и до настоящего времени эта тематика исследований представляет интерес [6] как с точки зрения математической физики, так и с точки зрения приложения результатов этих исследований к конкретным физическим ситуациям. Следует отметить, что успехи в исследовании фазовых переходов под воздействием шума в основном связаны с изучением одномерных динамических систем.

Математическое исследование генетической модели давалось в работах ее основоположников в различные годы (см., например, обзоры [7]–[9] и второе издание уже упомянутой монографии [10]). Однако в их работах не было дано полного аналитического исследования стационарного состояния генетической модели. При исследовании стационарных состояний для наборов значений параметров модели в общем положении в этих работах авторы переходили к численному моделированию.

В настоящей работе мы приводим результаты полного исследования стационарного состояния генетической модели при всех допустимых значениях ее параметров, предварительно опубликованные в работах [3], [11]–[13]. В разделе 2 мы кратко описываем конструкцию модели Хорстхемке–Лефевера и необходимые для дальнейшего изложения связанные с ней результаты. В разделе 3 ставится задача вычисления критической поверхности в пространстве параметров. В разделе 4 проводится полное аналитическое исследование критической поверхности. В разделе 5 критическая поверхность исследуется вблизи граничных значений параметра $\alpha = 0$ и $\alpha = 1$, в которых она теряет смысл.

2. КОНСТРУКЦИЯ МОДЕЛИ

Рассмотрим связанные пары химических реакций, которые осуществляются по следующей схеме:

$$A + X + Y \quad \underset{k_1}{\overset{k_2}{\leftrightarrows}} \quad 2Y + A^*, \qquad B + X + Y \quad \underset{k_3}{\overset{k_4}{\leftrightarrows}} \quad 2X + B^*,$$

где X, Y, A, B, A^*, B^* – символы химических реагентов, и при этом вещества, обозначаемые символами A, B, A^*, B^*, выполняют роль химической среды, в которой возможно протекание прямой и обратной реакции со сравнимыми друг с другом скоростями k_i, $i = 1, 2, 3, 4$. На основании базовых уравнений химической кинетики, описывающих динамику этой пары одновременно протекающих реакций, имеем, что[1)]

$$
\begin{aligned}
\dot{N}_t(X) &= k_2 N_t^2(Y) N_t(A^*) - k_1 N_t(X) N_t(Y) N_t(A) + k_3 N_t(X) N_t(Y) N_t(B) - \\
&\quad - k_4 N_t^2(X) N_t(B^*), \\
\dot{N}_t(Y) &= k_1 N_t(X) N_t(Y) N_t(A) - k_2 N_t^2(Y) N_t(A^*) + k_4 N_t^2(X) N_t(B^*) - \\
&\quad - k_3 N_t(X) N_t(Y) N_t(B),
\end{aligned}
$$

где $N_t(A)$, $N_t(A^*)$, $N_t(B)$, $N_t(B^*)$, $N_t(X)$, $N_t(Y)$ – зависящие от времени t числа частиц соответствующих реагентов. Из этой системы уравнений следует закон сохранения суммарного числа молекул обоих реагентов в каждом физически малом объеме термодинамической системы, так как сумма двух уравнений приводит к равенству $d(N_t(X) + N_t(Y))/dt = 0$. Тогда $N_t(X) + N_t(Y) = N = \text{const}$.

Обозначим через $x(t) = N_t(X)/N$, $1 - x(t) = N_t(Y)/N$ концентрации частиц соответственно реагентов X и Y в момент времени t. Пренебрегая малыми изменениями со временем величин $N_t(A)$, $N_t(A^*)$, $N_t(B)$, $N_t(B^*)$ по сравнению с самими этими величинами, т. е. считая, что они не зависят от t и при этом значения $N(A)$, $N(A^*)$, $N(B)$, $N(B^*)$ имеют один и тот же порядок величины, намного превосходящий числа $N_t(X)$ и $N_t(Y)$, перейдем к другому масштабу времени в кинетических уравнениях посредством замены $N[k_2 N(A^*) + k_4 N(B^*)]t$ на физически безразмерный параметр t. Тогда получается следующее уравнение для концентрации $x(t)$:

$$
\dot{x}(t) = \alpha - x(t) + \lambda x(t)\bigl(1 - x(t)\bigr), \qquad x(t) \in [0, 1], \tag{1}
$$

с безразмерными коэффициентами

$$
\alpha = \frac{k_2 N(A^*)}{k_2 N(A^*) + k_4 N(B^*)}, \qquad \lambda = \frac{k_3 N(B) + k_4 N(B^*) - k_1 N(A) - k_2 N(A^*)}{k_2 N(A^*) + k_4 N(B^*)}, \tag{2}
$$

$\alpha \in [0, 1]$, $\lambda \in \mathbb{R}$, которые являются характеристиками реакции. Уравнение (1) имеет устойчивую стационарную точку $\bar{x} = \bigl(\lambda - 1 + \sqrt{(\lambda - 1)^2 + 4\lambda\alpha}\bigr)/2\lambda$ внутри отрезка $[0, 1]$, к которой стремится любое решение с начальным значением $x_0 \in (0, 1)$. Значения $\alpha = 0, 1$ являются особыми, так как для них модель теряет свой физический смысл. Наличие одной устойчивой точки равновесия указывает на то, что в детерминированном случае модель (1) не допускает качественных изменений динамики при изменении ее параметров.

При учете термодинамических случайных флуктуаций чисел $N_t(A)$, $N_t(B)$ детерминированную модель (1) необходимо заменить на стохастическую с помощью аддитивных случайных возмущений параметров модели в виде стационарных эргодических случайных процессов. В стохастической модели Хорстхемке–Лефевера такое возмущение в виде белого шума $\sigma^2 \widetilde{\varphi}(t)$ вводится только для параметра λ:

[1)] По поводу методов построения уравнений химической кинетики см., например, статью [14].

$\lambda \Rightarrow \lambda + \sigma^2\widetilde{\varphi}(t)$, $\langle\widetilde{\varphi}(t)\rangle = 0$, $\langle\widetilde{\varphi}(t)\widetilde{\varphi}(0)\rangle = \delta(t)$, где здесь и далее знаком “тильда” отмечаются случайные величины, а угловыми скобками обозначены математические ожидания. Вводя стохастический дифференциал $d\widetilde{w}(t) = \widetilde{\varphi}(t)\,dt$, где $\widetilde{w}(t)$, $t \in \mathbb{R}_+$, – винеровский случайный процесс, приходим к генетической модели в виде стохастического дифференциального уравнения

$$d\tilde{x}(t) = \bigl[\alpha - \tilde{x}(t) + \lambda\tilde{x}(t)\bigl(1 - \tilde{x}(t)\bigr)\bigr]\,dt + \sigma\tilde{x}(t)(1 - \tilde{x}(t))\,d\widetilde{w}(t), \tag{3}$$

определяющего марковский диффузионный случайный процесс $\tilde{x}(t)$, $t \in \mathbb{R}_+$.

Для дифференциала $d\widetilde{w}(t)$ в уравнении (3) в зависимости от предназначения стохастической системы используются различные определения (см. по этому поводу книгу [15]). Для построения стохастических моделей физических систем естественно использовать уравнения, в которых дифференциал $d\widetilde{w}(t)$ понимается по Стратоновичу [16] в отличие от классического подхода на основе стохастического дифференциала Ито [17]. Вопросу обоснования этого положения посвящена обширная литература как теоретического характера (см., например, работы [18], [19]), основанная на теоремах приближения решений дифференциальных уравнений со случайными коэффициентами [20], [21], так и экспериментального характера, где сравнивались предсказания родственных стохастических моделей в конкретной физической ситуации, основанные на различных стохастических дифференциалах [22].

Известно, что совокупность случайных реализаций – решений стохастического дифференциального уравнения (3) составляет марковский диффузионный случайный процесс с траекториями, непрерывными с вероятностью единица. Этот факт является основным положением теории уравнений с дифференциалом Ито (см., например, книгу [23]). В случае уравнений с дифференциалом Стратоновича данный факт устанавливается на основе однозначной связи между этими дифференциалами (см., например, книгу [15]). Поэтому для плотности $p(x,t) = d\Pr\{\tilde{x}(t) < x\}/dx = \langle\delta(\tilde{x}(t) - x)\rangle$ маргинального распределения первого порядка этого процесса справедливо уравнение Фоккера–Планка[2)]

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial[f(x)p(x,t)]}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2[g^2(x)p(x,t)]}{\partial x^2} \equiv (\mathsf{H}p)(x,t), \tag{4}$$

$$f(x) = \alpha - x + \lambda x(1-x) + \frac{\sigma^2}{2}x(1-x)(1-2x), \qquad g(x) = x(1-x). \tag{5}$$

Для любого случайного значения $\tilde{x}(0) \in (0,1)$, статистически независимого от значений винеровского процесса $\widetilde{w}(t)$, $t \in \mathbb{R}_+$, уравнение (3) имеет единственное с точностью до стохастической эквивалентности решение, которое с вероятностью единица содержится в $(0,1)$ при всех $t \in \mathbb{R}_+$. Этот факт можно доказать на основе методов общей теории стохастических дифференциальных уравнений (см. книгу [23]). Более прозрачное доказательство строится (см. работу [25]) на основе представления белого шума в виде предела при $m \to \infty$ от последовательности $\{\widetilde{\varphi}^{(m)}(t), m \in \mathbb{N}\}$ *импульсных* процессов с траекториями

$$\tilde{\varphi}^{(m)}(t) = \sum_{n\in\mathbb{Z}} \widetilde{\alpha}_n^{(m)} u^{(m)}(t - \tilde{t}_n^{(m)}). \tag{6}$$

2) Заметим, что к такому же уравнению можно прийти посредством техники приближений, которая разрабатывалась в рамках общего подхода для эволюционных задач статистической физики (см., например, статью [24]).

Здесь $\{u^{(m)}(\cdot), m \in \mathbb{N}\}$ есть последовательность финитных локализованных около нуля гладких функций, которая стремится в слабом смысле к $\delta(t)$ при $m \to \infty$, $\{\{\widetilde{\alpha}_n^{(m)}, n \in \mathbb{Z}\}, m \in \mathbb{N}\}$ – последовательность одинаковых дихотомических независимых в совокупности случайных величин $\widetilde{\alpha}_n^{(m)}$, $n \in \mathbb{Z}$, с нулевым средним значением и таких, что $\widetilde{\alpha}_n^{(m)} \in \{\pm a^{(m)}\}$ при $n \in \mathbb{Z}$ и $a^{(m)} \to 0$ при $m \to \infty$. Последовательность $\{\{\tilde{t}_n^{(m)}, n \in \mathbb{Z}\}, m \in \mathbb{N}\}$ состоит из простейших пуассоновских случайных потоков $\tilde{t}_n^{(m)}$, $n \in \mathbb{Z}$, с плотностями $\rho_m = (a^{(m)})^{-2}$ таких, что при каждом фиксированном $m \in \mathbb{N}$ поток статистически независим от последовательности случайных величин $\{\widetilde{\alpha}_n^{(m)}, n \in \mathbb{N}\}$.

Известно, что ряд (6) сходится для каждого $m \in \mathbb{N}$ с вероятностью единица (см. статью [26]), что устанавливается посредством применения леммы Бореля–Кантелли. Указанное выше свойство решений уравнения (3) следует из того, что этим свойством обладают решения $\tilde{x}^{(m)}(t)$ дифференциального уравнения со случайными коэффициентами

$$\dot{\tilde{x}}^{(m)}(t) = \bigl[\alpha - \tilde{x}(t) + \lambda \tilde{x}(t)\bigl(1 - \tilde{x}(t)\bigr)\bigr] + \sigma \tilde{x}(t)\bigl(1 - \tilde{x}(t)\bigr)\widetilde{\varphi}^{(m)}(t)$$

при каждом фиксированном m. Применяя теорему Вонга–Закаи [20] к пределам $\tilde{x}(t)$ последовательности решений $\langle \tilde{x}^{(m)}(t), m \in \mathbb{N}\rangle$, что допустимо, так как последовательность случайных процессов $\{\widetilde{w}^{(m)}(t), m \in \mathbb{N}\}$ с траекториями

$$\widetilde{w}^{(m)}(t) = \int_0^t \widetilde{\varphi}^{(m)}(s)\, ds$$

поточечно стремится к стандартному винеровскому процессу при $m \to \infty$, получим, что предельные траектории $\tilde{x}(t)$ с вероятностью единица также полностью расположены в интервале $(0, 1)$.

Ввиду того, что траектории $\tilde{x}(t)$ диффузионного процесса полностью расположены в $(0, 1)$ при $\tilde{x}(0) \in (0, 1)$, носитель каждого решения $p(x, t)$ уравнения (4) с начальной плотностью распределения $p(x, 0)$ такой, что $\mathrm{supp}[p(x, 0)] \subset [0, 1]$, совпадает с $[0, 1]$. По этой причине для данного начального распределения соответствующее решение $p(x, t)$ удовлетворяет граничному условию равенства нулю потока вероятности

$$J[p(x,t)] = f(x)p(x,t) - \frac{\sigma^2}{2}\frac{\partial[g^2(x)p(x,t)]}{\partial x}$$

в *естественных* (в смысле книги [23]) граничных точках[3)] $x = 0, 1$.

Довольно просто можно найти стационарное решение $p(x)$ уравнения Фоккера–Планка (4), которое имеет вид $J[p(x)] = 0$ при естественных граничных условиях. Оно существует и единственно для каждого набора значений параметров $\alpha \in (0, 1)$, $\lambda \in \mathbb{R}$, $\sigma^2 > 0$ и представляется формулами (см. работы [3], [4])

$$p(x) = \frac{A}{x(1-x)}\left(\frac{x}{1-x}\right)^{\beta} \exp\left\{\frac{2}{\sigma^2}\left(\frac{\alpha - 1}{1-x} - \frac{\alpha}{x}\right)\right\}, \qquad \beta = \frac{2(2\alpha + \lambda - 1)}{\sigma^2}, \tag{7}$$

где постоянная A находится из условия $\int_0^1 p(x)\, dx = 1$,

$$A = \frac{1}{2}\exp\left\{\frac{2}{\sigma^2} + \beta \ln\sqrt{\frac{1-\alpha}{\alpha}}\right\}\left[K_{-\beta}\left(\frac{4}{\sigma^2}\sqrt{\alpha(1-\alpha)}\right)\right]^{-1},$$

[3)] Это свойство используется без обоснования в монографии [4].

$K_{-\beta}(\cdot)$ – модифицированная функция Бесселя второго рода с показателем $-\beta$, которая для любого показателя $\nu \in \mathbb{C}$ и положительного x определяется интегральным представлением (см. книгу [27])

$$K_\nu(x) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-x\,\mathrm{ch}\,u - \nu u}\,du, \qquad \mathrm{Re}\,x > 0.$$

Плотность распределения $p(x)$ теряет смысл при $\alpha = 0$, так как она не интегрируема в окрестности точки $x = 0$. По той же причине она теряет смысл при $\alpha = 1$, когда она не интегрируема в окрестности точки $x = 1$. При $\alpha \in (0, 1)$ имеет место равенство $p(0) = p(1) = 0$.

Диффузионный процесс $\tilde{x}(t)$, $t \in \mathbb{R}_+$, характеризуется тем, что границы отрезка $[0, 1]$, внутри которого расположены его траектории, не являются естественными в смысле Феллера (определение см., например, в монографии [4]). Это связано с тем, что критерием естественности границ по Феллеру является расходимость двух интегралов – характеристик процесса вблизи границ, которые в нашем случае имеют вид

$$\int p(x)g^2(x)\biggl(\int_{x'}^{x} p(y)\,dy\biggr)dx, \qquad \int p(x)\biggl(\int_{x'}^{x} p(y)g^2(y)\,dy\biggr)dx$$

и поэтому заведомо сходятся.

Таким образом, для процесса $\tilde{x}(t)$, $t \in \mathbb{R}_+$, применима теорема Эллиотта (см. работу [28]), на основании которой можно утверждать, что *спектр* $\{-\mu_m\}$ *дифференциального оператора* H *с граничными условиями* $(J[p(x,t)])_{x=0,1} = 0$ *является чисто дискретным и* $\mu_m \geqslant 0$, *а соответствующие собственные функции* $\psi_m(x)$ *образуют полную систему в пространстве* $\mathbb{L}_1(0,1)$.

Указанное свойство оператора H позволяет утверждать, что случайный процесс $\tilde{x}(t)$, $t \in \mathbb{R}$, имеет единственную финальную плотность распределения, которая совпадает с единственной собственной функцией оператора H с нулевым собственным значением. Это означает, что для любой начальной плотности $p(x,0)$ с носителем, сосредоточенным на $[0,1]$, и удовлетворяющей граничному условию $(J[p(x,0)])_{x=0,1} = 0$, соответствующее решение $p(x,t)$ уравнения (4) стремится к стационарной плотности $p(x)$ при $t \to \infty$. Более того, можно утверждать, что статистические характеристики случайного процесса $\tilde{x}(t)$, $t \in \mathbb{R}_+$, стремятся к соответствующим статистическим характеристикам диффузионного стационарного эргодического процесса $\tilde{x}_\infty(t)$, $t \in \mathbb{R}$, с маргинальной плотностью распределения первого порядка $p(x)$ и условной вероятностью перехода $p(x,t;y,s)$, удовлетворяющей уравнению (4) при $t \geqslant s$ и начальному условию $p(x,s;y,s) = \delta(x-y)$.

3. КРИТИЧЕСКАЯ ПОВЕРХНОСТЬ

Качественное устройство плотности $p(x)$, а именно число ее мод, характеризуется разбиением пространства наборов параметров $(\lambda, \sigma^2, \alpha)$ на области таким образом, что эта плотность имеет фиксированное число точек максимума (модальность плотности распределения) в каждой из этих областей. Такое разбиение по аналогии с термодинамикой будем называть фазовой диаграммой системы, а поверхность Σ, которая разделяет эти области, – *критической поверхностью*. Дальнейшее содержание статьи посвящено исследованию этой поверхности. Изменению модальности плотности $p(x)$ соответствует изменение числа решений уравнения $dp(x)/dx = 0$ при изменении параметров системы, т. е. такая бифуркация $p(x)$ связана с вырождением решений этого уравнения, которое приводится к виду

$$S(x) \equiv \alpha - x + \lambda x(1-x) - \frac{\sigma^2 x}{2}(1-x)(1-2x) = 0, \qquad x \in (0,1). \tag{8}$$

Качественный анализ критической поверхности при произвольных значениях параметров λ и α, который дается ниже в этом и следующем разделах, отсутствует в предыдущих публикациях, посвященных генетической модели.

Замечание 1. Кроме решений уравнения (8) формально условию наличия бифуркации плотности $p(x)$ удовлетворяют точки $x = 0$ и $x = 1$, так как в них верны равенства $p'(0) = p''(0) = p'(1) = p''(1) = 0$ при любых значениях параметров α, λ и σ^2, кроме $\sigma^2 = 0$ (когда плотность $p(x)$ не существует). Однако при фиксированном α значения параметров λ_{c} и σ^2_{c}, при которых могут возникнуть дополнительные экстремумы плотности в точках $x = 0, 1$, отсутствуют. В самом деле, если бы существовала экстремальная точка $x_{\mathrm{c}}(\lambda, \sigma^2)$, находящаяся внутри $(0,1)$ и такая, что $x_{\mathrm{c}}(\lambda, \sigma^2) \to 0$ либо $x_{\mathrm{c}}(\lambda, \sigma^2) \to 1$ при $\lambda \to \lambda_{\mathrm{c}}$ и $\sigma^2 \to \sigma^2_{\mathrm{c}}$ независимо от направления перехода к пределу в полуплоскости $(\lambda, \sigma^2 > 0)$, то в этой точке выполнялось бы равенство $S(x_{\mathrm{c}}(\lambda, \sigma^2)) = 0$. Но это невозможно, так как при $\lambda \to \lambda_{\mathrm{c}}$, $\sigma^2 \to \sigma^2_{\mathrm{c}}$ в последнем равенстве либо $S(0) = \alpha \neq 0$, либо $S(1) = \alpha - 1 \neq 0$.

Уравнение (8) может иметь либо одно, либо три вещественных решения. Тем его решениям, которые расположены на $(0,1)$, соответствуют экстремумы плотности $p(x)$. Один вещественный корень всегда находится внутри $(0,1)$, так как $S(1)S(0) < 0$, и поэтому $p(x)$ имеет один экстремум внутри интервала. Тогда в силу равенства $p(0) = p(1) = 0$ при наличии трех экстремумов, два из которых являются максимумами, а один – минимумом между ними, имеется три вещественных корня уравнения (8) внутри $(0,1)$.

Ввиду замечания 1 кратность решений уравнения $dp/dx = 0$ внутри интервала $(0,1)$ эквивалентна кратности корней уравнения $S(x) = 0$. Анализ существования кратного корня $x_0 \in \mathbb{R}$ у полинома $S(x)$ основан на том, что для него наряду с равенством $S(x_0) = 0$ должно выполняться равенство $S'(x_0) = 0$. Тогда условие существования кратного корня у полинома $S(x)$ в зависимости от его параметров следует из равенства нулю зависящего от параметров остатка, который получается в результате применения алгоритма Евклида к паре полиномов $S(x)$ и $S'(x)$. В результате приходим к уравнению критической поверхности Σ в виде

$$P(\lambda, \sigma^2, \varepsilon) \equiv \lambda^4 + \lambda^2\left(1 - 5\sigma^2 - \frac{\sigma^4}{2}\right) - \lambda\varepsilon(9\sigma^4 + 18\sigma^2 - 4\lambda^2) - 4\sigma^2\left(1 - \frac{\sigma^2}{4}\right)^3 - 27\sigma^4\varepsilon^2 = 0, \tag{9}$$

где $\varepsilon = \alpha - 1/2 \in [-1/2, 1/2]$. Выполнение этого равенства является необходимым и достаточным условием для существования кратного корня x_0 при наборе $(\lambda, \sigma^2, \alpha)$ разрешенных значений параметров, который, однако, может как принадлежать, так и не принадлежать интервалу $(0,1)$.

При выполнении равенства (9) из условия $S'(x_0) = 0$ определяется сам кратный корень:

$$
\begin{gathered}
S'(x_0) = 3\sigma^2\left(x_0 - \frac{1}{2}\right)^2 + 2\lambda\left(x_0 - \frac{1}{2}\right) + 1 - \frac{\sigma^2}{4} = 0, \\
x_0 = \frac{1}{2} + \frac{2\lambda(1+2\sigma^2) + (18\alpha - 9)\sigma^2}{12\sigma^2 - 3\sigma^4 - 4\lambda^2}.
\end{gathered} \tag{10}
$$

Причем такое решение возможно только тогда, когда дискриминант квадратного уравнения положителен, что означает выполнение условия $4\lambda^2 + 3\sigma^4 - 12\sigma^2 > 0$. Таким образом, критическая поверхность Σ должна располагаться вне поверхности $\{(\lambda, \sigma^2, \alpha)\colon \Delta \equiv 4\lambda^2 + 3\sigma^4 - 12\sigma^2 = 0\}$ эллиптического цилиндра и иметь с ней точки соприкосновения. В этих точках реализуется тройной корень уравнения $S(x) = 0$: $x_0 = 1/2 - \lambda/3\sigma^2$. Остальные точки поверхности соответствуют двойному корню. Для того чтобы кратный корень x_0 соответствовал бифуркации плотности $p(x)$, необходимо и достаточно, чтобы выполнялось неравенство $0 < x_0 < 1$, эквивалентное неравенству

$$
\left|\frac{4\lambda(1+2\sigma^2) + 36\varepsilon\sigma^2}{4\lambda^2 + 3\sigma^4 - 12\sigma^2}\right| < 1,
$$

которое накладывает дополнительное ограничение на параметры $(\lambda, \sigma^2, \alpha)$ (для тройного корня оно имеет вид $2|\lambda| < 3\sigma^2$). Вводя гиперболы

$$
G_\pm(\lambda, \sigma^2, \varepsilon) \equiv (2\sigma^2 + 1 \pm 2\lambda)^2 - (\sigma^2 + 2(4 \mp 9\varepsilon))^2 + 4(4 \mp 9\varepsilon)^2 - 1 = 0, \tag{11}
$$

запишем последнее неравенство в виде

$$
G_+(\lambda, \sigma^2, \varepsilon) G_-(\lambda, \sigma^2, \varepsilon) \geqslant 0. \tag{12}
$$

Оно определяет допустимую область для расположения точек критической поверхности, границами которой являются соответственно гиперболы $G_\pm = 0$. Это условие очень важно, так как поверхность, определяемая уравнением (9), не является связной[4)].

Замечание 2. Так как при $|\varepsilon| \neq 1/2$ кратное решение x_0 не может пересечь границы 0 и 1 (см. выражение (10)) при изменении параметров, кривая Σ_ε, определяемая уравнением (9), при $\varepsilon = \mathrm{const}$, $|\varepsilon| \neq 1/2$ (и в частности та ее часть, которая представляет собой пересечение критической поверхности с плоскостью $\varepsilon = \mathrm{const}$) не имеет общих точек с гиперболами $G_\pm = 0$. Исключение могут составлять точки с $\sigma^2 = 0$, где плотность $p(x)$ не существует, и точки, где $\Delta = 0$. Последнее связано с невозможностью рассуждать по непрерывности в формуле (10), которая теряет смысл, так как предел в точке, в которой $\Delta = 0$, по различным направлениям может быть разным.

Замечание 3. Пересечение кривой Σ_ε с гиперболами $G_\pm = 0$ в точках, для которых $\Delta = 0$, но $\sigma^2 \neq 0$, возможно только в точках соприкосновения кривой с эллипсом $\Delta = 0$, так как ее точки не могут находиться внутри этого эллипса.

[4)] Заметим, что при $\alpha \neq 1/2$ в том случае, когда параметры $\langle\lambda, \sigma^2, \alpha\rangle$ не соответствуют точке соприкосновения поверхности Σ с поверхностью цилиндра, кроме кратного корня x_0 имеется еще один не равный ему корень. При этом некратный корень соответствует максимуму плотности $p(x)$ и бифуркация состоит в одновременном рождении дополнительного максимума вместе с минимумом.

4. АНАЛИЗ КРИТИЧЕСКОЙ ПОВЕРХНОСТИ

Проведем исследование критической поверхности, изучая ее пересечения с плоскостями при фиксированных значениях $\alpha \in (0,1)$ (либо $\varepsilon \in (-1/2, 1/2)$). Эти пересечения представляют собой кривые Σ_ε четвертого порядка. Ту часть каждой из них, которая удовлетворяет условию (12), в дальнейшем будем называть *критической кривой*. Полная классификация кривых четвертого порядка отсутствует ввиду чрезвычайного разнообразия их качественного устройства (см. книгу [29]). В частности, они могут быть многосвязными, и при этом не существует общего метода определения числа их связанных компонент. Кривая Σ_ε как раз оказывается многосвязной, и поэтому возникает задача выделения именно той из ее компонент, которая соответствует критической кривой. В общем положении значений параметров $(\lambda, \sigma^2, \varepsilon)$ для выделения требуемой компонеты и ее исследования потребуется провести довольно сложный геометрический анализ.

Обозначим посредством $(\lambda_*(\varepsilon), \sigma_*^2(\varepsilon))$ координаты точек соприкосновения кривой Σ_ε с эллипсом $\Delta = 0$ на плоскости $\varepsilon = \mathrm{const}$, указав явно их зависимость от ε. Эти координаты находятся из совместного решения уравнений $P(\lambda_*, \sigma_*^2, \varepsilon) = 0$ и $4\lambda_*^2 + 3\sigma_*^4 - 12\sigma_*^2 = 0$. Из этих уравнений находим, что

$$\lambda_* = -\frac{9\varepsilon\sigma_*^2}{1+2\sigma_*^2}, \tag{13}$$

и, используя уравнение эллипса, получаем, что

$$4(\sigma_*^2 - 1)^3 = 27\sigma_*^2(1-4\varepsilon^2). \tag{14}$$

Уравнение (14) однозначно определяет неявным образом зависимость $\sigma_*^2(\varepsilon)$, так как оно имеет одно вещественное решение $\sigma_*^2(\varepsilon) \geqslant 1$ (при $\sigma_*^2 < 1$ уравнение не имеет решений, так как $|\varepsilon| < 1/2$). В самом деле, в правой части (14) находится линейная функция, а в левой – выпуклая при $\sigma_*^2 > 1$ функция. Следовательно, при $\sigma_*^2 > 1$ имеется не более двух вещественных решений. С другой стороны, значение линейной функции в правой части больше значения функции в левой части при $\sigma_*^2 = 1$. Поэтому имеется только одно пересечение прямой с выпуклой левой частью при указанных значениях σ_*^2. Это пересечение должно происходить при $\sigma_*^2 < 4$. Заметим также, что равенство $\sigma_*^2 = 1$ возможно только при $|\varepsilon| = 1/2$.

Перейдем в уравнении $P(\lambda, \sigma^2, \varepsilon) = 0$ к полярным координатам с центром в точке соприкосновения $(\lambda_*(\varepsilon), \sigma_*(\varepsilon))$:

$$\lambda = \lambda_* + \rho\cos\varphi, \qquad \sigma^2 = \sigma_*^2 + \rho\sin\varphi, \tag{15}$$

и разложим полином $P(\lambda, \sigma^2, \varepsilon)$ в ряд Тейлора около этой точки по степеням ρ. Это разложение обрывается на четвертой степени по ρ:

$$P(\lambda, \sigma^2, \varepsilon) = -3(\sigma_*^2 - 1)^2(\rho\sin\varphi)^2 Q_2(z) + \frac{1}{6}(\rho\sin\varphi)^3 Q_3(z) + \frac{1}{16}(\rho\sin\varphi)^4 Q_4(z),$$

где введена переменная $z = \operatorname{ctg}\varphi$ и ее значение $z_* = \operatorname{ctg}\varphi_* = \lambda_*/3\sigma_*^2 \in [-1/2, 1/2]$,

$$Q_2(z) = (z - z_*)^2 \geqslant 0, \qquad Q_4(z) = (4z^2 - 1)^2 \geqslant 0,$$

$$Q_3(z) = 8z_*(7\sigma_*^2 - 1)z^3 - 6(\sigma_*^2 + 5)z^2 + 18z_*(1+\sigma_*^2)z + \frac{3}{2}(\sigma_*^2 - 3).$$

Точки, для которых $\operatorname{ctg}\varphi = \pm\infty$, $\varphi = 0, \pi$, соответствуют пересечению кривой с уровнем $\sigma^2 = \sigma_*^2$.

Значение z_* принадлежит интервалу $(-1/2, 1/2)$. Это следует из того, что точка (λ_*, σ_*^2) лежит на эллипсе, и поэтому $z_*^2 = \lambda_*^2/9\sigma_*^4 = (4-\sigma_*^2)/12\sigma_*^2 \leqslant 1/4$ при $\sigma_*^2 \geqslant 1$.

Найдем аналитическое выражение, определяющее кривую Σ_ε. Поделим $P(\lambda, \sigma^2, \varepsilon)$ на ρ^2, исключив значение $\rho = 0$, которое соответствует точке соприкосновения кривой с эллипсом. Тогда получаем квадратное уравнение относительно ρ:

$$\frac{1}{16}\rho^2 Q_4(z)\sin^2\varphi + \frac{1}{6}\rho Q_3(z)\sin\varphi - 3(\sigma_*^2-1)^2 Q_2(z) = 0. \tag{16}$$

Дискриминант этого уравнения неотрицателен в силу определения функций $Q_2(z)$, $Q_4(z)$. Поэтому получаем две функции, определяемые уравнением (9):

$$\rho_\pm(\varphi) = \frac{4}{3Q_4(z)\sin\varphi}\Big(-Q_3(z) \pm \sqrt{Q_3^2(z) + 27(\sigma_*^2-1)^2 Q_4(z) Q_2(z)}\Big). \tag{17}$$

Они описывают кривую Σ_ε при тех значениях φ, при которых $\rho_\pm(\varphi) \geqslant 0$.

Ввиду неотрицательности дискриминанта функция $\rho_+(\varphi)$ неотрицательна при $\sin\varphi \geqslant 0$ и поэтому определяет кривую только при $\varphi \in [0, \pi]$. Наоборот, функция $\rho_-(\varphi)$ определяет кривую только при $\sin\varphi \leqslant 0$, $\varphi \in [-\pi, 0]$, независимо от знака $Q_3(z)$.

Используя связь $z_*^2 = (4-\sigma_*^2)/12\sigma_*^2$, находим, что коэффициент $Q_3(z_*)$ положителен:

$$Q_3(z_*) = \frac{8}{9\sigma_*^4}(\sigma_*^2-1)^3 > 0.$$

По непрерывности $Q_3(z) > 0$ в окрестности точки z_* при $\sigma_*^2 > 1$. Тогда функция $\rho_+(\varphi)$ определена при φ, находящемся в окрестности φ_*. При этом ввиду равенств $Q_2(z_*) = 0$, $\rho_+(\varphi_*) = 0$ в этой точке имеется соприкосновение кривой Σ_ε с эллипсом.

Функция $\rho_-(\varphi)$ может обращаться в нуль только в исключительном случае, когда одновременно $Q_3(z) = 0$ и $Q_2(z)Q_4(z) = 0$, что реализуется только при $\varepsilon = \pm 1/2$. Таким образом, при $|\varepsilon| < 1/2$ функция $\rho_-(\varphi) > 0$ при $\varphi \in (-\pi, 0)$.

Если дискриминант не равен нулю, т. е. $|\varepsilon| < 1/2$, то кривые, определяемые функциями $\rho_+(\varphi)$ и $\rho_-(\varphi)$, могут иметь общие точки только при $\varphi = 0, \pi$. Из уравнения (16) следует, что функции $\rho_\pm(\varphi)$, которые являются его решениями, имеют конечные совпадающие для них обеих пределы r_+ и r_- при $\varphi \to 0$ и $\varphi \to \pi$ соответственно, которые удовлетворяют уравнению

$$r_\pm^2 \pm \varkappa r_\pm - 3(\sigma_*^2-1)^2 = 0, \qquad \varkappa = 8z_*\frac{7\sigma_*^2-1}{6}.$$

Теорема 1. *Функции $\rho_+(\varphi)$ и $\rho_-(\varphi)$ определены и неотрицательны соответственно на отрезках $[-\pi, 0]$ и $[0, \pi]$ при $z \neq \pm 1/2$. При этом $\rho_+(\varphi) \to \infty$ при $z \to \pm 1/2$.*

Доказательство. Функция $\rho_+(\varphi)$ определена в окрестности угла φ_*, и при этом $z_* \in (-1/2, 1/2)$. Покажем, что она определена на всем интервале $(-1/2, 1/2)$.

На ограниченном интервале изменения $z = \operatorname{ctg}\varphi$ полиномы $Q_2(z)$, $Q_3(z)$, $Q_4(z)$ от z ограничены. Тогда, как следует из формулы (17), функция $\rho_+(\varphi)$ может стремиться к $+\infty$ только в том случае, когда $Q_4(z) \to 0$, т. е. $z \to \pm 1/2$ и $\sin\varphi \to 2\sqrt{5}/5$.

В этих условиях числитель в (17) стремится к ненулевому значению. Вычисление значений $Q_3(\pm 1/2)$ на основе выражения (13) для λ_* и $z_* = \lambda_*/3\sigma_*^2$ приводит к формуле $Q_3(\pm 1/2) = -12(1 \pm 2\varepsilon) < 0$. Ввиду отрицательности этой величины $\rho_+(\varphi) > 0$. Следовательно, получаем следующую асимптотическую формулу:

$$\rho_+(\varphi) = \frac{\sqrt{5}|Q_3(\pm 1/2)|}{12(z^2 - 1/4)^2}\big(1 + o(1)\big) \quad \text{при} \quad z \to \pm\frac{1}{2}.$$

Таким образом, положительная функция $\rho_+(\varphi)$ в области определения $[0, \pi]$ имеет по переменной z интервалы непрерывности $(-\infty, -1/2)$, $(-1/2, 1/2)$, $(1/2, \infty)$, а при $|z| = 1/2$ – разрывы второго рода.

Из (17) следует, что функция $\rho_-(\varphi)$ определена при всех $\varphi \in (-\pi, 0)$ за исключением, быть может, тех углов, для которых $Q_4(z) = 0$. Однако при $z = \pm 1/2$ она имеет конечный предел

$$\lim_{z \to \pm 1/2,\, \sin\varphi < 0} \rho_-(\varphi) = 9\sqrt{5}(\sigma_*^2 - 1)^2 \frac{Q_2(\pm 1/2)}{|Q_3(\pm 1/2)|}.$$

Следовательно, по непрерывности заключаем, что $\rho_-(\varphi)$ определена на всем интервале $(-\pi, 0)$.

Следствие 1. *Функции $\rho_\pm(\varphi)$ определяют двухсвязную кривую так, что одна ее компонента Σ_+ задается функцией $\rho_+(\varphi)$ при $\varphi \in (\psi, \pi - \psi)$, $\psi = \operatorname{arcctg}(1/2)$, а вторая компонента Σ_- задается на дополнении к $[-\pi, \pi] \backslash [\psi, \pi - \psi]$ функциями*

$$\begin{aligned} &\rho_+(\varphi), && \varphi \in [0, \psi),\\ &\rho_-(\varphi), && \varphi \in [-\pi, 0],\\ &\rho_+(\varphi), && \varphi \in (\pi - \psi, \pi]. \end{aligned}$$

Доказательство. Данное утверждение следует из того, что связная компонента кривой должна определяться непрерывной функцией, и того, что функции $\rho_\pm(\varphi)$ имеют совпадающие предельные значения при $\varphi = 0, \pi$.

Исследуем поведение компоненты Σ_+ в окрестности точки ее соприкосновения $(\lambda_*, \sigma_*^2 > 1)$ с эллипсом $\Delta = 0$, $\varepsilon = \text{const}$, т. е. при значениях φ в малой окрестности угла φ_* или, что то же самое, при значениях z в малой окрестности точки z_*, где $\rho_+(\varphi) = o(1)$. Покажем, что компонента имеет особенность типа касп с острием в этой точке, направленным в сторону эллипса, и касательной, направленной под углом φ_*.

Теорема 2. *В локальных декартовых координатах (u, v) с центром в точке $(\lambda_*, \sigma_*^2 > 1)$ кривая, представляемая функцией $\rho_+(\varphi)$, описывается асимптотической формулой*

$$u = \text{const}\,|v|^{2/3}, \qquad v \to 0,$$

в окрестности точки $(0, 0)$.

Доказательство. Найдем асимптотическое выражение для функции $\rho_+(\varphi)$ в окрестности точки (λ_*, σ_*^2) при малых значениях $z - z_*$. Так как $Q_3(z_*) > 0$, $Q_4(z_*) \neq 0$

и $Q_2(z) = (z - z_*)^2$, из (17) получаем следующую асимптотическую с точностью до $(z - z_*)^2$ формулу при $z \to z_*$:

$$\rho_+(\varphi) = 18\frac{(\sigma_*^2 - 1)^2}{\sin\varphi_*}\frac{Q_2(z)}{Q_3(z_*)} + O((z - z_*)^3).$$

Учитывая явное выражение для $Q_3(z_*)$ и то, что $\sin\varphi_* = (1 + z_*^2)^{-1/2}$, преобразуем эту формулу к виду

$$\rho_+(\varphi) = \frac{3\sigma_*^2}{1 - 4\varepsilon^2}(\sigma_*^2 - 1)^2(1 + z_*^2)^{1/2}(z - z_*)^2 + O((z - z_*)^3).$$

Перейдем в локальные декартовы координаты (u, v) с центром в точке (λ_*, σ_*^2) и u-осью, направленной под углом φ_*. При этом $u > 0$. В терминах таких координат кривая $\rho_+(\varphi)$ представляется уравнением $u^2 + v^2 = C \operatorname{arctg}^4(v/u)$, где $\rho_+(\varphi) = (v^2 + u^2)^{1/2}$, $\varphi - \varphi_* = \operatorname{arctg}(v/u)$. Из уравнения следует, что при $u, v \to 0$ вдоль кривой выполняется условие $v/u \to 0$. Следовательно, при указанном переходе имеет место асимптотическая эквивалентность $u^2 + v^2 \propto (v/u)^4$. В свою очередь это приводит к тому, что $u^6 \propto v^4$, т. е. $u \propto |v|^{2/3}$, $v \to 0$.

Согласно теореме 2 критическая кривая состоит из двух *ветвей*, сшитых в точке соприкосновения с эллипсом $\Delta = 0$.

Для установления возможности пересечения компонент $\Sigma_\pm$ с гиперболами $G_\pm = 0$ перейдем в уравнении (11) к полярным координатам $(\rho^{(\pm)}, \varphi)$:

$$\rho^{(\pm)}[4(1\pm\sin 2\varphi)-\sin^2\varphi]+2[2\cos\varphi(2(\lambda_*\pm\sigma_*^2)\pm1)+\sin\varphi(3\sigma_*^2\pm4\lambda_*-6\pm18\varepsilon)] = 0, \quad (18)$$

где при подстановке использовано, что $G_\pm(\lambda_*, \sigma_*^2, \varepsilon) = 0$, и опущено указание зависимости от ε в величинах λ_* и σ_*^2. Гиперболы $G_\pm = 0$ являются двухсвязными кривыми, но вид уравнения (18) указывает на то, что у каждой из них имеется компонента, которая проходит через центр полярной системы координат. У такой компоненты и только у нее найдется угол $\varphi^{(\pm)}$ такой, что $\rho^{(\pm)}(\varphi^{(\pm)}) = 0$. Углы $\varphi^{(\pm)}$ определяют наклоны касательных к гиперболам относительно направления λ-оси в центре координат. В результате получаем, что

$$z^{(\pm)} = \frac{6z_* \mp 3(\sigma_*^2 - 2)}{2(2\sigma_*^2 + 1) \pm 12\sigma_*^2 z_*}, \quad (19)$$

где введено обозначение $\operatorname{ctg}\varphi^{(\pm)} = z^{(\pm)}$. Здесь знаменатель больше нуля в силу неравенства

$$z_*^2 = \frac{4 - \sigma_*^2}{12\sigma_*^2} < \left(\frac{2\sigma_*^2 + 1}{6\sigma_*^2}\right)^2,$$

которое имеет место при $\sigma_*^2 > 1$.

Следующее утверждение дает ответ на вопрос, какая из компонент $\Sigma_\pm$ кривой соответствует пересечению критической поверхности с плоскостью $\varepsilon = \mathrm{const}$.

Теорема 3. *Критическая кривая представляется компонентой Σ_+ двухсвязной кривой Σ_ε.*

Доказательство. Необходимо установить, какая из непрерывных компонент кривой из указанных в следствии 1 удовлетворяет условию (12). Компоненты кривой могут пересекаться с гиперболами $G_\pm = 0$ либо в точках оси $\sigma^2 = 0$, либо в точках соприкосновения с эллипсом $\Delta = 0$ (см. замечания 2, 3). Из формулы (11) следует, что гиперболы могут пересекать ось $\sigma^2 = 0$ (λ-ось) в точках с $\lambda = 0, \pm 1$. Однако из уравнения (9) получаем, что точки $(\pm 1, 0)$ на плоскости (λ, σ^2) не лежат на кривой Σ_ε при $\varepsilon \neq 1/2$.

Кривая Σ_+ не проходит через точку $(0,0)$ на плоскости (λ, σ^2), которая имеет полярные координаты $((\sigma_*^4 + \lambda_*^2)^{1/2}, \varphi_0)$ относительно (λ_*, σ_*^2), где $\operatorname{ctg} \varphi_0 = \lambda_*/\sigma_*$, $\sin \varphi_0 < 0$, так как $\rho_+(\varphi_0) < 0$. Поскольку кривая Σ_ε проходит через точку $(0,0)$ $(P(0,0,\alpha) = 0)$, через эту точку должна проходить и кривая Σ_-.

Кривая Σ_- согласно ее определению не проходит через точку соприкосновения, в которой допустимо пересечение гипербол $G_\pm = 0$ с Σ_ε (см. замечание 3). Поэтому компонента Σ_- может иметь общие точки с этими гиперболами только при $\sigma^2 = 0$, но такое пересечение согласно вышесказанному имеет место только в точке $(0,0)$, которая принадлежит им обеим.

Докажем теперь, что кривая Σ_- при $\varphi \neq \varphi_0$ находится внутри области, определяемой неравенством $G_+ G_- < 0$. Так как компонента Σ_- может иметь только одну общую точку $(0,0)$ с каждой из гипербол $G_\pm = 0$, она находится в указанной области, если неравенство имеет место в окрестности этой общей точки.

Доказательство выполнимости неравенства $G_- G_+ < 0$ для точек кривой Σ_-, сколь угодно близких к декартовой точке $(0,0)$, основано на уравнении $P(\lambda, \sigma^2, \varepsilon) = 0$, которому она удовлетворяет. Определим исходя из него направление касательной к этой компоненте в точке $(0,0)$, где $\varphi = \varphi_0$.

Так как $(\partial P/\partial \sigma^2)_{(0,0)} = -4$, $(\partial P/\partial \lambda)_{(0,0)} = 0$, по теореме о неявной функции для $\sigma^2(\lambda)$ верно равенство $(d\sigma^2/d\lambda)_{(0,0)} = 0$, т. е. касательная к кривой $\rho_-(\varphi)$ в нулевой точке направлена по прямой с $\sigma^2 = 0$.

Каждая из гипербол пересекает компоненту в нулевой точке. Покажем, что каждая из них имеет в этой точке касательную, пересекающую ось $\sigma^2 = 0$, рассматривая эти гиперболы как функции $\sigma_\pm^2(\lambda)$. Из уравнений (11) следует, что неявные функции $\sigma_\pm^2(\lambda)$ в точке $(0,0)$ имеют касательные с коэффициентами

$$\left(\frac{d\sigma_\pm^2}{d\lambda}\right)_{(0,0)} = [3(1 \mp 3\varepsilon)]^{-1} \neq 0.$$

Тогда точки $(\lambda \in \mathbb{R}, 0)$ компоненты Σ_-, достаточно близкие к нулевой точке, находятся в области $G_- G_+ < 0$, так как гиперболы в этих точках принимают значения $G_\pm(\lambda, 0, \varepsilon) = \pm 4\lambda + O(\lambda^2)$ при $\lambda \to 0$.

Рассмотрим возможность пересечения гипербол компонентой Σ_+. Такая возможность имеется только в точке соприкосновения $(\lambda_*(\varepsilon), \sigma_*(\varepsilon))$, т. е. в центре полярной системы координат. Компонента Σ_+ разделяет полуплоскость $(\lambda, \sigma^2 \geqslant 0)$ на две части. Если имеется пересечение какой-либо из гипербол $G_+ = 0$ или $G_- = 0$ с этой компонентой, то такая гипербола должна при непрерывном изменении угла φ перейти из одной части плоскости в другую, проходя через точку соприкосновения так, что $\rho^{(+)}(\varphi^{(+)}) = 0$ и соответственно $\rho^{(-)}(\varphi^{(-)}) = 0$. Покажем, что такие переходы невозможны.

Так как компонента Σ_+ обладает каспом в точке соприкосновения (λ_*, σ_*), т. е. имеет точку поворота при $\varphi = \varphi_*$ с касательной в виде луча, исходящего под углом φ_* из этой точки, для доказательства невозможности перехода гипербол из одной части полуплоскости в другую достаточно показать, что $\varphi^{(-)} > \varphi_* > \varphi^{(+)}$, т. е. имеют место неравенства $z^{(+)} < z_* < z^{(-)}$. Эти неравенства эквивалентны в силу равенства $z_*^2 = (4 - \sigma_*^2)/12\sigma_*^2$ неравенствам $(1 \pm 2z_*)(\sigma_*^2 - 1) > 0$. Последние справедливы при $z_* \in (-1/2, 1/2)$ и $\sigma_*^2 > 1$.

Следствие 2. *При $\sigma^2 \to \infty$ ветви $\lambda_\pm(\sigma^2)$ критической кривой Σ_+ в плоскости (λ, σ^2) при фиксированном значении ε удовлетворяют неравенствам*

$$-\frac{1}{2}(\sigma^2 - \sigma_*^2) + \lambda_* < \lambda_-(\sigma^2) < \lambda_+(\sigma^2) < \frac{1}{2}(\sigma^2 - \sigma_*^2) + \lambda_*$$

и имеют следующее асимптотическое поведение:

$$\lambda_-(\sigma^2) = -\frac{\sigma^2}{2} + \sigma\sqrt{2(1-2\varepsilon)} + O(1), \qquad \lambda_+(\sigma^2) = \frac{\sigma^2}{2} - \sigma\sqrt{2(1+2\varepsilon)} + O(1).$$

Доказательство. Так как критическая кривая определяется функцией $\rho_+(\varphi)$ при $\varphi \in (\psi, \pi - \psi)$, неравенства следуют из равенств (15) и ограничения на угол $|\operatorname{ctg}\varphi| = |z| < 1/2$.

Разделим уравнение (9) на $\sigma^8/16$ и введем новую переменную $a = 2\lambda/\sigma^2$:

$$\widetilde{P}(a, \sigma^2, \varepsilon) \equiv a^4 + 4a^2\left(\sigma^{-4} - 5\sigma^{-2} - \frac{1}{2}\right) - 8\varepsilon a(9\sigma^{-2} + 18\sigma^{-4} - a^2\sigma^{-2}) - \\ - \left(\frac{4}{\sigma^2} - 1\right)^3 - 432\sigma^{-4}\varepsilon^2 = 0.$$

Согласно теореме 1 и равенствам (15) ветви $\lambda_\pm(\sigma^2)$ компоненты Σ_+ стремятся к бесконечности при $\sigma^2 \to \infty$. Поэтому асимптотики функций $a_\pm(\sigma^2) = 2\lambda_\pm(\sigma^2)/\sigma^2$, которые удовлетворяют этому уравнению, вычисляются на его основе переходом к пределу $\sigma^2 \to \infty$. В силу доказанных неравенств для ветвей $\lambda_\pm(\sigma^2)$ функции $a_\pm(\sigma^2)$ ограничены. Тогда каждая из них имеет предел a_*. Для этих предельных значений получаем уравнение $(a_*^2 - 1)^2 = 0$ такое, что $a_* = \pm 1$ – его двукратно вырожденные корни. Подстановка выражений $a = \pm 1 + b$ в уравнение $\widetilde{P}(a, \sigma^2, \varepsilon) = 0$, где $b = o(1)$ при $\sigma^2 \to \infty$, приводит с точностью до $O(b^3)$ к уравнению $b^2 - 8\sigma^{-2} + 16\varepsilon\sigma^{-2}\operatorname{sgn} a_* = 0$, что дает четыре значения для функций $b = \pm 2\sqrt{2}(1 + 2\varepsilon\operatorname{sgn} a_*)^{1/2}/\sigma$.

Из теоремы 3 следует, что наименьшее и наибольшее из этих четырех значений соответствуют асимптотической кривой Σ_- при $\sigma^2 \to \infty$, что завершает доказательство утверждения.

Пример 1. В качестве примера рассмотрим критическую кривую, которая описывается биквадратным уравнением, в симметричном случае, когда $\alpha = 1/2$. Эта кривая двухсвязна. Решение биквадратного уравнения, соответствующее разрешенной области значений параметров, имеет вид

$$\lambda^2 = \frac{1}{2}\left[\frac{\sigma^4}{2} + 5\sigma^2 - 1 - (2\sigma^2 + 1)^{3/2}\right].$$

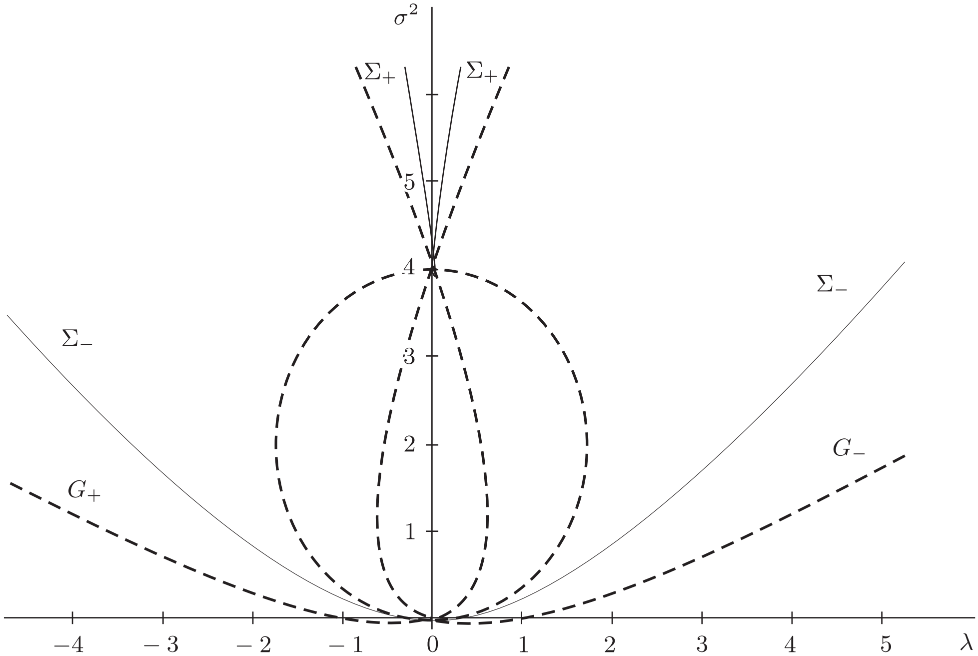

Рис. 1. Кривые Σ_+ и Σ_- в симметричном случае (сплошные линии). Штриховыми линиями показаны гиперболы $G_\pm$ и эллипс $\Delta = 0$.

Из условия $\lambda^2 > 0$ следует ограничение $\sigma^2 \geqslant 4$, т. е. критическая кривая расположена выше эллипса.

В точке соприкосновения $(0, 4)$ кривая $\sigma^2(\lambda)$ имеет касп, который характеризуется *критическим индексом* $2/3$, так как асимптотика кривой в точке $\lambda = 0$ имеет вид $\sigma^2 = 4 + 3(2|\lambda|)^{2/3}\big(1 + o(1)\big)$. График функции $\sigma^2(\lambda)$ приведен на рис. 1.

5. ИССЛЕДОВАНИЕ КРИТИЧЕСКОЙ КРИВОЙ В ПРЕДЕЛЬНЫХ СЛУЧАЯХ

Выводы, полученные в предыдущем разделе, справедливы при любых значениях $\varepsilon \in (-1/2, 1/2)$. Однако с точки зрения получения удобных расчетных формул для критической кривой в разложении выражения (17) для $\rho_+(\varphi)$ вблизи точки каспа можно ограничиться первыми слагаемыми только в случае, когда значение $|\varepsilon|$ не очень близко к $1/2$. В этом разделе мы изучим противоположный случай, когда величина $1/2 - |\varepsilon|$ является малым параметром. Такой анализ опирается на решения уравнения (9) при $|\varepsilon| = 1/2$ несмотря на то, что они являются нефизическими.

Эти решения представляются явными формулами для кривой Σ_ε. Принимая во внимание симметрию критической поверхности при замене ε на $-\varepsilon$, мы изучим решения только в случае $\varepsilon = -1/2$. Непосредственно проверяется, что они описываются следующим образом. Кривая Σ_ε состоит из двукратной прямой $\lambda_0 = 1 + \sigma^2/2$ и кривой, состоящей из двух полупарабол $\lambda_\pm = (-\sigma^2 \pm 4\sigma)/2$, которые сшиваются в точке $\lambda = 3/2$, $\sigma^2 = 1$.

В рассматриваемом нами случае нужно построить теорию возмущений для формы кривой Σ_+ при малых значениях α. Для того чтобы установить тип асимптотического разложения, которое мы строим в виде степенного разложения по дробным степеням α, совершим подстановку

$$\lambda = \frac{3+u+v}{2}, \qquad \sigma^2 = 1+v-u, \tag{20}$$

так что уравнение (9) и уравнение эллипса в этих переменных принимают соответственно вид

$$u^2(v^2+4u)+\alpha\{36u^2+6uv(u+v-3)-4(u^3+v^3)\}-27\alpha^2(1+v-u)^2=0,$$
$$u^2+v^2-uv+3u=0.$$

Найдем правильную асимптотику изменения переменных u, v в окрестности точки $(0,0)$ при $\alpha\to 0$. Для этого произведем замену $u\leftarrow\alpha^a u$, $v\leftarrow\alpha^b v$ с показателями $a>0$, $b>0$, которые выбираются из условия существования в левой части уравнения группы не менее чем из двух слагаемых с одинаковыми минимальными значениями степеней. Анализ возможностей такого выбора параметров a и b после произведенной замены приводит к единственным значениям $a=2/3$, $b=1/3$.

В результате, отбирая слагаемые с минимальной степенью величины α, равной двум, т. е. пренебрегая слагаемыми более высокого порядка по степеням α, имеем выражение

$$R(u)\equiv 4(u^3-v^3-27)+(uv-9)^2=0. \tag{21}$$

Уравнение (21) описывает кривую третьего порядка на плоскости (u,v). Покажем, что она двухсвязна, и выделим из ее компонент ту, которая, как и кривая Σ_+, обладает точкой поворота.

При $v\to\infty$ неявная функция $u(v)$ не может быть ограниченной. Поэтому имеются асимптотики кривой $u(v)\to\infty$ при $v\to\infty$. Из уравнения (21) следует, что возможны следующие типы асимптотического поведения: $u\sim v^2$, $u\sim v^{1/2}$. Положим $u=kv^2\bigl(1+o(1)\bigr)$. После подстановки этого выражения в (21) с удержанием главных членов, пропорциональных v^6, получаем условие $k=-1/4$ для их исчезновения. После этого находим, что $o(1)=-2/v+o(v^{-1})$ является главным поправочным слагаемым. Асимптотика второго типа получается из первой на основе соображений симметрии уравнения (21) относительно u и v.

Применим алгоритм Евклида к паре полиномов $R(u)$ и $R'(u)$. Остаток после применения алгоритма пропорционален $(v^3-27)^3$ и обращается в нуль при $v=3$. Следовательно, имеется точка $(-3,3)$, в которой $R(u)$ имеет кратный корень, т. е. в ней может реализоваться либо самопересечение кривой, либо точка поворота (касп).

Кривая симметрична относительно диагонали $u=-v$ и имеет с ней две точки пересечения. Эти точки определяются из уравнения $(u-1)(u+3)^3=0$, из которого видно, что точка $(-3,3)$ является особой. Во второй точке $(1,-1)$ кривая пересекает трансверсально диагональ $v=-u$, так как в ней невозможно ее самопересечение.

По топологическим соображениям наличие двух точек пересечения с побочной диагональю вместе с неограниченностью кривой позволяет сделать заключение о ее двухсвязности.

Таким образом, согласно поставленной выше задаче нужно выбрать компоненту кривой, на которой находится точка $(-3, 3)$, и исследовать эту компоненту в окрестности этой особой точки. С этой целью перейдем к полярным координатам с центром в точке $(-3, 3)$: $u = -3 + \rho\cos\varphi$, $v = 3 + \rho\sin\varphi$. В результате после исключения тривиального корня $\rho = 0$ получаем следующее квадратное относительно ρ уравнение:

$$\frac{\rho^2}{4}\sin^2 2\varphi + \rho(\cos\varphi - \sin\varphi)(4 + 5\sin 2\varphi) - 27(1 + \sin 2\varphi) = 0.$$

Так как последнее слагаемое отрицательно, положительное решение $\rho(\varphi)$ имеет вид

$$\rho(\varphi) = \frac{2}{\sin^2 2\varphi}\left[(\sin\varphi - \cos\varphi)(4 + 5\sin 2\varphi) + \big(2(2 + \sin 2\varphi)\big)^{3/2}\right]. \tag{22}$$

Для того чтобы выяснить характер особенности при $\rho = 0$, нужно найти углы, при которых $\rho(\varphi)$ стремится к нулю в особой точке с $\rho = 0$. При этом можно в уравнении пренебречь слагаемым, пропорциональным ρ^2. Отсюда находим, что функция $\rho(\varphi)$ стремится к нулю при $\varphi \to -\pi/4$ или $\varphi \to 3\pi/4$. Вводя отклонения $\chi = \varphi + \pi/4$ либо $\chi = \varphi - 3\pi/4$ от этих значений угла φ, находим из уравнения, что в окрестности каждого из них, т. е. при $\chi \to 0$, выполняется условие

$$\rho^2 \pm 4\sqrt{2}\rho - 216\chi^2 + o(\chi^2) = 0,$$

где верхние знаки соответствуют равенству $\varphi = 3\pi/4$, а нижние -- равенству $\varphi = -\pi/4$. Отсюда следует, что $\rho(\varphi) = 27\sqrt{2}\chi^2 + o(\chi^2)$ при знаке $+$, а при знаке $-$ решение $\rho(\varphi)$, обращающееся в нуль при $\chi = 0$, ввиду неотрицательности $\rho(\varphi)$ возможно только для изолированного значения $\chi = 0$. Это означает, что кривая $\rho(\varphi)$ может подходить к особой точке только под углом $\varphi = 3\pi/4$, поэтому она является точкой поворота кривой, т. е. в ней реализуется касп.

Перейдем теперь на выбранной компоненте кривой (22), которая содержит касп, к исходным переменным λ, σ^2. Учитывая все сделанные выше в процессе анализа замены переменных, находим, что

$$\lambda = \frac{1}{2}[3(1 + \alpha^{1/3} - \alpha^{2/3}) + \alpha^{1/3}\rho\sin\varphi - \alpha^{2/3}\rho\cos\varphi] + O(\alpha),$$
$$\sigma^2 = 1 + 3(\alpha^{1/3} + \alpha^{2/3}) + \alpha^{1/3}\rho\sin\varphi - \alpha^{2/3}\rho\cos\varphi + O(\alpha),$$

где $\rho(\varphi)$ определяется равенством (22). Эти формулы параметрически определяют приближенно, с точностью до $\alpha^{2/3}$, критическую кривую при малых значениях α. При $\rho = 0$ они описывают сдвиг точки каспа $(\lambda_*(\varepsilon), \sigma_*^2(\varepsilon))$ при ε вблизи значения $-1/2$.

6. ЗАКЛЮЧЕНИЕ

В отличие от работ других авторов в настоящей статье проведен полный анализ критической поверхности Σ в модели Хорстхемке–Лефевра, разбивающей ее пространство параметров $(\lambda, \sigma^2 > 0,\ \alpha \in (0, 1))$ на две области, в каждой из которых она имеет два качественно различных стационарных динамических режима. Переход между этими двумя режимами при достаточно медленном (квазистатическом) изменении параметров системы представляет с физической точки зрения фазовый переход между двумя "фазами": унимодальной и бимодальной. Динамический режим в бимодальной фазе состоит из последовательно сменяющих друг друга временных интервалов случайной длительности, в которых относительная концентрация реагентов флуктуирует вблизи значения одного из двух максимумов плотности $p(x)$.

Рассматривая бифуркационную перестройку динамического режима системы как термодинамический фазовый переход, для его количественной характеризации необходимо ввести параметр порядка. В качестве такового, по-видимому, нужно выбрать половину расстояния между концентрациями, соответствующими двум модам плотности распределения $p(x)$. Проведем классификацию фазовых переходов в рассмотренной системе, приняв за основу их разделение на два типа согласно следующему признаку: появляется ли в результате перехода отличное от нуля значение параметра порядка скачкообразно (переход первого рода) или непрерывно (переход второго рода). Тогда в том случае, когда перестройка плотности $p(x)$ происходит с образованием не более чем двукратного корня уравнения $dp(x)/dx = 0$, второй максимум плотности возникает отдельно от уже существующего у нее максимума. Поэтому расстояние между этими максимумами не равно нулю в момент перехода и можно говорить о переходе первого рода. С аналитической точки зрения переход реализуется в виде катастрофы *складки* согласно классификации Тома. Если же перестройка плотности происходит так, что уравнение $dp(x)/dx = 0$ имеет трехкратный корень, то из исчезающего максимума рождается сразу два новых максимума. Поэтому в этом случае параметр порядка непрерывным образом начинает возрастать, начиная с нулевого значения, и нужно говорить о фазовом переходе второго рода.

В соответствии с проведенным анализом модели второй случай реализуется в точке каспа критической кривой, которая находится на эллипсе $\Delta = 0$. При этом мода, в которой происходит бифуркация, расположена в точке $x_0 = 1/2 - \lambda/3\sigma^2$. Согласно классификации Тома этот переход происходит в результате катастрофы *сборки*. Применимость такой классификации связана с тем, что $p(x)$ аналитически зависит от параметров λ и σ^2. Если положить, что роль температуры в рассматриваемой системе выполняет интенсивность шума σ^2, то критический индекс параметра порядка в точках соприкосновения критической кривой равен $1/2$, как это имеет место для катастрофы сборки:

$$(\sigma^2 - \sigma_*^2)\left(\frac{\partial p'}{\partial \sigma^2}\right)_{x_*} + \frac{1}{2}(x_0 - x_*)^2 p'''(x_*) = 0, \qquad x_0 - x_* \sim (\sigma^2 - \sigma_*^2)^{1/2}.$$

Вместе с тем нужно отметить, что фазовый переход первого рода в системе происходит без дополнительных затрат теплоты на образование новой фазы, если в качестве термодинамической энтропии S системы выбрать энтропию Шеннона

$$\int_0^1 p(x) \ln p(x)\, dx,$$

которая изменяется непрерывно с изменением параметров системы. Тогда термодинамическая связь $\delta Q = T\,\delta S$ указывает на отсутствие теплового скачка при переходе из унимодальной фазы в бимодальную.

Список литературы

[1] M. Kimura, T. Ohta, *Theoretical Aspects of Population Genetics*, Monographs in Population Biology, **4**, Princeton Univ. Press, Boston, 1971.

[2] L. Arnold, W. Horsthemke, R. Lefever, *Z. Phys. B*, **29**:4 (1978), 367–373.

[3] Т. М. Фам, Ю. П. Вирченко, *Научные ведомости БелГУ. Сер. Матем. Физика*, **12(155)**:31 (2013), 130–146.

[4] В. Хорстхемке, Р. Лефевр, *Индуцированные шумом переходы*: *теория и применение в физике*, *химии и биологии*, Мир, М., 1987.

[5] J. Smythe, F. Moss, P. V. E. McClintock, *Phys. Rev. Lett.*, **51**:12 (1983), 1062–1065.

[6] P. S. Landa, P. V. E. McClintock, *Phys. Rep.*, **323**:1, 1–80.

[7] W. Horsthemke, "Noise-induced transitions", *Stochastic Nonlinear Systems in Physics*, *Chemistry*, *and Biology* (University of Bielefeld, West Germany, October 5–11, 1980), eds. L. Arnold, R. Lefever, Springer, Berlin, 1981, 116–126.

[8] R. Lefever, "Noise-induced transitions in biological systems", *Stochastic Nonlinear Systems in Physics*, *Chemistry*, *and Biology* (University of Bielefeld, West Germany, October 5–11, 1980), eds. L. Arnold, R. Lefever, Springer, Berlin, 1981, 127–136.

[9] W. Horsthemke, R. Lefever, "Noise-induced transitions", *Noise and Nonlinear Dynamical Systems*, v. 2: *Theory of Noise-Induced Processes in Special Applications*, eds. F. Moss, P. V. E. McClintock, Cambridge, 2009, 179–208.

[10] W. Horsthemke, R. Lefever, *Noise-induced transitions*: *Theory and Applications in Physics*, *Chemistry and Biology*, Springer Series in Synergetics, **15**, Springer, Berlin, 2006.

[11] Т. М. Фам, Ю. П. Вирченко, *Научные ведомости БелГУ. Сер. Матем. Физика*, **26(169)**:33 (2013), 57–63.

[12] Т. М., Ю. П. Вирченко, *Научные ведомости БелГУ. Сер. Матем. Физика*, **5(176)**:34 (2014), 103–111.

[13] Фам Минь Туан, Ю. П. Вирченко, *Научные ведомости БелГУ. Сер. Матем. Физика*, **25(196)**:37 (2014), 108–118.

[14] Г. С. Яблонский, В. И. Быков, А. Н. Горбань, *Кинетические модели каталитических реакций*, Наука, Новосибирск, 1983.

[15] В. С. Пугачев, И. Н. Синицын, *Стохастические дифференциальные системы. Анализ и фильтрация*, Наука, Новосибирск, 1990.

[16] R. L. Stratonovich, *SIAM J. Control*, **4** (1966), 362–371.

[17] K. Itô, *Nagoya Math. J.*, **1** (1950), 35–47.

[18] N. G. van Kampen, *J. Statist. Phys.*, **24**:1 (1981), 175–187.

[19] W. Moon, J. S. Wettlaufer, *New J. Phys.*, **16** (2014), 055017, 14 pp.

[20] E. Wong, M. Zakai, *Ann. Math. Stat.*, **36** (1965), 1560–1564.

[21] G. Blankenship, G. C. Papanicolaou, *SIAM J. Appl. Math.*, **34**:3 (1978), 437–476.

[22] J. Smythe, F. Moss, P. V. E. McClintock, D. Clarkson, *Phys. Lett. A.*, **97**:3 (1983), 95–98.

[23] И. И. Гихман, А. В. Скороход, *Стохастические дифференциальные уравнения*, Наукова думка, Киев, 1968.

[24] Н. В. Ласкин, С. В. Пелетминский, В. И. Приходько, *ТМФ*, **34**:2 (1978), 244–255.

[25] Т. М. Фам, Ю. П. Вирченко, *Научные ведомости БелГУ. Сер. Матем. Физика*, **11(208)**:39 (2014), 161–166.

[26] Ю. П. Вирченко, Н. В. Ласкин, "Огрубленное описание распределения решений уравнения Ланжевена", *ТМФ*, **41**:3 (1979), 406–417.

[27] А. Н. Тихонов, А. А. Самарский, *Уравнения математической физики*, Изд-во Моск. ун-та, М., 1999.

[28] J. Elliott, *Trans. Amer. Math. Soc.*, **78** (1955), 406–425.

[29] А. А. Савелов, *Плоские кривые. Систематика*, *свойства*, *применения*, Физматгиз, М., 1960.

Поступила в редакцию 3.11.2015,
после доработки 22.11.2015