

СОДЕРЖАНИЕ

Валеев К. Г. Исследование неподвижных точек отображений	333
Коган Н. Я. К теории одноосного гиростабилизатора. I	345
Коган Н. Я. К теории одноосного гиростабилизатора. II	358
Леонов Н. Н. Об экстремальных системах с преимущественным поиском	371
Леонов Н. Н. Об оптимальном управлении в экстремальных системах	383
Кориков А. М. Применение разностных уравнений для исследования установившегося режима работы градиентной экстремальной системы с инерционным объектом	388
Валеев К. Г., Яковенко В. С. О конечных решениях системы линейных дифференциальных уравнений с периодическими коэффициентами	399
Бочков Г. Н., Музычук О. В. О стохастической устойчивости линейных систем с нестационарно флуктуирующими параметрами	403
Постников Л. В. Некоторые вопросы взаимодействия квазигармонических колебаний в слабонелинейных системах	414
Баталова З. С., Иванова И. М., Неймарк Ю. И. О классификации по синдрому	423
Стронгин Р. Г. Выбор испытаний и условие остановки в одномерном глобальном поиске	432
Бедина А. А., Коган Д. И. Некоторые классы дискретных управляемых систем и языки (алгоритмические проблемы)	441
Перельмутер В. М. Об идентификации объектов методом максимального правдоподобия	450
Мошкина Т. В., Тарапович А. С. К теории резистивного параметрического генератора	459
Глик А. Л. Анализ электропневматического сервомеханизма с сухим трением способом гармонической линеаризации нелинейностей	468
Батищев Д. И., Литвер А. В. Тестирование метода кусочно-линейной аппроксимации на одном классе многоэкстремальных функций	472
Краткие сообщения и письма в редакцию	
Рахимов Б. Э. О влиянии формы характеристики ФД на работу системы ФАП при действии помех	476
Бакаев Ю. Н. Критерий синхронизации для системы фазовой автоподстройки с нелинейным фильтром	479

УДК 517.949

ИССЛЕДОВАНИЕ НЕПОДВИЖНЫХ ТОЧЕК ТОЧЕЧНЫХ ОТОБРАЖЕНИЙ

K. Г. Валеев

Указывается общий асимптотический метод исследования устойчивости неподвижных точек. Дается способ построения решения системы разностных уравнений в окрестности неподвижных точек. Результаты применяются к линейным разностным уравнениям с переменными коэффициентами.

1. НЕПОДВИЖНЫЕ ТОЧКИ ОТОБРАЖЕНИЙ

Многие задачи механики и теории колебаний [1] приводят к рассмотрению системы уравнений

$$Y = F(X), \quad (1.1)$$

где X, Y — n -мерные векторы. Рассматриваем проекции векторов как координаты точек в n -мерном пространстве. Вектор и соответствующую точку обозначаем одинаково. Точка X отображается в точку Y с помощью уравнений (1.1). При многократном повторении отображений приходим к системе разностных уравнений

$$X_{k+1} = F(X_k) \quad (k = 0, 1, 2, \dots). \quad (1.2)$$

Вместо системы (1.2) можно рассматривать эквивалентную систему разностных уравнений с непрерывно изменяющимся аргументом

$$\begin{aligned} W(t+h) &= F(W(t)), \\ X_k &= W(kh) \quad (k = 0, 1, 2, \dots). \end{aligned} \quad (1.3)$$

Неподвижные точки отображения находятся из уравнений

$$C = F(C).$$

Рассмотрим малую окрестность неподвижной точки, полагая

$$W(t) = C + \mu Y(t). \quad (1.4)$$

Предполагая достаточную дифференцируемость $F(Y)$ по всем аргументам y_1, \dots, y_n — проекциям вектора Y , придем к квазилинейной системе разностных уравнений

$$Y(t+h) = A Y(t) + \mu F_1(Y(t)) + \mu^2 F_2(Y(t)) + \dots \quad (1.5)$$

Матрица A является матрицей Якоби в точке $X = C$:

$$A = \frac{DF(C)}{DC}, \quad A = \|a_{ij}\|_1^n, \quad a_{ij} = \frac{\partial f_i}{\partial c_j}, \quad F \equiv \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}, \quad C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}. \quad (1.6)$$

Проекции векторов $F_j(Y)$ являются однородными многочленами степени $j+1$. Линейной заменой переменных можно привести систему

(1.5) к уравнениям того же вида, где вместо матрицы A будет матрица J_1 — каноническая жорданова форма. Введением нового малого параметра и дополнительной заменой можно систему (1.5) привести к каноническому виду

$$X(t+h) = JX(t) + \mu H_1(X(t)) + \mu^2 H_2(X(t)) + \dots, \quad H_j(0) = 0, \quad (1.7)$$

где J — диагональная матрица с собственными числами $\lambda_1, \dots, \lambda_n$. В дальнейшем рассматривается система (1.7), где $H_j(\lambda)$ — векторы с полиномиальными проекциями.

2. ВСПОМОГАТЕЛЬНАЯ СИСТЕМА ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

В важном частном случае, когда J — единичная матрица, имеем

$$X(t+h) - X(t) = \mu H_1(X(t)) + \mu^2 H_2(X(t)) + \dots. \quad (2.1)$$

В работе [2] показано, что, приняв $h = \mu$, можно построить семейство частных решений системы (2.1)

$$X(t) = Z_0(t) + \mu Z_1(t) + \mu^2 Z_2(t) + \dots \equiv Z(t, \mu), \quad (2.2)$$

где $Z(t, \mu)$ удовлетворяет вспомогательной системе дифференциальных уравнений

$$\begin{aligned} \frac{dZ}{dt} &= S_0(Z) + \mu S_1(Z) + \mu^2 S_2(Z) + \dots, \\ S_0(Z) &\equiv H_1(Z). \end{aligned} \quad (2.3)$$

Разложение (2.2) можно найти, подставляя $Z(t, \mu)$ в (2.1). Определение векторов $Z_j(t)$ приводит к решению систем дифференциальных уравнений, решение которых зависит от разрешимости системы (2.3) при $\mu = 0$. Построение вспомогательной системы (2.3) осуществляется с помощью дифференцирований. В работе [2] показан асимптотический при $\mu \rightarrow 0$ характер разложения (2.2).

3. ПРИБЛИЖЕННОЕ РЕШЕНИЕ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

При решении системы уравнений первого порядка

$$\frac{dY(t)}{dt} = F(t, Y(t)), \quad Y(0) = Y_0 \quad (3.1)$$

с помощью метода Эйлера или методов типа Рунге—Кутта [3] приходим к системе разностных уравнений

$$Y_{k+1} - Y_k = hH(kh, Y_k, h) \quad (k = 0, 1, 2, \dots; h > 0). \quad (3.2)$$

Вспомогательная система дифференциальных уравнений примет вид

$$\frac{dZ}{dt} = S(t, Z, h), \quad Z \equiv Z(t, h), \quad Z(0, h) = Y_0, \quad (3.3)$$

при этом выполняется условие

$$Y_k = Z(kh, h) \quad (k = 0, 1, 2, \dots). \quad (3.4)$$

Поэтому, чтобы сравнивать решение системы (3.1) и ее приближенного решения, определяемого разностными уравнениями (3.2), достаточно сравнивать решения систем обыкновенных дифференциальных уравнений (3.1), (3.3). Отметим, что

$$\lim_{h \rightarrow 0} S(t, Z, h) = F(t, Z), \quad \lim_{h \rightarrow 0} Z(t, h) = Y(t).$$

Для методов типа Адамса [3] приходим к системе разностных уравнений

$$Y_{k+1} - Y_k = hH(kh, Y_{k+1}, Y_k, Y_{k-1}, \dots, Y_{k-l}, h). \quad (3.5)$$

Можно доказать существование вспомогательной системы вида (3.3), такой, что любое решение $Z(t, h)$ определяет по формуле (3.4) некоторое решение системы (3.5). Независимо от начальных значений $Y_0, Y_{-1}, \dots, Y_{-l}$, любое решение Y_k системы (3.5) при возрастании k стремится к одному из решений вида (3.4). На основе этих выводов можно построить всю теорию численного интегрирования уравнений (3.1).

4. ИНТЕГРАЛЫ РАЗНОСТНЫХ УРАВНЕНИЙ С ПОСТОЯННЫМИ КОЭФФИЦИЕНТАМИ

Рассматривается система разностных уравнений

$$X(t + h) = JX(t). \quad (4.1)$$

Функцию $v(t, X)$ будем называть интегралом системы разностных уравнений (4.1), если для любого решения $X(t)$ системы (4.1) имеем

$$v(t, X(t)) = \psi(t), \quad \psi(t + h) \equiv \psi(t).$$

Функция $v(X)$ будет интегралом, если

$$v(JX) \equiv v(X). \quad (4.2)$$

Ищем полиномиальные интегралы в виде произведения

$$v(X) = x_1^{m_1} x_2^{m_2} \dots x_n^{m_n} \quad (m_j \geq 0). \quad (4.3)$$

Система (4.1) в скалярной форме имеет вид

$$x_j(t + h) = \lambda_j x_j(t) \quad (j = 1, \dots, n). \quad (4.4)$$

Выражение $v(X)$ (4.3) будет интегралом, если выполнено соотношение

$$\lambda_1^{m_1} \lambda_2^{m_2} \dots \lambda_n^{m_n} = 1 \quad (m_j = 0, 1, 2, \dots; \sum_{j=1}^n m_j > 0) \quad (4.5)$$

или

$$m_1 p_1 + m_2 p_2 + \dots + m_n p_n = 0 \pmod{2\pi i}, \quad (4.6)$$

$$p_j = \ln \lambda_j.$$

Полиномиальные интегралы системы (4.1) существуют, если показатели $p_j = \ln \lambda_j$ являются рационально зависимыми. Если интеграл вида (4.3) выражается полиномиально через другие интегралы, то эти интегралы полиномиально зависимы. Выберем полиномиально независимые интегралы $v_1(X), \dots, v_q(X)$, такие, что остальные интегралы выражаются полиномиально через $v_1(X), \dots, v_q(X)$. Введем векторы M, P :

$$M(m_1, \dots, m_n), \quad P(p_1, \dots, p_n) \quad (m_j = 0, 1, 2, \dots).$$

Множество векторов, удовлетворяющих условию

$$MP \equiv m_1 p_1 + m_2 p_2 + \dots + m_n p_n = 0 \pmod{2\pi i}, \quad M \neq 0,$$

обозначим через $\{M\}$. Выберем наименьшее множество векторов M_1, \dots, M_q , таких, что любой вектор $M \in \{M\}$ представим в виде

$$M = v_1 M_1 + v_2 M_2 + \dots + v_n M_n \quad (v_j = 0, 1, 2, \dots).$$

Обозначая через m_{js} проекции вектора M_j , получим полиномиально независимые интегралы в виде

$$v_j(X) = x_1^{m_{j1}} x_2^{m_{j2}} \dots x_n^{m_{jn}} \quad (j = 1, \dots, q). \quad (4.7)$$

Если одна из проекций $m_{js} \neq 0$, то x_s будем называть резонансной переменной, λ_s — резонансным мультиплликатором, $p_s = \ln \lambda_s$ — резонансным показателем. Остальные величины называем нерезонансными.

Пример 1. Рассматривается система разностных уравнений

$$\begin{aligned} x_1(t+h) &= x_1(t), & x_2(t+h) &= e^{i\varphi} x_2(t), \\ x_3(t+h) &= e^{-i\varphi} x_3(t). \end{aligned}$$

Эта система имеет два полиномиально независимых интеграла

$$v_1 = x_1, \quad v_2 = x_2 x_3.$$

5. РАЗНОСТНЫЕ УРАВНЕНИЯ С ЧАСТНЫМИ РАЗНОСТЯМИ

Ищется полиномиальное решение уравнения

$$\varphi(JX) - \varphi(X) = \psi(X),$$

$$\varphi(X) = \sum \psi_{m_1, m_2, \dots, m_n} x_1^{m_1} x_2^{m_2} \dots x_n^{m_n}. \quad (5.1)$$

Если в известной правой части отсутствуют члены, являющиеся интегралами, то имеем решение

$$\varphi(X) = \sum \psi_{m_1, m_2, \dots, m_n} (\lambda_1^{m_1}, \lambda_2^{m_2}, \dots, \lambda_n^{m_n} - 1)^{-1} x_1^{m_1} x_2^{m_2} \dots x_n^{m_n}. \quad (5.2)$$

Из этого вывода следует теорема.

Теорема 1. В разностном уравнении с частными разностями

$$\varphi(JX) - \varphi(X) + \chi(v_1, \dots, v_q) = \psi(X), \quad (5.3)$$

где $\psi(X)$ — произвольный многочлен, всегда можно выбрать многочлен $\chi(v_1, \dots, v_q)$ так, чтобы уравнение (5.3) имело полиномиальное решение. Здесь $v_1(X), \dots, v_q(X)$ — полиномиально независимые интегралы системы (4.1), через которые выражаются все полиномиальные интегралы системы (4.1).

Обозначая через $\langle f(t) \rangle$ не зависящий от времени член в разложении функции $f(t)$ в ряд по экспоненциальным функциям, найдем, что

$$\chi(v_1, \dots, v_q) = \langle \psi(J^{th-1} X) \rangle. \quad (5.4)$$

Если $|\lambda_j| \leq 1$ ($j = 1, \dots, n$), то можно использовать формулу

$$\chi(v_1, \dots, v_q) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \psi(J^{th-1} X) dt. \quad (5.5)$$

Особенно важен случай, когда $|\lambda_j| = 1$ ($j = 1, \dots, n$). В этом случае для отыскания χ применимо усреднение по времени

$$\chi(v_1, \dots, v_q) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \psi(J^{th-1} X) dt. \quad (5.6)$$

Пример 2. Рассматривается уравнение вида (5.3)

$$\varphi(2x_1, 05x_2) - \varphi(x_1, x_2) + \chi(v) = 1 + x_1 + x_1 x_2 + 3x_2^2.$$

Полагая $h = 1$, находим по формуле (5.4) $\chi(v)$:

$$\chi(v) = \langle 1 + 2^t x_1 + x_1 x_2 + 3 \cdot 2^{-2t} x_2^2 \rangle = 1 + x_1 x_2, \quad v = x_1 x_2.$$

Выбирая $\chi(v)$ указанным способом, находим полиномиальное решение

$$\varphi(x_1, x_2) = x_1 - 4x_2^2, \quad \chi(v) = 1 + v, \quad v = x_1 x_2.$$

6. МЕТОД РАСШИРЕНИЯ СИСТЕМЫ УРАВНЕНИЙ

Рассматривается система разностных уравнений вида (1.7), где $H_i(X)$ — полиномиальные векторы,

$$X(t+h) = JX(t) + \mu H_1(X(t)) + \mu^2 H_2(X(t)) + \dots \quad (6.1)$$

Вводим вспомогательные переменные $z_1(t), \dots, z_q(t)$ объединенные в вектор $Z(t)$:

$$Z(t) = \Phi_0(X(t)) + \mu \Phi_1(X(t)) + \mu^2 \Phi_2(X(t)) + \dots \quad (6.2)$$

Предполагаем, что вектор $Z(t)$ удовлетворяет системе уравнений вида (2.1)

$$Z(t+h) = Z(t) + \mu Q_1(Z(t)) + \mu^2 Q_2(Z(t)) + \dots \quad (6.3)$$

Основная идея заключается в рассмотрении расширенной системы (6.1), (6.3) вместо системы (6.1). Система (6.3) строится одновременно с уравнениями (6.2). Уравнения (6.2) определяют соотношения между переменными $X(t)$, $Z(t)$. Эти соотношения удовлетворяют системе (6.1), (6.3) в силу самих соотношений (6.2). Заменяя в (6.2) t на $t+h$, имеем

$$Z(t) + \sum_{s=1} \mu^s Q_s(Z(t)) = \sum_{s=0} \mu^s \Phi_s(X(t+h)). \quad (6.4)$$

Исключая в помощь (6.2) $Z(t)$ и с помощью (6.1) $X(t+h)$, получим уравнения

$$\sum_{s=0} \mu^s \Phi_s(X) + \sum_{k=1} \mu^k Q_k \left(\sum_{s=0} \mu^s \Phi_s(X) \right) = \sum_{k=0} \mu^k \Phi_k \left(JX + \sum_{s=1} \mu^s H_s(X) \right). \quad (6.5)$$

Раскладывая функции в (6.5) по степеням μ , приходим к уравнениям с частными разностями для отыскания $\Phi_k(X)$, $Q_k(Z)$:

$$\Phi_0(JX) - \Phi_0(X) = 0, \quad (6.6)$$

$$\Phi_k(JX) - \Phi_k(X) - Q_k(\Phi_0) = \Psi_k(X) \quad (6.7)$$

$$(k = 1, 2, \dots).$$

Правая часть $\Psi_k(X)$ известна, если известны $\Phi_0(X), \dots, \Phi_{k-1}(X)$, $Q_1(Z), \dots, Q_{k-1}(Z)$. Выберем в качестве проекций вектора $\Phi_0(X)$ все полиномиально независимые интегралы системы разностных уравнений (4.1). В этом случае вектор $Q_k(Z)$ всегда можно выбрать так, чтобы вектор-функция $\Psi(X)$, где

$$\Psi(X) = \Psi_k(X) + Q_k(\Phi_0),$$

не содержала в своем разложении интегралов системы (4.1). В этом случае находим $\Phi_k(X)$ с полиномиальными проекциями. Обрывая этот процесс на некотором шаге, получаем приближенные выражения для $Z(t)$ (6.2) и системы (6.3).

Система (6.3) принадлежит к изученному виду (2.1). Приближен-

ное решение системы (6.3) можно построить, если полностью интегрируется система обыкновенных дифференциальных уравнений

$$\frac{dZ}{dt} = Q_1(Z). \quad (6.8)$$

Стационарные решения системы (6.3) определяет многообразие решений системы (6.1). Общее решение системы (6.3) определяет интегралы в специальной форме (6.2) системы (6.1).

Отметим, что все результаты имеют асимптотический характер при $\mu \rightarrow 0$. Все разложения по степеням малого параметра обычно расходятся.

7. ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ НЕПОДВИЖНОЙ ТОЧКИ

Рассмотрим вещественную систему (6.1) в предположении о том, что собственные числа λ_j матрицы J удовлетворяют условию $|\lambda_j| \leq 1$ и имеются резонансные собственные числа, модуль которых равен единице. Это наиболее сложные резонансные случаи, когда применим метод расширения системы. В этих случаях при разных λ_j имеем полиномиальные интегралы—проекции вектора $\Phi_0(X)$ в (6.2):

$$\begin{aligned} \lambda_1 &= 1, \quad v = x_1; \quad \lambda_2 = -1, \quad v = x_2^2; \\ \lambda_3 &= e^{i\varphi}, \quad \lambda_4 = e^{-i\varphi} \quad (0 < \varphi < \pi), \quad v = x_3 x_4 = |x_3|^2. \end{aligned} \quad (7.1)$$

Если $H_j(0) = 0$ ($j = 1, 2, \dots$), то всегда можно выбрать векторы $\Phi_j(X)$, $Q_j(Z)$, чтобы выполнялось условие

$$\Phi_j(0) = 0 \quad (j = 0, 1, \dots), \quad Q_j(0) \quad (j = 1, 2, \dots).$$

В этом случае нулевому решению $X = 0$ системы (6.1) соответствует нулевое решение $Z = 0$ системы (6.3). При этом, если решение $Z = 0$ системы (6.3) неустойчиво, то неустойчиво решение $X = 0$ системы (6.1). Если решение $Z = 0$ системы (6.3) устойчиво, то из (7.1) следует ограниченность значений резонансных переменных, так как остальные нерезонансные переменные $x_j \rightarrow 0$ при $t \rightarrow +\infty$. Имеем теорему.

Теорема 2. Если нулевое решение вспомогательной системы (6.3) устойчиво (асимптотически устойчиво, неустойчиво) при учете членов порядка μ^N , то нулевое решение системы (6.1) устойчиво (асимптотически устойчиво, неустойчиво) в N -м приближении, т. е. при учете лишь членов порядка не выше чем μ^N .

Теорема 2 дает удобный способ для исследования устойчивости в критических случаях.

Пример 3. Рассмотрим систему разностных уравнений

$$\begin{aligned} x_1(t+h) &= e^{i\varphi} x_1(t) + \mu \sum a_{ks} x_1^k(t) x_2^s(t) + \mu^2 \sum b_{ks} x_1^k(t) x_2^s(t) + \dots, \\ x_2(t+h) &= e^{-i\varphi} x_2(t) + \mu \sum a'_{ks} x_2^k(t) x_1^s(t) + \mu^2 \sum b'_{ks} x_2^k(t) x_1^s(t) + \dots. \end{aligned} \quad (7.2)$$

Штрихом обозначаем комплексно-сопряженную величину. Так как порождающая система при $\mu = 0$ имеет один интеграл, то вспомогательное уравнение (6.3) имеет первый порядок

$$z(t+h) = z(t) + \mu Q_1(z(t)) + \mu^2 Q_2(z(t)) + \dots \quad (7.3)$$

Вычисления показывают, что

$$Q_1(z) = 2 \sum_{k=s+1} \operatorname{Re} (e^{-i\varphi} a_{ks}) z^k,$$

$$Q_2(z) = 2 \sum_{k=s+1} \operatorname{Re} (e^{-i\varphi} b_{ks}) z^k + \sum_{k+\gamma=s+r} a_{ks} a'_{r\gamma} z^{k+\gamma} + \\ + 2 \sum_{\substack{k \neq s+1 \\ r+k-1=\gamma+s}} \operatorname{Re} \frac{k a_{ks} a_r e^{-2i\varphi} + (s+1) a_{ks} a'_{r-1,\gamma}}{\exp[i\varphi(s+1-k)] - 1} z^{r+s}. \quad (7.4)$$

Индексы пробегают все значения, удовлетворяющие указанным условиям. Исследование устойчивости уравнения (7.3) сводится к исследованию устойчивости дифференциального уравнения первого порядка

$$h \frac{dz}{dt} = \mu Q_1(z) + \mu^2 Q_2(z) - 0,5 \mu^2 Q_1(z) \frac{dQ_1(z)}{dz} + \dots \quad (7.5)$$

В более конкретных примерах исследование устойчивости легко проводится с помощью уравнения (7.5).

Пример 4. Рассмотрим точечное отображение (7.2) в частном случае

$$\begin{aligned} x_1(t+h) &= e^{i\varphi} x_1(t) + \mu(ax_1^2(t) + bx_1(t)x_2(t) + cx_2^2(t)), \\ x_2(t+h) &= e^{-i\varphi} x_2(t) + \mu(a'x_2^2(t) + b'x_1(t)x_2(t) + c'x_1^2(t)). \end{aligned} \quad (7.6)$$

Уравнение (7.3) примет вид

$$\begin{aligned} z(t+h) &= z(t) + \mu^2 g z^2(t) + \dots, \quad g = -bb' - 2cc' + \\ &+ \operatorname{Re} \frac{4e^{-i\varphi} - 2e^{-2i\varphi}}{1 - e^{-i\varphi}} ab. \end{aligned}$$

Дифференциальное уравнение (7.5)

$$h \frac{dz}{dt} = \mu^2 g z^2 + \dots \quad (z \geq 0)$$

имеет устойчивое положительное решение при $g < 0$. Следовательно, решения системы (7.6) устойчивы при $g < 0$ и неустойчивы при $g > 0$. При $g = 0$ имеем сомнительный случай.

8. ПОСТРОЕНИЕ РЕШЕНИЯ СИСТЕМЫ РАЗНОСТНЫХ УРАВНЕНИЙ

Ищем решение системы разностных уравнений (6.1)

$$X(t+h) = JX(t) + \mu H_1(X(t)) + \mu^2 H_2(X(t)) + \dots \quad (8.1)$$

в виде асимптотического при $\mu \rightarrow 0$ разложения. Возьмем отдельную проекцию $x_j(t)$. Вводим новую переменную $w_j(t)$

$$w_j(t) = [1 + \mu a_1(X) + \mu^2 a_2(X) + \dots] x_j(t). \quad (8.2)$$

Предполагаем, что $w_j(t)$ удовлетворяет простому разностному уравнению

$$w_j(t+h) = [\lambda_j + \mu b_1(Z(t)) + \mu^2 b_2(Z(t)) + \dots] w_j(t). \quad (8.3)$$

Подставляем (8.2) в (8.3) и приходим к уравнениям для определения $a_k(X)$, $b_k(Z)$ вида

$$\begin{aligned} a_k(JX) \lambda_j - a_k(X) &= b_k(\Phi_0) + \psi_k(X) \\ (k = 1, 2, \dots). \end{aligned} \quad (8.4)$$

Здесь $\psi_k(X)$ — известная функция, которая может содержать отрицательные степени в разложении по степеням x_j . Всегда можно выбрать

$b_k(\Phi_0)$ в виде дробно-рациональной функции так, чтобы существовало дробно-рациональное решение $a_k(X)$. Окончательно решение системы (8.1) приводится к решению уравнений первого порядка (8.3). Коэффициенты этих уравнений являются медленно изменяющимися, так как зависят от переменных $Z(t)$, удовлетворяющих системе (6.3). Как следствие из уравнений вида (8.3) вытекает теорема.

Теорема 3. Если собственные числа $\lambda_1, \dots, \lambda_n$ матрицы J удовлетворяют условию

$$\lambda_1^{m_1} \lambda_2^{m_2} \dots \lambda_n^{m_n} \neq 1 \quad \left(m_j = 0, 1, 2, \dots; \sum_{j=1}^n m_j > 0 \right), \quad (8.5)$$

то существует формальная замена переменных

$$X(t) = W(t) + \mu \Pi_1(W(t)) + \mu^2 \Pi_2(W(t)) + \dots, \quad (8.6)$$

переводящая систему (8.1) в систему вида

$$W(t+h) = J W(t). \quad (8.7)$$

Общее решение системы (8.7) имеет вид

$$W(t) = J^{th-1} Q(t), \quad Q(t+h) \equiv Q(t). \quad (8.8)$$

Здесь $Q(t)$ — произвольный периодический вектор с периодом h . Эта теорема — аналог теоремы Пуанкаре для дифференциальных уравнений. Она дает общий вид аналитического решения системы (8.1) при условии (8.5).

Пример 5. Ищем условия устойчивости решений уравнения [4]

$$y_{n+1} = y_n + 2\mu \cos nh y_n, \quad (8.9)$$

С помощью введения новых переменных можно свести уравнение (8.9) к системе вида (1.7)

$$\begin{aligned} x_1(t+h) &= e^{ih} x_1(t), \quad x_2(t+h) = e^{-ih} x_2(t), \\ x_3(t+h) &= x_3(t) + \mu x_3(t) [x_1(t) + x_2(t)], \\ x_3(nh) &= y_n. \end{aligned}$$

В общем случае, когда $2\pi h^{-1}$ является иррациональным числом, порождающая система при $\mu = 0$ имеет интегралы

$$v_1 = x_3, \quad v_2 = x_1 x_2 \equiv 1 \quad (x_1(t) = e^{it}, \quad x_2(t) = e^{-it}).$$

Построение вспомогательной переменной $z(t)$ (6.2) и уравнения (6.3) приводит к линейным уравнениям

$$z(t) = x_3(t) + 2\mu \operatorname{Re} \frac{e^{it}}{1 - e^{ih}} x_3(t) + 2\mu^2 \operatorname{Re} \frac{e^{2it}}{(1 - e^{ih})(1 - e^{2ih})} + \dots, \quad (8.10)$$

$$z(t+h) = z(t) - \mu^2 z(t) + O(\mu^3) z(t).$$

Так как $z(t) \rightarrow 0$ при $t \rightarrow \infty$, нулевое решение уравнения (8.9) устойчиво. Находим $z(t)$ и, обращая уравнение (8.10), получаем решение y_n

$$\begin{aligned} y_n &= (1 - \mu^2 + \dots)^n \times \\ &\times \left[1 + \mu \frac{\sin \left(nh + \frac{h}{2} \right)}{\sin \left(\frac{h}{2} \right)} + \mu^2 \frac{2 \cos \left(\frac{h}{2} \right) - \cos \left(2nh + \frac{3h}{2} \right)}{2 \sin \left(\frac{h}{2} \right) \sin h} + \dots \right] c. \end{aligned}$$

Если $2\pi h^{-1}$ — рациональное число, то решение отличается, так как появляются новые интегралы.

9. РАЗНОСТНЫЕ НЕАВТОНОМНЫЕ УРАВНЕНИЯ

Рассмотрим разностные уравнения вида

$$X_{k+1} = JX_k + \mu H_1(kh, X_k) + \mu^2 H_2(kh, X_k) + \dots \quad (9.1)$$

или

$$X(t+h) = JX(t) + \mu H_1(t, X(t)) + \mu^2 H_2(t, X(t)) + \dots, \quad (9.2)$$

где проекции векторов $H_j(t, X)$ являются полиномами относительно x_1, \dots, x_n с коэффициентами, экспоненциально зависящими от времени t :

$$H_j(t, X) = \sum H_{j, m_1, \dots, m_n, r} x_1^{m_1} \dots x_n^{m_n} e^{\alpha_{jr} t}. \quad (9.3)$$

Один путь состоит в введении новых переменных

$$x_{jr}(t) = e^{\alpha_{jr} t}.$$

с добавлением в системе (9.2) новых уравнений

$$x_{jr}(t+h) = e^{\alpha_{jr} h} x_{jr}(t). \quad (9.4)$$

Таким путем система (9.2) может быть приведена к виду (1.7). Другой путь состоит в использовании метода расширения. Вводится новый вектор

$$Z(t) = \Phi_0(t, X(t)) + \mu \Phi_1(t, X(t)) + \mu^2 \Phi_2(t, X(t)) + \dots, \quad (9.5)$$

удовлетворяющий системе разностных уравнений

$$Z(t+h) = Z(t) + \mu Q_1(Z(t)) + \mu^2 Q_2(Z(t)) + \dots. \quad (9.6)$$

Вектор $\Phi_0(t, X)$ удовлетворяет системе уравнений

$$\Phi_0(t+h, JX) - \Phi_0(t, X) = 0, \quad (9.7)$$

В качестве проекций $\Phi_0(t, X)$ берем независимые решения $v_1(t, X), \dots, v_q(t, X)$ уравнения

$$\varphi(t+h, JX) - \varphi(t, X) = 0, \quad (9.8)$$

такие, что любое решение $\varphi(t, X)$ полиномиально выражается через $v_1(t, X), \dots, v_q(t, X)$. Отметим, что в качестве $v_j(t, X)$ можно взять интегралы

$$v_j(t, X) \equiv \lambda_j^{-th^{-1}} x_j,$$

где x_j — резонансные переменные. Определение $\Phi_k(t, X)$ приводит к системе уравнений с частными разностями

$$\Phi_k(t+h, JX) - \Phi_k(t, X) - Q_k(\Phi_0) = \Psi_k(t, X) \quad (k = 1, 2, \dots). \quad (9.9)$$

Всегда можно найти полиномиальный вектор $Q_k(Z)$, такой, что система (9.9) имеет полиномиальное решение. Если $H_k(t, 0) \equiv 0$, то можно найти $Q_k(Z)$, такие, что $Q_k(0) = 0$. При $\operatorname{Re} \alpha_{jr} \leq 0$ устойчивость системы (9.6) соответствует устойчивости решения $X = 0$ системы (9.2). Для построения решения вводим вектор W

$$W(t) = X(t) + \mu A_1(t, X(t)) + \mu^2 A_2(t, X(t)) + \dots, \quad (9.10)$$

который удовлетворяет системе разностных уравнений

$$W(t+h) = JW(t) + \mu B_1(Z(t)) W(t) + \mu^2 B_2(Z(t)) W(t) + \dots. \quad (9.11)$$

Здесь $B_j(Z)$ — диагональные матрицы, элементы которых являются дробно-рациональными функциями медленно изменяющихся переменных $Z(t)$. Система (9.11) распадается на уравнения первого порядка и легко интегрируется при известном $Z(t)$.

Пример 6. Ищем решение разностного уравнения

$$x(t+h) = x(t) + \mu x^2(t) e^{-t}. \quad (9.12)$$

Вводим вспомогательную переменную $z(t)$:

$$\begin{aligned} z(t) &= x(t) + \mu \varphi_1(t, x(t)) + \mu^2 \varphi_2(t, x(t)) + \dots, \\ z(t+h) &= z(t) + \mu Q_1(z(t)) + \mu^2 Q_2(z(t)) + \dots. \end{aligned} \quad (9.13)$$

Заменив в (9.13) t на $t+h$ и исключив $z(t+h)$, $z(t)$, $x(t+h)$, приходим к уравнениям для $\varphi_k(t, x)$, $Q_k(z)$. При $k=1$ имеем

$$x^2 e^{-t} + \varphi_1(t+h, x) - \varphi_1(t, x) = Q_1(x),$$

отсюда

$$Q_1(x) \equiv 0, \quad \varphi_1(t, x) = x^2 e^{-t} (1 - e^{-h})^{-1}.$$

Аналогично находим $\varphi_2(t, x)$ и т. д. Получим

$$\begin{aligned} z(t) &= x(t) + \mu \frac{x^2(t) e^{-t}}{1 - e^{-h}} + \mu^2 \frac{x^3(t) e^{-2t-h}}{(1 - e^{-h})(1 - e^{-3h})} + \dots, \\ z(t+h) &= z(t). \end{aligned}$$

Разрешая это уравнение относительно $x(t)$, найдем общее решение уравнения (9.12)

$$x(t) = z(t) - \mu z^2(t) \frac{e^{-t}}{1 - e^{-h}} + 2 \mu^2 z^3(t) \frac{e^{-2t}}{(1 - e^{-h})(1 - e^{-2h})} + \dots.$$

Здесь $z(t)$ — произвольная периодическая функция периода h .

10. ЛИНЕЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ С ПЕРИОДИЧЕСКИМИ КОЭФФИЦИЕНТАМИ

К разностным уравнениям с периодическими коэффициентами приводят разностные приближенные методы [4]. Покажем, что к ним может приводить метод точечных отображений [1]. Рассмотрим систему линейных дифференциальных уравнений с квазипериодическими коэффициентами

$$\begin{aligned} \frac{dy}{dt} &= A(\omega t, \nu t) Y, \quad Y(0) = Y_0, \\ A(t+2\pi, \tau) &\equiv A(t, \tau+2\pi) \equiv A(t, \tau). \end{aligned} \quad (10.1)$$

Рассмотрим более общую систему с дополнительными параметрами α , β :

$$\frac{dX}{dt} = A(\omega t + \alpha, \nu t + \beta) X, \quad X(0) = Y_0. \quad (10.2)$$

Фундаментальную нормированную при $t=0$ матрицу решений при закрепленных ω , ν обозначим через $N(t, \alpha, \beta)$. Ясно, что $N(t, \alpha, \beta)$ — матрица, периодическая по α , β с периодом 2π . При $\alpha = \beta = 0$ придем к решению системы (10.1). Рассмотрим моменты t_k , кратные T . Значения $Y(t)$, $X(t)$ при $t_k = kT$ обозначим через Y_k , X_k :

$$X_k = N(kT, \alpha, \beta) X_0 \quad (k = 0, 1, 2, \dots). \quad (10.3)$$

Принимая $t_k = kT$ за начало отсчета, получим уравнения при $\alpha = \beta = 0$:

$$Y_{k+1} = N(T, \omega T k, \nu T k) Y_k \quad (k = 0, 1, 2, \dots). \quad (10.4)$$

Имеем разностные уравнения с квазипериодическими коэффициентами. Пусть $\omega T = 2\pi$. Придем к системе разностных уравнений с периодическими коэффициентами

$$\begin{aligned} Y_{k+1} &= B(kh) Y_k, \quad B(kh) \equiv N(T, 0, 2\pi\nu^{-1} k), \\ B(t + 2\pi) &\equiv B(t) \quad (k = 0, 1, 2, \dots). \end{aligned} \quad (10.5)$$

Решение системы (10.5) в определенном смысле равносильно решению системы дифференциальных уравнений (10.1) с квазипериодическими коэффициентами. При построении решений возникает проблема «малых знаменателей». Частично возникающие трудности преодолены в работах [5, 6], выполненных под руководством автора. Рассмотрим разностные уравнения в канонической форме

$$\begin{aligned} X(t + h) &= JX(t) + \mu H_1(t) X(t) + \mu^2 H_2(t) X(t) + \dots, \\ H_j(t + 2\pi) &\equiv H_j(t). \end{aligned} \quad (10.6)$$

В силу линейности уравнений вспомогательная система (9.6) и соотношения (9.5) будут линейными. Вводим переменные $Z(t)$:

$$Z(t) = [\Phi_0(t) + \mu\Phi_1(t) + \mu^2\Phi_2(t) + \dots] X(t), \quad (10.7)$$

$$Z(t + h) = Z(t) + \mu Q_1 Z(t) + \mu^2 Q_2 Z(t) + \dots, \quad (10.8)$$

$$Q_k = \text{const.}$$

Размерность q вектора $Z(t)$ определяется числом резонансных переменных. Если собственные числа $\lambda_1, \dots, \lambda_n$ матрицы J удовлетворяют условию $|\lambda_j| \ll 1$, то q — число собственных значений, для которых $|\lambda_j| = 1$. Разностные уравнения (10.8) легко решаются, а поэтому введением новых переменных W (9.10).

$$W(t) = X(t) + \mu A_1(t) X(t) + \mu^2 A_2(t) X(t) + \dots \quad (10.9)$$

сводится к решаемой системе уравнений первого порядка

$$W(t + h) = JW(t) + \mu B_1 W(t) + \mu^2 B_2 W(t) + \dots, \quad (10.10)$$

$$B_k = \text{const.}$$

Если выполнено условие

$$\lambda_j \neq e^{ikh} \lambda_s \quad (j, s = 1, \dots, n; k = 0, \pm 1, \pm 2, \dots), \quad (10.11)$$

то имеем нерезонансный случай и $B_k \equiv 0$ ($k = 1, 2, \dots$). Аналогичные результаты получаются для уравнений с квазипериодическими коэффициентами.

Изложенный способ решения применим к линейным разностным уравнениям с экспоненциальными коэффициентами [7].

Выводы. Приближенное исследование системы разностных уравнений (1.7) вблизи неподвижной точки $X = 0$ зависит от решения вспомогательной системы дифференциальных уравнений (6.8)

$$\frac{dZ}{dt} = Q(Z). \quad (10.12)$$

Порядок этой системы можно взять равным числу резонансных переменных $x_j(t)$, т. е. числу собственных чисел λ_j , входящих в одно из соотношений

$$\lambda_1^{m_1} \lambda_2^{m_2} \dots \lambda_n^{m_n} = 1 \quad (m_j = 0, 1, 2, \dots)$$

при $m_j > 0$. Проекции $Q_1(Z)$ являются многочленами порядка не ниже второго. Отметим, что решение системы (10.12) на промежутке $\tau \in [0, L]$ позволяет построить приближенное решение системы (1.7) на промежутке $t \in [0, L \mu^{-1}]$ с любой конечной степенью точности относительно μ .

ЛИТЕРАТУРА

- 1 Ю. И. Неймарк, Тр. Международного симпозиума по нелинейным колебаниям, 2, изд. АН УССР, Киев, 1963.
- 2 К. Г. Валеев, Изв. высш. уч. зав. — Радиофизика, 11, № 7, 985 (1968).
- 3 И. С. Березин и Н. П. Жидков, Методы вычислений, 2, Физматгиз, М., 1959.
- 4 К. Г. Валеев, Ю. В. Ракитский, сб. Численные методы решения дифференциальных и интегральных уравнений и квадрупольной формулы, изд. Наука, М., 1964.
- 5 А. Н. Жернак, Изв. высш. уч. зав. — Математика, № 7, 48 (1969).
- 6 А. Н. Жернак, Журн. выч. мат. и матем. физ., 10, № 1, 230 (1970).
- 7 К. Г. Валеев, Дифференциальные уравнения, 5, № 3, 499 (1969).

Киевский институт инженеров
гражданской авиации

Поступила в редакцию
10 июня 1970 г

INVESTIGATION OF POINT MAPPINGS FIXED POINTS

K. G. Valeyev

A general asymptotic method for a fixed points stability investigation is given. A way to construct the solution for a set of difference equations in the fixed points vicinity is presented. The results are applied to linear difference equations with variable coefficients.

УДК + 531 383

К ТЕОРИИ ОДНООСНОГО ГИРОСТАБИЛИЗАТОРА. I

Н. Я. Коган

Изучение динамики одноосного гиростабилизатора при учете сухого трения в осях стабилизации и прецессии сведено к исследованию точечных отображений плоскости склейки фазового пространства в себя. Выяснены возможные виды отображений и проведено их частичное исследование.

Уравнения одноосного гиростабилизатора без коррекции с линейной характеристикой разгрузочного двигателя при учете моментов сухого трения в осях стабилизации и прецессии могут быть записаны [1] в виде

$$\dot{u} = \omega - nu + mv - L_1 f(u), \quad \dot{\omega} = -u - L_2 f(w), \quad \dot{v} = w, \quad (1)$$

где $f(x) = \text{sign } x$ при $x \neq 0$ и $-1 \leq f(0) \leq 1$ и $n > 0, m > 0, L_1 > 0, L_2 > 0$. В работах [1, 2] рассмотрены частные случаи системы (1): $L_1 = 0, L_2 \neq 0$ и $L_1 \neq 0, L_2 = 0$ при условии $n \geq m$ и показано, что отрезки состояний равновесия систем, соответствующих этим случаям, асимптотически устойчивы в целом. В настоящей работе рассмотрены те же частные случаи системы (1), но при условии $n < m$.

Для удобства будем обозначать случай $L_1 = 0, L_2 \neq 0$ буквой A и случай $L_1 \neq 0, L_2 = 0$ — буквой B . В случае B уравнения (1) при помощи линейной замены переменных и времени $x = -\sqrt[3]{m}w, y = \sqrt[3]{m^2}v, z = u, \tau = \sqrt[3]{m}t$ сводятся к уравнениям задачи Вышнеградского, исследованной Андроновым и Майером [3]. Целесообразность рассмотрения этого случая в настоящей работе продиктована тем, что обе задачи сводятся к одинаковым по виду точечным отображениям, которые исследуются методами, отличными от применяемых в работе [3].

Перейдем к исследованию системы (1) в случаях A и B . При $n < m$ характеристическое уравнение линейной системы, соответствующей системе (1), имеет один действительный отрицательный корень $p_1 = -k$ и два комплексных сопряженных корня $p_{2,3} = \rho \pm i\sigma$, причем $n < k < m$ и $\rho > 0$. По формулам Виета связь между коэффициентами m, n и корнями характеристического уравнения имеет следующий вид:

$$k\Delta^2 = m, \quad k - 2\rho = n, \quad \Delta^2 - 2k\rho = 1, \quad (2)$$

где $\Delta^2 = \rho^2 + \sigma^2$.

Случай A . Введем замену переменных

$$\xi = \frac{1}{2\rho L_2} (-u + \Delta^2 v - 2\rho w),$$

$$\eta = \frac{1}{2\rho L_2} [(k - 2\rho)u - k\Delta^2 v + 2k\rho w], \quad w = \frac{1}{2\rho L_2} w$$

и, учитывая соотношения (2), запишем систему (1) в новых переменных

$$\dot{\xi} = -k\xi + f(w), \dot{\eta} = -k\dot{\xi} + 2\rho\eta - 2\rho\Delta^2 w, \dot{w} = -\frac{1}{2\rho}\dot{\xi} + \frac{1}{2\rho}\eta. \quad (3)$$

Переменные ξ , η , w выберем в качестве фазовых координат системы (3). Обозначим фазовое пространство буквой R . Плоскость $w = 0$ делит пространство R на два полупространства R_+ ($w > 0$) и R_- ($w < 0$), в каждом из которых поведение изображающей точки описывается линейной системой, получающейся из (3) при $w > 0$ и $w < 0$ соответственно (рис. 1 а). Двигаясь по фазовой траектории из некоторой конеч-

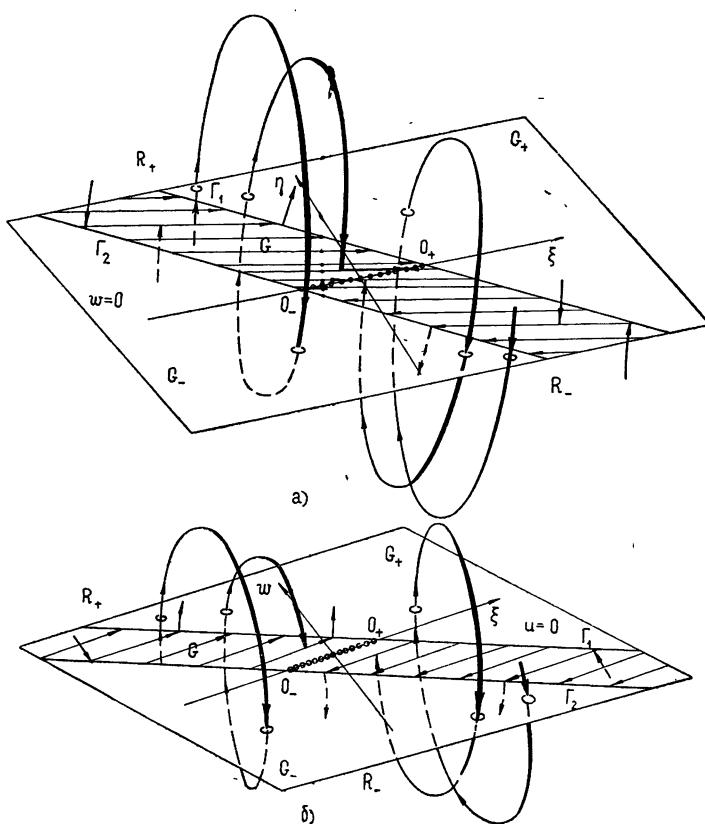


Рис. 1.

ной части R , изображающая точка по истечении конечного промежутка времени попадает на плоскость $w = 0$ и затем при $t \rightarrow \infty$ вновь и вновь пересекает ее [1]. В силу этого исследование системы (3) можно свести к точечному отображению плоскости $w = 0$ в себя [4, 5].

Рассмотрим поведение изображающей точки на плоскости $w = 0$. Прямые $\Gamma_1(\eta + k\xi - 1 = 0)$ и $\Gamma_2(\eta + k\xi + 1 = 0)$ делят плоскость $w = 0$ на три части: G_+ и G_- , где фазовые траектории пересекают плоскость $w = 0$ в направлении из R_- в R_+ и из R_+ в R_- соответственно, и полоску скользящих движений G , где фазовые траектории подходят к плоскости $w = 0$ как из R_+ , так и из R_- . На полоске скользящих движений поведение изображающей точки описывается уравнениями

$$w = 0, \eta - \eta_0 = -(k - 2\rho)(\xi - \xi_0), \eta = \eta_0 \exp[-(k - 2\rho)t].$$

Так как $n > 0$, то из соотношений (2) следует, что $k - 2\rho > 0$, и так как $\rho > 0$, то траектории скользящих движений расположены под большим углом к оси η , чем прямые Γ_1 и Γ_2 . На прямых Γ_1 и Γ_2 поведение изображающей точки определяется следующими условиями: $\Gamma_1 \subset G_+$ при $\eta > 0$, $\Gamma_1 \subset G$ при $\eta < 0$, $\Gamma_2 \subset G_-$ при $\eta < 0$ и $\Gamma_2 \subset G$ при $\eta > 0$. Изображающая точка, двигаясь по траектории скользящих движений, попадает на $\Gamma_1 \subset G_+$ при $\eta > 0$ и на $\Gamma_2 \subset G_-$ при $\eta < 0$. При работе гиростабилизатора в скользящем режиме гироскоп остается неподвижным, а рама поворачивается вокруг оси стабилизации до тех пор, пока гирокопический момент, действующий по оси прецессии, не достигает максимального значения момента трения. На полоске скользящих движений расположен отрезок состояний равновесия системы (3) $O_- O_+$: $w = 0$, $\eta = 0$, $-1/k \leq \xi \leq 1/k$.

Очевидно, что все траектории скользящих движений, расположенные внутри параллелограмма, образованного прямым Γ_1 , Γ_2 и траекториями скользящих движений, проходящими через концевые точки отрезка состояний равновесия, втыкаются в отрезок $O_- O_+$. Обозначим эту

область \tilde{G} . Изображающая точка, двигаясь по траектории области \tilde{G} , асимптотически приближается к соответствующему этой траектории состоянию равновесия. Состоянием равновесия системы (3) соответствуют равномерные вращательные движения рамы гиростабилизатора [1].

Перейдем к общему описанию фазового пространства системы (1) в случае В. Введем новые переменные $\xi = (1/L_1)[-u + \Delta^2 v - (\Delta^2 - 1)w/k]$, $u = (1/L_1)u$, $w = (\Delta^2/L_1)w$. Тогда система (1) запишется следующим образом:

$$\dot{\xi} = -k\xi + f(u), \quad u = w + 2\rho u - \xi, \quad \dot{w} = -u\Delta^2. \quad (4)$$

В качестве фазовых переменных возьмем ξ , u , w . Фазовое пространство системы (4) будем также обозначать буквой R . Плоскость $u = 0$ делит пространство R на две части R_+ ($u > 0$) и R_- ($u < 0$), в каждом из которых поведение изображающей точки описывается линейной системой, получающейся из (4) при $u > 0$ и $u < 0$ соответственно (рис. 1 б). Плоскость $u = 0$ пряммыми $\Gamma_1(w + k\xi - 1 = 0)$ и $\Gamma_2(w + k\xi + 1 = 0)$ разбивается на три части: G_+ и G_- , где фазовые траектории пересекают плоскость $u = 0$ из R_- в R_+ и из R_+ в R_- соответственно, и полоску скользящих движений G , где фазовые траектории подходят к плоскости $u = 0$ как из R_+ , так и из R_- . Уравнение траекторий скользящих движений имеет вид $u = 0$, $w = w_0$, $\xi = w_0 t + \xi_0$. На прямых Γ_1 и Γ_2 поведение изображающей точки определяется следующими условиями: $\Gamma_1 \subset G_+$ при $w > 0$, $\Gamma_1 \subset G$ при $w < 0$, $\Gamma_2 \subset G_-$ при $w < 0$, $\Gamma_2 \subset G_+$ при $w > 0$. Изображающая точка, двигаясь по траекториям скользящих движений, попадает на $\Gamma_1 \subset G_+$ при $w > 0$ и на $\Gamma_2 \subset G_-$ при $w < 0$. При работе гиростабилизатора в скользящем режиме рама остается неподвижной, а гироскоп совершает инерционное движение до тех пор, пока сумма гирокопического момента, возникающего при движении гироскопа вокруг оси прецессии и момента, развиваемого разгрузочным двигателем, не достигнет максимального значения момента трения. На полоске скользящих движений расположена отрезок состояний равновесия системы (4) $O_- O_+$: $u = 0$, $w = 0$, $-1/k \leq \xi \leq 1/k$. Состояниям равновесия соответствуют состояния покоя гиростабилизатора. Заметим, что как в случае А, так и в случае В фазовые траектории симметричны относительно начала координат.

Перейдем к рассмотрению точечных отображений плоскостей $w = 0$

в себя (случай A) и $w = 0$ в себя (случай B), которые получаются в силу фазовых траекторий системы (3) и (4) соответственно. Оба случая будем рассматривать одновременно.

Так как области G_+ и G_- симметричны относительно начала координат и фазовые траектории симметричны относительно начала координат, то достаточно рассмотреть отображение T области G_+ в себя. Обозначая через T_1 отображение плоскости $w = 0$ ($w = 0$) в себя, порождаемое траекториями полупространства R_+ , T_2 , T_3 — отображение в силу траекторий скользящих движений области G в $\Gamma_1 \subset G_+$ при $\eta > 0$ ($w > 0$) и в $\Gamma_2 \subset G_-$ при $\eta < 0$ ($w < 0$) и T_4 — отображение плоскости $w = 0$ ($w = 0$) в себя, порождаемое траекториями полупространства R_- , построим возможные виды отображения области G_+ в себя:

$$T = T_1 T_4, \quad T = T_1 T_3 T_4, \quad T = T_1 T_3 T_4 T_2, \quad T = T_1 T_2, \quad T = T_1 T_4 T_2. \quad (5)$$

Прежде чем перейти к конкретному построению отображений и отысканию областей их определения, докажем следующую лемму.

Лемма 1. Фазовые траектории системы (3) и (4) после конечного числа пересечений плоскости $w = 0$ (случай A) и $w = 0$ (случай B) попадают внутрь слоя $-1/k \leq \xi \leq 1/k$ и в нем остаются.

Доказательство проведем для случая A . В случае B доказательство аналогичное. Интегрируя первое уравнение системы (3) при $w > 0$ и $w < 0$, получим соответственно

$$\xi = \frac{1}{k} + \left(\xi_0 - \frac{1}{k} \right) e^{-kt}, \quad \xi = -\frac{1}{k} + \left(\xi_0 + \frac{1}{k} \right) e^{-kt}. \quad (6)$$

Из (6) следует, что плоскости $\xi = 1/k$ и $\xi = -1/k$, являются устойчивыми интегральными плоскостями линейных систем, получающихся из (3) при $w > 0$ и $w < 0$ соответственно. Следовательно, фазовые траектории системы (3) приближаются к плоскости $\xi = 1/k$ в R_+ и k плоскости $\xi = -1/k$ в R_- , никогда не пересекая их. Будем обозначать участки траекторий системы (3) в R_+ через L_+ , в R_- через L_- и в G через L . Тогда из (6) следует, что изображающая точка, двигающаяся по траектории, склеенной из L_+ и L_- , через конечный промежуток времени окажется внутри слоя $-1/k \leq \xi \leq 1/k$ и по траекториям такого вида из него не выйдет (рис. 2). Если изображающая точка попадает на полоску G , то дальнейшее ее движение будет проходить по траектории L в направлении слоя $-1/k \leq \xi \leq 1/k$, причем если она оказалась внутри этого слоя, то по траектории L выйти из него она не может. Лемма доказана.

Если обозначить часть области G_+ , для которой $-1/k \leq \xi \leq 1/k$, через G'_+ , то в силу доказанной леммы изучение отображения G_+ в себя сводится к изучению отображения G'_+ в себя.

Кроме того, используя симметрию фазовых траекторий относительно начала координат вместо отображения T_4 , будем рассматривать отображение S , переводящее точки области G_- и G при $\eta < 0$ ($w < 0$) в симметричные относительно начала координат точки области G_+ и G при $\eta > 0$ ($w > 0$). При этом неподвижные точки преобразования T и их устойчивость не изменяются, но число возможных видов отображения G'_+ в себя уменьшается до трех [5]:

$$T = T_1 S, \quad T = T_1 S T_2, \quad T = T_1 T_2. \quad (7)$$

Перейдем к построению отображений (7) и нахождению областей

их определения. Будем обозначать исходную точку любого из отображений T , T_1 , T_2 , S через (ξ, η) в случае A и (ξ, w) в случае B . Точку, полученную в результате применения любого из отображений T_1 , T_2 или S , будем отмечать волнистой линией сверху, а результат применения отображения T — чертой сверху. Используя уравнения Γ_1 , из уравнений траекторий скользящих движений найдем выражение для T_2 .

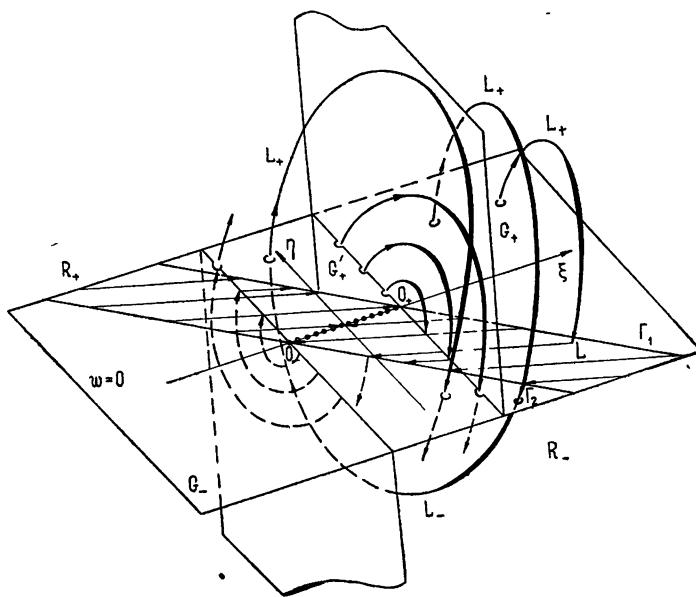


Рис. 2.

Случай A :

$$\tilde{\xi} = -\frac{k-2\rho}{2\rho}\xi - \frac{1}{2\rho}\eta + \frac{1}{2\rho}, \quad \tilde{\eta} = \frac{k(k-2\rho)}{2\rho}\xi + \frac{k}{2\rho}\eta - \frac{k-2\rho}{2\rho}. \quad (8)$$

Случай B :

$$\tilde{\xi} = -\frac{1}{k}w + \frac{1}{k}, \quad \tilde{w} = w. \quad (9)$$

Выражение для отображения T_1 получаем из общего вида решения системы (3) при $w > 0$ в случае A и (4) при $u > 0$ в случае B , выражая постоянные интегрирования через координаты исходной точки. Для случая A имеем

$$\begin{aligned} \tilde{\xi} &= \frac{1}{k} + \left(\xi - \frac{1}{k} \right) e^{-k\tau}, \\ \tilde{\eta} &= C_1 e^{-k\tau} + e^{p\tau} \left[(\eta - C_1) \cos(\sigma\tau) + \frac{p}{\sigma} \left(\frac{\Delta^2 - k^2}{\Delta^2 + k^2} C_1 + \eta \right) \sin(\sigma\tau) \right], \end{aligned} \quad (10)$$

где

$$C_1 = \frac{(k^2 + \Delta^2)(-k\xi + 1)}{a^2}, \quad a^2 = k^2 + 2k\rho + \Delta^2$$

и τ — наименьший положительный корень уравнения

$$\frac{1}{k^2 + \Delta^2} C_1 e^{-k\tau} + e^{\rho\tau} \left[-\frac{1}{\Delta^2 + k^2} C_1 \cos(\sigma\tau) + \right. \\ \left. + \frac{1}{2\rho\sigma} \left(\eta + \frac{\rho^2 - \sigma^2 - k^2}{\Delta^2 + k^2} C_1 \right) \sin(\sigma\tau) \right] = 0. \quad (11)$$

Для случая B имеем

$$\tilde{\xi} = \frac{1}{k} + \left(\xi - \frac{1}{k} \right) e^{-k\tau},$$

$$\tilde{w} = \frac{1}{k} C_1 e^{-k\tau} + e^{\rho\tau} \left[\left(\frac{1}{k} C_1 - w \right) \cos(\sigma\tau) + \left(\frac{\rho}{\sigma} w - \frac{k+\rho}{\sigma} C_1 \right) \sin(\sigma\tau) \right], \quad (12)$$

где $C_1 = \Delta^2 k (-k\xi + 1)/a^2$ и τ — наименьший положительный корень уравнения

$$C_1 e^{-k\tau} + e^{\rho\tau} \left[-C_1 \cos(\sigma\tau) + \left(\frac{\Delta^2}{\sigma} w - \frac{\Delta^2 + k\rho}{k\sigma} C_1 \right) \sin(\sigma\tau) \right] = 0. \quad (13)$$

Введем следующие обозначения: $\beta = (-k\xi + 1)/\eta$ в случае A и $\beta = (-k\xi + 1)/w$ в случае B и, используя эти обозначения, перепишем (10) — (13). В случае A получим

$$\tilde{\xi} = \frac{1}{k} + \left(\xi - \frac{1}{k} \right) e^{-k\tau}, \quad \tilde{\eta} = K_A(\beta, \tau) \eta, \quad (14)$$

где

$$K_A(\beta, \tau) = -e^{\rho\tau} \left[\frac{\sigma^2 + k^2 - \rho^2}{2\rho\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) - \frac{k^2 - 2k\rho + \Delta^2}{2\rho\sigma} \beta \sin(\sigma\tau) \right], \quad (15)$$

$$F_A(\beta, \tau) = e^{-k\tau} + e^{\rho\tau} \left[\frac{k+\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) - \frac{a^2}{2\rho\sigma} \left(1 - \frac{1}{\beta} \right) \sin(\sigma\tau) \right] = 0.$$

В случае B получим

$$\tilde{\xi} = \frac{1}{k} + \left(\xi - \frac{1}{k} \right) e^{-k\tau}, \quad \tilde{w} = K_B(\beta, \tau) w, \quad (16)$$

где

$$K_B(\beta, \tau) = -e^{\rho\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) + \frac{\Delta^2}{k\sigma} (1 - \beta) \sin(\sigma\tau) \right], \quad (17)$$

$$F_B(\beta, \tau) = e^{-k\tau} + e^{\rho\tau} \left[\frac{k+\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) + \frac{a^2}{k\sigma} \left(\frac{1}{\beta} - 1 \right) \sin(\sigma\tau) \right] = 0.$$

Имея аналитические выражения отображений T_1 и T_2 , несложно записать любое из отображений (7). Так, например, отображение $T = T_1 S$ получается из (14), (16) сменой знака в правых частях, т. е. $\tilde{\xi} = -\xi$, $\tilde{\eta} = -\eta$ в случае A и $\tilde{\xi} = -\xi$, $w = -\tilde{w}$ в случае B .

Итак, мы получили, что изучение случаев A и B сводится к исследованию отображений (7) области G'_+ плоскости склейки фазового пространства в себя. Область G'_+ представляет собой полубесконечную полосу $-1/k < \xi < 1/k$, $\eta > 0$ ($w > 0$), ограниченную снизу прямой Γ_1 . Замечая, что в обоих случаях отображения T области G'_+ в себя отличаются только видом T_2 , а T_1 отличаются только видом функций $K(\beta, \tau)$ и $F(\beta, \tau)$, будем в дальнейшем исследовать оба случая вместе. Для

этого введем единые обозначения: $x = \xi$, $y = \eta$ в случае A и $x = \xi$, $y = \omega$ в случае B . Тогда отображения T_1 (14) — (17) запишутся в виде

$$\tilde{x} = \frac{1}{k} + \left(x - \frac{1}{k} \right) e^{-k\tau}, \quad \tilde{y} = K(\beta, \tau)y, \quad (18)$$

где τ — наименьший положительный корень уравнения $F(\beta, \tau) = 0$, а β определяется из соотношения $\beta = (-kx + 1)/y$. Функции $K(\beta, \tau)$ и $F(\beta, \tau)$ определяются из (14), (15) в случае A и из (16), (17) в случае B .

Перейдем к исследованию отображения (18). Из уравнений (15) и (17) следует, что каждому значению $\beta = \text{const}$ отвечает одно определенное значение τ . Точки области G'_+ , для которых $\beta = \text{const}$, лежат на лучах, проведенных из конца отрезка состояний равновесия точки O_+ с координатами $x = 1/k$, $y = 0$. Уравнение семейства лучей $\beta = \text{const}$ имеет вид

$$\beta y = -kx + 1. \quad (19)$$

При изменении β от 0 до 1 лучи (19) выметают всю область G'_+ , причем $\beta = 0$ соответствует лучу $x = 1/k$, а $\beta = 1$ — лучу Γ_1 . Очевидно, что нас интересуют только те точки лучей $\beta = \text{const}$, которые принадлежат области G'_+ , поэтому в дальнейшем, говоря, что рассматриваются точки луча $\beta = \text{const}$, будем иметь в виду только те точки, которые принадлежат G'_+ .

Докажем ряд утверждений, касающихся свойств отображения (18).

1) Функция $\tau = \tau(\beta)$, задаваемая уравнением $F(\beta, \tau) = 0$, взаимно однозначная и монотонная с $\frac{d\tau}{d\beta} > 0$. При изменении β от 0 до 1 τ изменяется от π/σ до $\tau_\Gamma < 2\pi/\sigma$ (Приложение 1).

2) По крайней мере для тех значений β , при которых $K(\beta, \tau) < 0$, функция $K(\beta, \tau)$ меняется монотонно, причем $\frac{dK(\beta, \tau)}{d\beta} > 0$. (Прилож. 2).

Лемма 2. Отображение T_1 преобразует семейство лучей (19) в лучи, имеющие тот же центр, но расположенные вне G'_+ .

Доказательство. Исключая x и y из (18), получим

$$\tilde{\beta}\tilde{y} + \tilde{k}\tilde{x} - 1 = 0, \quad (20)$$

где $\tilde{\beta} = \beta e^{-k\tau}/K(\beta, \tau)$. Так как точкам луча $\beta = \text{const}$ соответствует одно значение τ , то (20) представляет собой уравнение луча с центром в точке O_+ ($x = 1/k$, $y = 0$) и угловым коэффициентом $k/\tilde{\beta}$. Отображение T_1 плоскости склейки фазового пространства в себя порождается траекториями систем (3), (4) в R_+ и, следовательно, переводит точки G_+ в точки областей G и G_- . Отсюда следует, что лучи (20) расположены вне G'_+ . Лемма доказана.

Лучи (19), для которых $K(\beta, \tau) \geq 0$, переходят при применении T_1 в лучи (20) с $\tilde{\beta} > 0$, лежащие на полоске скользящих движений (рис. 3 а). Лучи, для которых $K(\beta, \tau) < 0$, преобразуются в лучи (20) с $\tilde{\beta} < 0$, проходящие через область G и G'_+ (рис. 3 б). Отображением $T = T_1 S$ лучи (19) преобразуются в лучи с центром в точке O_- ($x = -1/k$, $y = 0$), проходящие через G и G'_+ (рис. 3 в). Уравнение

образов лучей (19) при применении отображения $T = T_1 S$ имеет вид $\delta y = kx + 1$, где $\delta = -\tilde{\beta} > 0$.

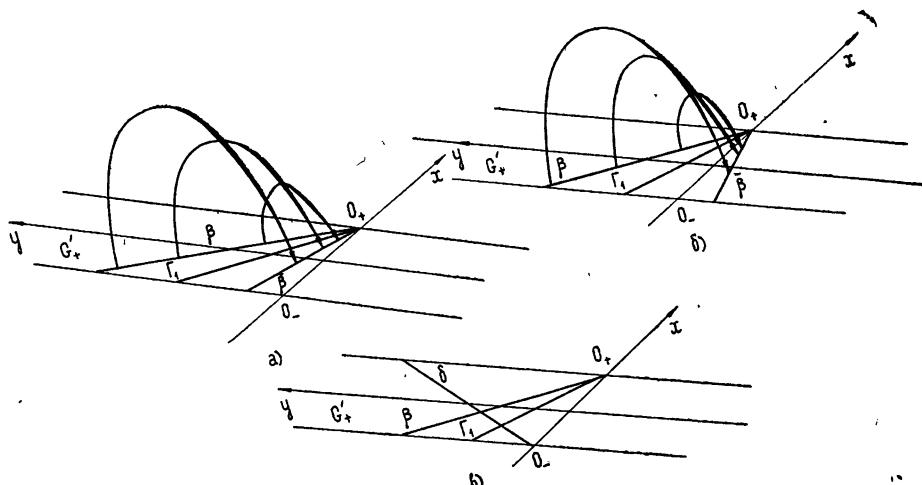


Рис. 3.

Перейдем к отысканию областей определения отображений (7). Прежде всего установим область D_3 определения отображения $T = T_1 T_2$. Так как отображение $T = T_1 T_2$ порождается фазовыми траекториями, попадающими на полосу скользящих движений при $y > 0$, то область D_3 состоит из точек, для которых $K(\beta, \tau) \geq 0$. Границами области определения этого отображения будут Γ_1 и луч $\beta = \underline{\beta}$, для которого $K(\underline{\beta}, \tau) = 0$ (рис. 4 а). Присоединяя к этому уравнению уравнение $F(\underline{\beta}, \tau) = 0$, по-

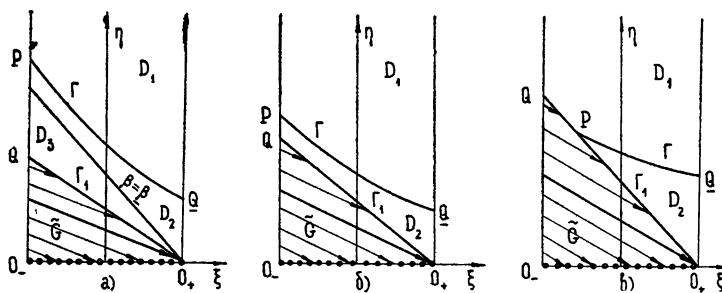


Рис. 4.

лучим систему трансцендентных уравнений для определения $\underline{\beta}$ и τ . Эта система совместна не при всех значениях параметров. Очевидно, что граница в плоскости параметров n и m , отделяющая область существования отображения $T = T_1 T_2$, находится из условия $\underline{\beta} = 1$, $\tau = \tau_r$ и определяется уравнениями $K(\underline{\beta} = 1, \tau_r) = 0$, $K(\underline{\beta} = 1, \tau_l) = 0$. Пусть ω_3 — область в плоскости параметров n и m , для которой существует отображение $T = T_1 T_2$, и l_2 — граница ω_3 . Уравнения, определяющие l_2 , имеют вид для случая А

$$e^{\rho\tau} \left[\frac{k - \rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] = 0,$$

$$e^{-(k+\rho)\tau} + \frac{k+\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) = 0; \quad (21)$$

для случая B

$$\begin{aligned} e^{\rho\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] &= 0, \\ e^{-(k+\rho)\tau} + \frac{k+\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) &= 0. \end{aligned} \quad (22)$$

Из первого уравнения (21) $\operatorname{tg}(\sigma\tau) = \sigma/(k-\rho)$ и $\sigma\tau = \pi + \psi$, где $\psi = \operatorname{arctg}[\sigma/(k-\rho)]$. Подставляя $\sigma\tau = \pi + \psi$ во второе уравнение (21), получим уравнение кривой l_2 :

$$\exp \left[\left(-\frac{k+\rho}{\sigma} \right) \left(\pi + \operatorname{arctg} \frac{\sigma}{k-\rho} \right) \right] = \frac{2\rho}{b},$$

где $b^2 = (k-\rho)^2 + \sigma^2$. Аналогично определим l_2 для случая B . Из первого уравнения (22) определим $\sigma\tau = \pi + \varphi$, где $\varphi = \operatorname{arctg}(\sigma/\rho)$, затем, подставив $\sigma\tau = \pi + \varphi$ во второе уравнение (22), получим $\exp \left(-\frac{k+\rho}{\sigma} \right) \times$
 $\times \left(\pi + \operatorname{arctg} \frac{\sigma}{\rho} \right) = \frac{k}{\Delta}$.

Для построения l_2 в плоскости n, m удобно ввести вспомогательные параметры $\theta = k/\sigma$, $x = \rho/\sigma$, которые связаны с n и m следующими соотношениями:

$$n = \frac{\theta - 2x}{(x^2 - 2\theta x + 1)^{1/2}}, \quad m = \frac{\theta(1+x^2)}{(x^2 - 2\theta x + 1)^{3/2}}.$$

Уравнения l_2 в плоскости вспомогательных параметров записутся следующим образом: для случая A

$$\exp \left[-(\theta + x) \left(\pi + \operatorname{arctg} \frac{1}{\theta - x} \right) \right] = \frac{2x}{[(\theta - x)^2 + 1]^{1/2}},$$

для случая B

$$\exp \left[-(\theta + x) \left(\pi + \operatorname{arctg} \frac{1}{x} \right) \right] = \frac{\theta}{(x^2 + 1)^{1/2}}.$$

Кривые l_2 просчитаны на машине и приведены на рис. 5.

Перейдем к отысканию области определения отображения $T = T_1 S T_2$. Оно существует при всех значениях параметров. Введем в рассмотрение кривую Γ , которая отображением $T_1 S$ переводится в луч Γ_1 . Один конец этой кривой — точка Q — принадлежит лучу $\xi = 1/k$ и отображается в точку Q пересечения Γ_1 с прямой $\xi = -1/k$, причем ордината точки Q большие ординаты точки Q . Это следует из (14) и (16) при $\bar{\xi} = -\tilde{\xi}$, $\bar{\eta} = -\tilde{\eta}$ ($\bar{w} = -\tilde{w}$). Другой конец кривой Γ , именно точка P , находится либо на прямой $x = -1/k$, либо на луче Γ_1 . Область D_2 , ограниченная кривой Γ и лучом $\beta = \beta$ при $n, m \in \omega_3$ либо кривой Γ и лучом или частью луча Γ_1 при $n, m \in \omega_3$, является областью определения отображения $T = T_1 S T_2$ (рис. 4а, б, в). Точки областей D_2 и D_3 отображениями $T = T_1 S T_2$ и $T = T_1 T_2$ соответственно переводятся в точки луча Γ_1 — границы полоски скользящих движений. В остальной части области G'_+ , которую обозначим через D_1 , определено отображение $T = T_1 S$. Область

D_1 представляет собой полубесконечную полосу $-1/k \leq x \leq 1/k$, ограниченную снизу либо кривой Γ , либо кривой Γ и отрезком PQ луча Γ_1 .

Таким образом, исследование уравнений динамики одноосного гиростабилизатора для частных случаев, когда сухое трение учитывается только в оси прецессии, либо только в оси стабилизации, сводится к изучению возможных отображений границы полоски скользящих движений в себя и отображения $T = T_1 S$ полубесконечной полосы D_1 в плоскости склейки в себя. Причем для обоих случаев отображение $T = T_1 S$ имеет вид

$$\bar{x} = -\frac{1}{k} - \left(x - \frac{1}{k} \right) e^{-k\tau}, \quad \bar{y} = K^*(\beta, \tau) y,$$

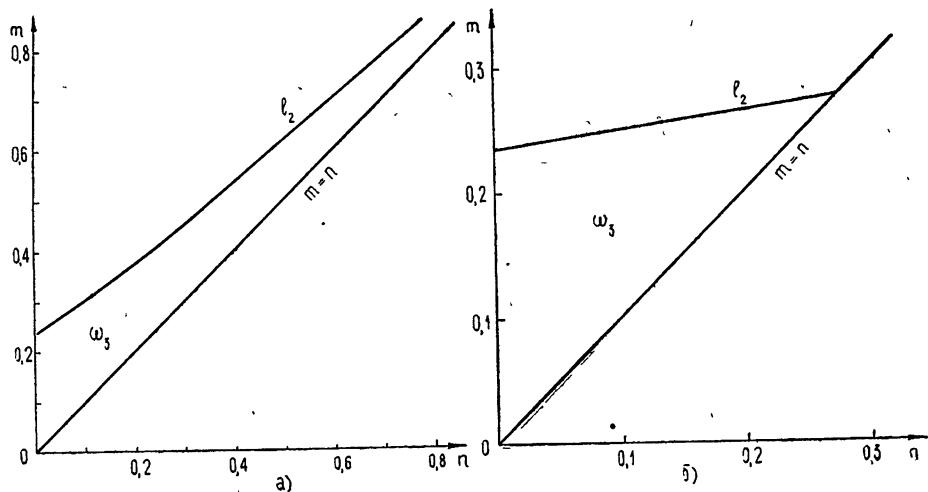


Рис. 5 а) Случай А, б)-случай В

где τ — наименьший положительный корень уравнения $F(\beta, \tau) = 0$, β определяется из соотношения $\beta = (-kx + 1)/y$ и $K^*(\beta, \tau) = -K(\beta, \tau)$.

Приложение 1

Из уравнений (11) и (13) получим, что при $\beta = 0$ $C_1 = 0$ и $\tau = \pi/\sigma$. При $\beta = 1$ значение τ , соответствующее лучу Γ_1 , обозначим через τ_r . Анализируя уравнения (15) и (17) при $\beta = 1$, получим, что $\pi/\sigma < \tau_r < 2\pi/\sigma$. Для оценки производной $\frac{d\tau}{d\beta}$ выразим β через τ из (15) и (17):

случай А

$$\beta = -\frac{a^2}{2\rho\sigma} \frac{\sin(\sigma\tau)}{\Phi_A(\tau)}, \quad (23)$$

случай В

$$\beta = -\frac{a^2}{k\sigma} \frac{\sin(\sigma\tau)}{\Phi_B(\tau)}, \quad (24)$$

где

$$\Phi_A = e^{-(k+\rho)\tau} - \left[\frac{k^2 + \sigma^2 - \rho^2}{2\rho\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right],$$

$$\Phi_B = e^{-(k+\rho)\tau} - \left[\frac{\Delta^2 + k\rho}{k\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right].$$

Дифференцируя (23) и (24) по τ , получим для случая A

$$\frac{d\beta}{d\tau} = -\frac{a^2}{2\rho\sigma} \frac{e^{-(k+\rho)\tau} [(k+\rho)\sigma^{-1} \sin(\sigma\tau) + \cos(\sigma\tau)] - 1}{\Phi_A^2(\tau)},$$

для случая B

$$\frac{d\beta}{d\tau} = -\frac{a^2}{k} \frac{e^{-(k+\rho)\tau} [(k+\rho)\sigma^{-1} \sin(\sigma\tau) + \cos(\sigma\tau)] - 1}{\Phi_B^2(\tau)}.$$

Функция $e^{-(k+\rho)\tau} \left[\frac{k+\rho}{\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right] - 1 < 0$ при $\pi < \sigma\tau < 2\pi$.

Отсюда следует, что $\frac{d\beta}{d\tau} > 0$ и, следовательно, существует однозначная обратная функция $\tau = \tau(\beta)$ ($0 \leq \beta \leq 1$) и $\frac{d\tau}{d\beta} > 0$.

Приложение 2.

Из (14) и (16) получим, что при $\beta = 0$ $K_A = K_B = -e^{-\rho\sigma/\pi}$ и при $\beta = 1$

$$K_A = -\exp(\rho\tau_F) \left[\frac{k-\rho}{\sigma} \sin(\sigma\tau_F) - \cos(\sigma\tau_F) \right],$$

$$K_B = -\exp(\rho\tau_F) \left[\frac{\rho}{\sigma} \sin(\sigma\tau_F) - \cos(\sigma\tau_F) \right].$$

Дифференцируя $K_B(\beta, \tau)$ по β , получим $\frac{dK_B}{d\beta} = \frac{\Delta^2}{k\sigma} \left(1 - \frac{1}{\beta} \right) e^{\rho\tau} \sin(\sigma\tau) \geq 0$ при $0 < \beta \leq 1$. Знак равенства имеет место при $\beta = 1$.

Подставляя (23) в $K_A(\beta, \tau)$, получим

$$K_A(\tau) = -e^{-(k+2\rho)\tau} \frac{f_1(\tau)}{f_2(\tau)}, \quad (25)$$

где

$$f_1(\tau) = e^{-k\tau} + e^{-\rho\tau} \left[\frac{k^2 + \sigma^2 - \rho^2}{2\rho\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right]$$

и

$$f_2(\tau) = f_1(-\tau).$$

Вычисляя непосредственно $\frac{dK_A}{d\tau}$ при $\tau = \pi/\sigma$ и $\tau = \tau_F$, получим

$$\frac{dK_A}{d\tau} \Big|_{\tau=\pi/\sigma} = -\frac{k^2 + \sigma^2 - 3\rho^2}{2\rho} K_A\left(\frac{\pi}{\sigma}\right) > 0$$

и

$$\frac{dK_A}{d\tau} \Big|_{\tau=\tau_F} = -(k - 2\rho) K_A(\tau_F).$$

Знак производной при $\tau = \tau_g$ определяется знаком $K_A(\tau_g)$, который может быть как положительным, так и отрицательным.

Покажем, что, по крайней мере, для тех значений τ , для которых $K_A(\tau) \leq 0$. $\frac{dK_A}{d\tau} \geq 0$. Введем функцию $\Phi(\tau) = e^{(k-2\rho)\tau} K_A(\tau) = -f_1(\tau)/f_2(\tau)$. Дифференцируя $\Phi(\tau)$, получим

$$\frac{d\Phi}{d\tau} e^{-(k-2\rho)\tau} = (k-2\rho)K_A(\tau) + \frac{dK_A}{d\tau}. \quad (26)$$

Покажем, что $\frac{d\Phi}{d\tau} \geq 0$ при $\pi/\sigma \leq \tau \leq \tau_g$, причем знак равенства имеет место при $\tau = \tau_g$:

$$\left. \frac{d\Phi}{d\tau} \right|_{\tau=\pi/\sigma} = -\frac{c^2}{2\rho} K_A\left(\frac{\pi}{\sigma}\right) e^{(k-2\rho)\pi} > 0, \quad \left. \frac{d\Phi}{d\tau} \right|_{\tau=\tau_g} = 0,$$

где $c^2 = k^2 - 2k\rho + \Delta^2$.

Для оценки знака $\frac{d\Phi}{d\tau}$ запишем ее в виде

$$\frac{d\Phi}{d\tau} = -\frac{f(\tau)}{f_2^2(\tau)},$$

где

$$f(\tau) = f'_1(\tau)f_2(\tau) - f_1(\tau)f'_2(\tau).$$

На концах интервала изменения $\tau \in [\pi/\sigma, \tau_g]$ $f(\tau)$ отрицательна и равна нулю соответственно. Производная

$$f'(\tau) = \frac{-(k^2 + 1)a^2}{\rho\sigma} [ch(k + \rho)\tau + \cos(\sigma\tau)] \sin(\sigma\tau) > 0$$

и, следовательно, при всех $\pi/\sigma \leq \tau \leq \tau_g$ $f(\tau) \leq 0$ и $\frac{d\Phi}{d\tau} \geq 0$, причем знак равенства имеет место при $\tau = \tau_g$. Из (26) следует, что для всех значений τ , при которых $K_A(\tau) \leq 0$, $\frac{dK_A}{d\tau} \geq 0$, т. е. в области отрицательности $K_A(\tau)$ меняется монотонно.

ЛИТЕРАТУРА

- Н. Я. Коган, Изв. высш. уч. зав. — Приборостроение, 11, № 12, 59 (1968).
- Н. Я. Коган, Изв. высш. уч. зав. — Приборостроение, 12, № 3, 79 (1969).
- А. А. Андronov, А. Г. Майер, Задача Вышнеградского в теории прямого регулирования, Собрание трудов, изд. АН СССР, М., 1956.
- А. А. Андронов, А. А. Витт, С. Э. Хайкин, Теория колебаний, Физматгиз, М., 1959.
- Ю. И. Неймарк, Изв. высш. уч. зав. — Радиофизика, 1, № 1, 41 (1958).

TO THE THEORY OF A SINGLE-AXIS GYROSTABILIZER. I

N. Ya. Kogan

An investigation of single-axis gyrostabilizer dynamics with regard of dry friction in stabilization and precession axes is reduced to an investigation of phase space gluing plane point transformation into itself. Possible kinds of mappings are found and partial investigation of them is carried out.

УДК 531.383

К ТЕОРИИ ОДНООСНОГО ГИРОСТАБИЛИЗАТОРА. II

Н. Я. Коган

Исследуется динамика одноосного гиростабилизатора с сухим трением в оси прецессии или в оси стабилизации. Исследование проводится методом точечных преобразований

В настоящей работе путем исследования точечных отображений, полученных в [1], полностью выяснена структура разбиения на траектории фазового пространства систем, описывающих динамику одноосного гиростабилизатора с сухим трением в оси прецессии или в оси стабилизации. Показано, что либо имеет место глобальная неустойчивость, либо имеется единственный устойчивый элемент притяжения в виде устойчивого отрезка состояний равновесия с некоторой областью притяжения. Произведено разбиение плоскости параметров на области условной устойчивости и неустойчивости. В случае условной устойчивости определена область притяжения отрезка состояний равновесия и дан алгоритм отыскания ее границ.

Прежде чем приступить к исследованию точечных отображений, сформулируем некоторые необходимые в дальнейшем сведения, полученные в работе [1].

1) Исследование динамики рассматриваемого одноосного гиростабилизатора с трением сводится к изучению отображений T_1T_2 и T_1ST_2 , граници Γ_1 полоски скользящих движений в себе и отображения T_1S области плоскости склейки в себя. В силу леммы 1 предшествующей работы эти точечные отображения достаточно исследовать только в полосе $-1/k \leq x \leq 1/k$, $y > 0$.

2) Уравнение прямой Γ_1 имеет вид [1]

$$y = -kx + 1. \quad (1)$$

3) Отображение T_1 имеет вид

$$\tilde{x} = 1/k + (x - 1/k) e^{-k\tau}, \quad \tilde{y} = K(\beta, \tau) y, \quad (2)$$

где τ — наименьший положительный корень уравнения $F(\beta, \tau) = 0$, в котором $\beta = (-kx + 1)/y$. Вид функций $F(\beta, \tau)$ и $K(\beta, \tau)$ приведен в работе [1]. Отображение T_1S получается из (2) сменой знака в правых частях. Отображение T_2 в случае A имеет вид

$$\begin{aligned} \tilde{x} &= -\frac{k - 2\rho}{2\rho} x + \frac{1}{2\rho} y + \frac{1}{2\rho}, \\ \tilde{y} &= \frac{k(k - 2\rho)}{2\rho} x + \frac{k}{2\rho} y - \frac{k - 2\rho}{2\rho} \end{aligned} \quad (3)$$

и в случае B

$$\tilde{x} = -\frac{1}{k} y + \frac{1}{k}, \quad \tilde{y} = y. \quad (4)$$

4) Для точек прямой Γ_1 имеем $\beta = 1$, $\tau = \tau_\Gamma$ и уравнение $F(\beta, \tau) = 0$ принимает вид

$$e^{-k\tau} + e^{\rho\tau} \left[\frac{k + \rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] = 0, \quad (5)$$

а функция $K(\beta, \tau)$ в случаях A и B равна соответственно

$$\begin{aligned} K_\Gamma &= -\exp(\rho\tau_\Gamma) \left[\frac{k - \rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right], \\ K_\Gamma &= -\exp(\rho\tau_\Gamma) \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right]. \end{aligned} \quad (6)$$

1. ИССЛЕДОВАНИЕ ОТОБРАЖЕНИЙ $T_1 T_2$ И $T_1 S T_2$

При значениях параметров из области ω_3 плоскости параметров n, m ($n, m \in \omega_3$) прямая Γ_1 принадлежит области D_3 определения отображения $T_1 T_2$ [1]. Согласно леммы 2 предшествующей работы луч Γ_1 отображением T_1 переводится в луч $\tilde{\Gamma}_1$ ($\tilde{\beta}y = -kx + 1$), где $\tilde{\beta} = \exp \times (-k\tau_\Gamma) K_\Gamma^{-1}$, причем $0 < K_\Gamma < 1$. Неравенство $K_\Gamma < 1$, как необходимое условие, следует из утверждений леммы 2 и условия $\tilde{x} > x$, выполняющегося в силу леммы 1. Следовательно, $\tilde{\Gamma}_1$ представляет собой луч, целиком лежащий на полоске скользящих движений. В случае A при $k \exp(k\tau_\Gamma) K_\Gamma < k - 2\rho$ $\tilde{\Gamma}_1$ принадлежит области \tilde{G} притяжения отрезка состояний равновесия $O_- O_+$ и в силу траекторий скользящих движений отображается в точки этого отрезка.

В случае A при $k \exp(k\tau_\Gamma) K_\Gamma > k - 2\rho$ и в случае B точки луча $\tilde{\Gamma}_1$ отображением T_2 переводятся в точки луча Γ_1 , причем $\bar{y} < y$. Следовательно, итерации точек луча Γ_1 при последовательном применении отображения $T_1 T_2$ сходятся к концевой точке O_+ отрезка состояний равновесия. При значениях параметров n, m , принадлежащих границе l_2 области ω_3 ($m, n \in l_2$), луч $\beta = \underline{\beta}$, для которого $K(\beta, \tau) = 0$, совпадает с Γ_1 и область D_1 вырождается в прямую Γ_1 . Все точки луча Γ_1 отображением T_1 переводятся в точки отрезка $O_- O_+$.

При $n, m \in \omega_3$ либо луч Γ_1 принадлежит области D_2 определения отображения $T_1 S T_2$, либо отрезок луча Γ_1 от точки O_+ до P принадлежит D_2 , а от P до Q области D_1 определения отображения $T_1 S$ [1]. При $n, m \notin \omega_3$ $K_\Gamma < 0$. Будем обозначать $-K_\Gamma$ через K_Γ . В силу леммы 2 луч Γ_1 отображением $T_1 S$ переводится в луч $\tilde{\Gamma}_1$, уравнение которого $\delta y = kx + 1$, где $\delta = \exp(-k\tau_\Gamma) K_\Gamma^{-1}$. Рассмотрим два случая $0 < K_\Gamma < 1$ и $K_\Gamma > 1$. Области значений параметров, соответствующих этим случаям, обозначим ω_2 и ω_1 . Области ω_2 и ω_1 разделяются кривой l_1 , для точек которой $K_\Gamma = 1$. Из (5) и (6) получаем уравнение для l_1 :

Случай A

$$\begin{aligned} e^{\rho\tau} \left[\frac{k - \rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] &= 1, \\ e^{-k\tau} + e^{\rho\tau} \left[\frac{k + \rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] &= 0. \end{aligned} \quad (7)$$

Случай B

$$e^{\rho\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] = 1,$$

$$e^{-k\tau} + e^{\rho\tau} \left[\frac{k+\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] = 0. \quad (8)$$

Кривые l_1 просчитаны и приведены на рис. 1.

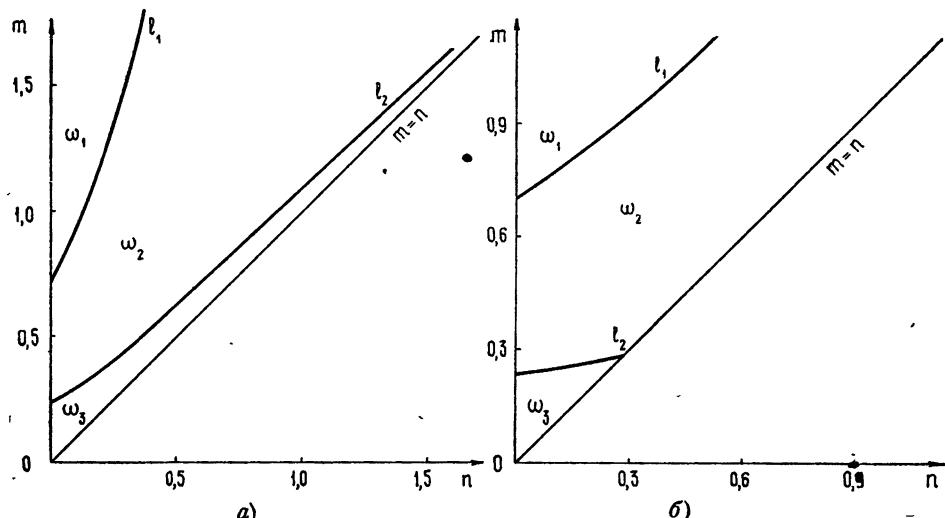


Рис. 1. а) Случай A; б) случай B.

Перейдем к рассмотрению отображения луча Γ_1 в себя при $n, m \in \omega_2$. Если ордината точки Q меньше ординаты точки P , т. е. $y_P \geq y_Q$, то Γ_1 отображением T_1S переводится в отрезок $O_- \bar{Q}$ луча $\tilde{\Gamma}_1$, целиком лежащий на полоске скользящих движений. В силу траекторий скользящих движений отрезок $O_- \bar{Q}$ отображается либо в отрезок $O_- O_+$, что может иметь место для точек, принадлежащих \tilde{G} в случае A, либо в Γ_1 , причем $y < y$. Следовательно, при неоднократном применении отображения T_1ST_2 последовательные итерации точек луча Γ_1 сходятся к отрезку состояний равновесия $O_- O_+$, причем в случае B все итерации сходятся к точке O_+ . Если точка P лежит на Γ_1 , т. е. $y_P < y_Q$, то на отрезке $O_+ P$ луча Γ_1 определено отображение T_1ST_2 , а на отрезке PQ отображение T_1S . Покажем, что в этом случае все последовательные итерации точек луча Γ_1 сходятся к отрезку $O_- O_+$. Точка P отображением T_1S переводится в точку \bar{P} пересечения Γ_1 и $\tilde{\Gamma}_1$, причем $y_{\bar{P}} < y_P$. Следовательно, все точки отрезка $O_+ P$ переходят в точки отрезка $O_- \bar{P}$, целиком лежащего в полоске скользящих движений, и в силу траекторий скользящих движений точки $O_- \bar{P}$ переходят либо на отрезок $O_- O_+$, либо на Γ_1 с уменьшением ординаты. Отрезок PQ луча Γ_1 переводится отображением T_1S в отрезок $\bar{P}\bar{Q}$ луча $\tilde{\Gamma}_1$, где $\bar{Q} = (T_1S)Q$. В свою очередь, отрезок $\bar{P}M$, где M — точка пересечения $\tilde{\Gamma}_1$ с Γ , отображением T_1ST_2 переводится в Γ_1 и так далее, т. е. в этом случае для точек луча Γ_1 , принадлежащих области D_1 , возможно отображение $T_1ST_1ST_2$ части луча Γ_1 в себя. Если выполняется условие

$$K_F K_M < 1, \quad (9)$$

где K_M — значение функции $K(\beta, \tau)$ в точке M , то после конечного числа повторений отображения $T_1ST_1ST_2$ последующие итерации точек попадают на отрезок O_+P определения отображения T_1ST_2 . Докажем условие (9) от противного. Пусть существует точка (n_1, m_1) в области ω_2 плоскости параметров, в которой условие (9) не выполняется, т. е. $K_{\Gamma} K_M > 1$. Рассмотрим параметры (n_0, m_0) , для которых точка P совпадает с точкой Q . Из очевидных геометрических рассуждений следует, что в этом случае $K_{\Gamma} K_M < 1$. Соединим точку (n_0, m_0) с точкой (n_1, m_1) непрерывной кривой l . В силу того, что отображения в рассматриваемых задачах порождаются фазовыми траекториями линейных систем, точечные отображения непрерывно зависят от параметров. Отсюда следует, что на кривой l существует по крайней мере одна точка (n, m) , для которой $K_{\Gamma} K_M = 1$. Так как кривая Γ и отрезок PQ луча Γ_1 принадлежат области определения отображения T_1S , то полученное условие означает, что M является двукратной неподвижной точкой отображения T_1S . В разделе 2 будет доказано, что отображение T_1S не может иметь кратных неподвижных точек.

При $n, m \in \omega_1$ $K_{\Gamma} > 1$. Точка P лежит на Γ_1 и $y_P > y_P$. На отрезке O_+P прямой Γ_1 определено отображение T_1ST_2 , на остальной части Γ_1 — отображение T_1S .

Рассмотрим сначала случай A . Покажем, что на отрезке O_+P существует простая неподвижная точка отображения T_1ST_2 . Из (2) и (3) находим функцию последования

$$\bar{y} = \left[\frac{k - 2\rho}{2\rho} \exp(-k\tau_{\Gamma}) + \frac{k}{2\rho} K_{\Gamma} \right] y - \frac{k - 2\rho}{\rho}. \quad (10)$$

Функция (10) приведена на рис. 2. Точка $y^* = \frac{k-2\rho}{\rho} \left[\frac{k-2\rho}{2\rho} \times \exp(-k\tau_{\Gamma}) + \frac{k}{2\rho} K_{\Gamma} - 1 \right]^{-1}$ пересечения прямой (10) с прямой $\bar{y} = y$

будет неподвижной точкой отображения T_1ST_2 прямой Γ_1 в себя только тогда, когда она принадлежит области определения этого отображения, т. е. $y^* < y_P$. Из (2) найдем $y_P = 2 [K_{\Gamma} + \exp(-k\tau_{\Gamma})]^{-1}$. Сравнивая y^* и y_P , после несложных вычислений получим, что $y^* < y_P$ при $K_{\Gamma} > 1$.

Следовательно, при $n, m \in \omega_1$ на Γ_1 существует неустойчивая неподвижная точка отображения T_1ST_2 . Точки луча Γ_1 , для которых $y < y^*$, при неоднократном применении отображения T_1ST_2 приближаются к отрезку O_-O_+ и на некотором конечном шаге попадают в область G . Точки, для которых $y^* < y$, после конечного числа итераций покидают луч Γ_1 и переходят в область D_1 определения отображения T_1S .

Перейдем к рассмотрению случая B . Так как при отображении T_2 значение ординаты точки не меняется (4), то при $n, m \in \omega_1$ точки отрезка O_+P отображением T_1ST_2 переводятся в точки Γ_1 с $\bar{y} > y$. Следовательно, при неоднократном применении отображения T_1ST_2 итерации точек луча Γ_1 удаляются от отрезка O_-O_+ и покидают область D_2 , переходя в область D_1 определения отображения T_1S . Отрезок состояний равновесия в этом случае неустойчив.

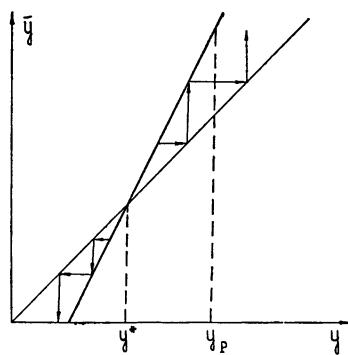


Рис. 2

2. ИССЛЕДОВАНИЕ ТОЧЕЧНОГО ОТОБРАЖЕНИЯ T_1S

Точечное отображение $T = T_1S$ существует при всех значениях параметров $m > n$ и определено в области D_1 плоскости склейки фазового пространства. Область D_1 представляет собой полосу $-1/k \leq x \leq 1/k$, $y > 0$, ограниченную снизу кривой Γ , либо кривой Γ и отрезком PQ луча Γ_1 . Аналитическое выражение для отображения T_1S получено в [1] и имеет вид

$$\bar{x} = -1/k - (x - 1/k)e^{-k\tau}, \quad \bar{y} = K(\beta, \tau)y, \quad (11)$$

где τ — наименьший положительный корень уравнения $F(\beta, \tau) = 0$, в котором $\beta = (-kx + 1)/y$ и $K(\beta, \tau) = -K(\beta, \tau)$.

Используя то, что точечное отображение T_1S лучи $\beta = \text{const}$ ($\beta y = -kx + 1$) переводят в лучи $\delta = \text{const}$ ($\delta y = kx + 1$), из (11) несложно установить, что при $K_g < 1$ в области D_1 существует единственная простая неустойчивая неподвижная точка отображения T_1S . Исследовать тип неподвижной точки трудно в силу сложности выражений коэффициентов характеристического полинома. Числовые просчеты на ЭВМ для нескольких значений параметров показали, что неподвижная точка седлового типа. Однако этого результата недостаточно для полного выяснения структуры разбиения фазового пространства на траектории. Необходимо выяснить вопрос о существовании других предельных элементов фазового пространства. В настоящем разделе показано, что в области D_1 существует кривая, которая является единственным асимптотически устойчивым многообразием точечного отображения T_1S , и, следовательно, отображение области D_1 в себя сводится к исследованию отображения линий в себя.

Покажем, что в области D_1 существует и единственno асимптотически устойчивое многообразие точечного отображения T_1S . Рассмотрим множество кривых Ω вида

$$y = f(x), \quad (y, x) \in D_1, \quad (12)$$

удовлетворяющих условиям

$$\frac{df}{dx} \geq \frac{q}{\gamma} \quad \left(\frac{q}{\gamma} > 0 \right); \quad \frac{df}{dx} \geq 0 \quad \left(\frac{q}{\gamma} < 0 \right). \quad (13)$$

Выражения для q и γ определяются из соотношений

$$d\bar{x} = -qdx + \gamma dy, \quad d\bar{y} = \delta dx + pdy, \quad (14)$$

где

$$q(\beta) = e^{-k\tau} \left(1 - k\beta \frac{d\tau}{d\beta} \right), \quad \gamma(\beta) = \beta^2 e^{-k\tau} \frac{d\tau}{d\beta},$$

$$\delta(\beta) = -k \frac{dK}{d\beta}, \quad p(\beta) = K - \beta \frac{dK}{d\beta},$$

причем $\delta, p > 0$ и $\gamma \geq 0$. Так как $\frac{d\tau}{d\beta} > 0$ и в области D_1 $K(\beta, \tau) > 0$, то $\frac{dK}{d\beta} \leq 0$ [1]. Точечное отображение T_1S порождает некоторое отображение кривых множества Ω . Введем метрику в множестве кривых Ω . Под расстоянием между двумя кривыми $y = f_1(x)$ и $y = f_2(x)$ будем понимать наибольшее значение разности абсолютных точек кривых, лежащих на одном и том же луче $\beta = \text{const}$:

$$\rho(f_1, f_2) = \sup_{\beta \in D_1} |x_1(\beta) - x_2(\beta)|. \quad (15)$$

Докажем следующую лемму.

Точечное отображение T_1S при выполнении условия $p - q \geq 0$ отображает множество кривых Ω в себя и в метрике (15) является сжимающим.

Доказательство. В силу того, что отображение T_1S порождается траекториями линейной системы, образ кривой $y = f(x)$ после применения отображения T_1S представим в виде (12). Пусть кривая $J(y = f(x))$ отображением T_1S переводится в кривую $\bar{J}(y = \bar{f}(x))$. Подставляя $y = f(x)$ и $y = \bar{f}(x)$ в (14) и учитывая, что $\frac{dx}{dx} = -q + \gamma \frac{df}{dx}$, найдем

$$\frac{d\bar{f}}{dx} = \left(\delta + p \frac{df}{dx} \right) \left(-q + \gamma \frac{df}{dx} \right)^{-1}.$$

Требуя выполнения условий (13) для кривой \bar{J} , получим $(p - q) \frac{df}{dx} \geq \frac{df}{dx} \geq -\frac{q^2}{\gamma} - \delta$. Очевидно, что при $p - q \geq 0$ условия (13) выполняются и $\bar{J} \subset \Omega$.

Докажем сжимаемость отображения T_1S множества Ω в себя в метрике (15). Пусть кривые J_1 и J_2 преобразуются в \bar{J}_1 и \bar{J}_2 . Рассмотрим произвольный луч $\beta = \beta_1$, пересекающий кривые \bar{J}_1 и \bar{J}_2 в точках (\bar{x}_1, \bar{y}_1) и (\bar{x}_2, \bar{y}_2) соответственно. Пусть точки (\bar{x}_1, \bar{y}_1) , (\bar{x}_2, \bar{y}_2) переводятся отображением T_1S в точки (\bar{x}_1, \bar{y}_1) и (\bar{x}_2, \bar{y}_2) кривых \bar{J}_1 и \bar{J}_2 . Точки (\bar{x}_1, \bar{y}_1) и (\bar{x}_2, \bar{y}_2) лежат на линии $\delta y = kx + 1$, где $\delta = \beta_1 e^{-k\tau} K^{-1}(\beta, \tau)$ [1].

Проведем луч $\delta y = -kx + 1$ через любую из точек (\bar{x}_1, \bar{y}_1) или (\bar{x}_2, \bar{y}_2) , например, (\bar{x}_1, \bar{y}_1) . Пусть этот луч пересекает кривую \bar{J}_2 в точке (\bar{x}_2, \bar{y}_2) . Так как углы наклона касательных к кривым множества Ω неотрицательны, то из очевидных геометрических соображений следует $|\bar{x}_2 - \bar{x}_1| \leq |\bar{x}_2 - \bar{x}_1|$. С другой стороны, из (11) следует $x_2 - x_1 = e^{-k\tau} (x_2 - x_1)$. Используя это соотношение, приходим к неравенству $|\bar{x}_2 - \bar{x}_1| \leq e^{-k\tau} |x_2 - x_1|$, откуда в силу произвольности $\beta = \beta_1$ получим

$$\rho(\bar{J}_1, \bar{J}_2) \leq \exp\left(-\frac{k}{\sigma}\pi\right) \rho(J_1, J_2). \quad (16)$$

Из (16) следует сжимаемость рассматриваемого отображения кривых Ω в себя с коэффициентом сжатия, не превосходящим $\exp\left(-\frac{k}{\sigma}\pi\right)$.

Лемма доказана. Путем непосредственного исследования функции $p - q$ несложно установить, что для $(x, y) \in D_1$ $p - q > 0$, следовательно, утверждения леммы справедливы в нашем случае (Приложение 1).

Так как область D_1 замкнутая и множество Ω полное, то в силу принципа сжатых отображений в Ω существует единственная инвариантная по отношению к отображению T_1S кривая J^- [2]. Кривая J^- представима в виде (12), имеет в каждой точке касательную с положительным углом наклона и при $x \rightarrow x^\infty = -1/k \operatorname{th}(k\pi/2\sigma)$ имеет вертикальную асимптоту. Снизу J^- определена вплоть до границы области D_1 . Из неравенства (16), при подстановке в него J^- вместо, например, J_1 , следует, что J^- — асимптотически устойчивое многообразие отображе-

ния T_1S . Следовательно, при неоднократном применении отображения T_1S к любой точке D_1 последовательность итераций либо асимптотически приближается к J^- , либо на некотором конечном шаге покидает область D_1 . Таким образом, изучение отображения области D_1 в себя сводится к исследованию отображения кривой J^- в себя.

Перейдем к исследованию отображения T_1S кривой J^- в себя. Так как в D_1 $K(\beta, \tau) > 0$, $\frac{dK}{d\tau} < 0$ и $K(\beta=0, \tau=\pi/\sigma)=\exp\left(\frac{\rho}{\sigma}\pi\right)$, то характер

поведения последовательности итераций точек кривой J^- при неоднократном применении отображения T_1S определяется величиной K_g . Рассмотрим два случая.

1) $K_g > 1$. В этом случае $K(\beta, \tau) > 1$ для всех $(x, y) \in D_1$, и в силу (11) после применения отображения T_1S ордината каждой последующей итерации увеличивается, т. е. $T^n(x, y) \rightarrow (x^\infty, \infty)$ при $n \rightarrow \infty$.

2) $K_g < 1$. В этом случае в силу условия $\frac{dK}{d\beta} < 0$ существует единственный луч $\beta = \beta^*$, для которого $K(\beta^*, \tau^*) = 1$. Точка $M^*(x^*, y^*)$ пересечения луча β^* с кривой J^- является простой неподвижной точкой преобразования T_1S . Очевидно, что по отношению к преобразованию J^- в себя это неустойчивая неподвижная точка, так как при $\beta > \beta^*$ $K(\beta, \tau) < 1$, а при $\beta < \beta^*$ $K(\beta, \tau) > 1$. С другой стороны, несложно показать, что неподвижная точка по отношению к преобразованию T_1S области D_1 в себя является неподвижной точкой седлового типа. Докажем это.

Предположим, что точка M^* — устойчивая неподвижная точка отображения $(T_1S)^{-1}$. Тогда существуют окрестности неподвижной точки, ограниченные замкнутыми кривыми без контакта $V(x, y) = c$, для которых $V(x, y) > 0$ и $V(T(x, y)) - V(x, y) > 0$, но это противоречит тому, что кривая, проходящая через M^* , является асимптотически устойчивым многообразием отображения T_1S . В силу полученного противоречия неподвижная точка M^* имеет седловой характер, и, следовательно, через нее проходят две сепаратрисы, одной из которых является кривая J^- , а другая J^+ состоит из точек, асимптотически приближающихся к M^* при неоднократном применении отображения T_1S . Кривая J^+ ограничивает область, точки которой при неоднократном применении отображения T_1S покидают область D_1 и переходят в области определения отображений T_1T_2, T_1ST_2 .

3. БИФУРКАЦИЯ НЕПОДВИЖНОЙ ТОЧКИ ОТОБРАЖЕНИЯ T_1S

При $n, m \in \omega_2 \cup \omega_3$ у отображения T_1S существует единственная простая неподвижная точка седлового типа. Рассмотрим бифуркацию неподвижной точки M^* при переходе параметров через кривую l_1 , разделяющую области ω_1 и ω_2 в плоскости параметров. Пусть $n, m \in \omega_2 \cup \omega_3$ и по некоторой непрерывной кривой приближаются к l_1 . В этом случае неподвижная точка по J^- смещается к нижней границе области D_1 . При $n, m \in l_1$ неподвижная точка совпадает с точкой P на луче Γ_1 , отрезок PQ которого вместе с кривой Γ ограничивает в этом случае область D_1 снизу. Сепаратриса J^+ при этом исчезает, сливаясь с неподвижной точкой.

Сначала рассмотрим бифуркацию неподвижной точки в случае A. При переходе параметров через l_1 в область ω_1 неподвижная точка покидает область D_1 и смещается по Γ_1 в область определения отображения T_1ST_2 . Устойчивость неподвижной точки при этом не изменяется. Область притяжения отрезка O_-O_+ в этом случае в пределах полосы

$-1/k < x < 1/k$ ограничена траекторией L^* скользящего движения, проходящей через M^* , и кривой Γ' , которая отображением T_1S переводится в L^* . Таким образом, в случае A имеет место простая бифуркация неподвижной точки, связанная с переходом ее через границу области определения точечного отображения [3].

Перейдем к рассмотрению бифуркации неподвижной точки в случае B . При $n, m \in I_1$ неподвижная точка из области D_1 переходит на границу полоски скользящих движений. При этом один из корней характеристического полинома, от расположения корней которого зависит устойчивость неподвижной точки, обращается в 0, а другой — в +1. Отрезок луча Γ_1 от O_- до P состоит из неподвижных точек отображения T_1ST_2 . Этот континuum неподвижных точек неустойчиво по отношению к преобразованию T_1S . При дальнейшем изменении параметров в направлении ω_1 неподвижные точки исчезают и одновременно отрезок состояний равновесия меняет свою устойчивость. Таким образом, в этом случае происходит совпадение двух видов бифуркаций: бифуркации при переходе параметров через границу N_{+1} и бифуркации, связанной с переходом неподвижной точки через границу области определения точечного преобразования [3, 4].

4. СТРУКТУРА РАЗБИЕНИЯ ФАЗОВОГО ПРОСТРАНСТВА

Используя полученные результаты исследования точечных отображений, рассмотрим разбиение фазового пространства на траектории при различных значениях параметров n и m . В силу леммы 1 предшествующей работы все фазовые полураектории, начинающиеся в любой конечной части фазового пространства, по истечении конечного промежутка времени попадают внутрь слоя $-1/k < x < 1/k$ и в нем остаются. При $m > n$ внутри слоя $-1/k < x < 1/k$ фазового пространства существует единственная симметричная относительно начала координат асимптотически устойчивая интегральная поверхность S^- , образованная траекториями, проходящими через кривую J . Следовательно, все фазовые траектории либо неограниченно приближаются к S^- при $t \rightarrow \infty$, либо попадают на полоску скользящих движений и стремятся к отрезку O_-O_+ состояний равновесия (рис. 3). В случае A отрезку состояний равновесия соответствуют стационарные вращения рамы гиростабилизатора, а в случае B — состояния покоя. При $n, m \in \omega_3$ отрезок асимптотически устойчив. На поверхности S^- существует симметричный неустойчивый предельный цикл седлового типа, соответствующий простой неподвижной точке отображения T_1S . Через него проходит сепаратрисная поверхность S^+ , образованная траекториями, проходящими через J^+ и асимптотически приближающимися к предельному циклу. Поверхность S^+ ограничивает область притяжения отрезка состояний равновесия. Фазовые траектории, попадающие на плоскость XY внутрь области Σ , представляющей собой сечение поверхности S^+ плоскостью XY , асимптотически приближаются к отрезку состояний равновесия. Все траектории, попадающие на плоскость XY вне Σ , при $t \rightarrow \infty$ уходят в бесконечность. Остановимся подробнее на определении границы области Σ притяжения отрезка состояний равновесия. В силу центральной симметрии фазового пространства рассмотрим часть границы области Σ , расположенную в G_+ и G при $y < 0$ и состоящую из кривых J^+, L, C (рис. 3). Для определения кривой J^+ , которая является сепаратрисой неподвижной точки преобразования T_1S , построен алгоритм, использующий вспомогательные отображения, введенные Неймарком в [5]. Описание алгоритма приведено в Приложении 2. По этому алго-

ритму рассчитаны сепаратрисы для ряда значений параметров. В результате расчетов на ЭВМ установлено, что сепаратриса J^+ имеет ко-

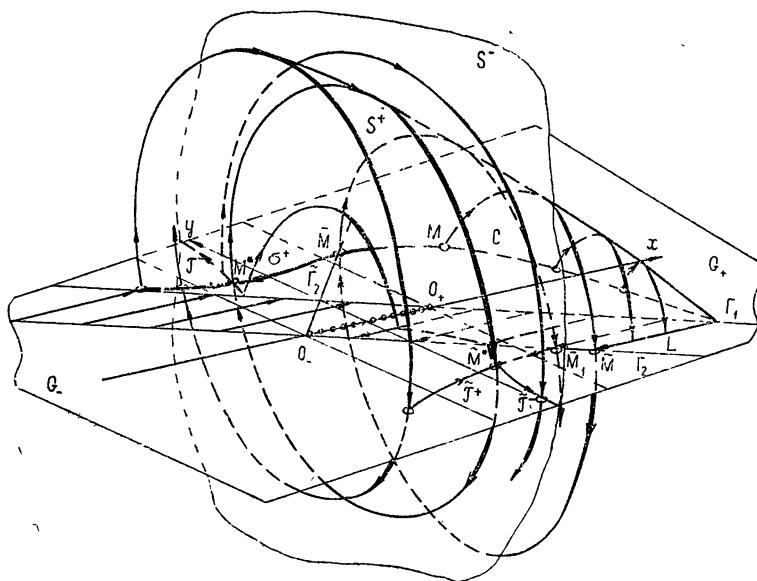


Рис. 3.

нечную длину, причем концевые точки \tilde{M}' , \tilde{M} ее расположены на лучах Γ_1 и $\bar{\Gamma}_1 = (T_1 S) \Gamma_1$ (рис. 3).

Покажем, каким образом можно определить кривую C . В работе [1] было введено отображение T_4 области G_- на плоскость XY , порожданное траекториями из R_- . Прямая Γ_2 при $y < 0$ принадлежит G_- и отображением T_4 переводится в прямую Γ_2 . Точка \tilde{M} , симметричная точке \tilde{M}' , отображением T_4 переводится в точку \tilde{M} , и пусть $\tilde{M}_1 = T_1(\tilde{M})$. Очевидно, что траектория L в полоске скользящих движений, проходящая через \tilde{M} , принадлежит границе области Σ . В углу, составленном полу-прямыми Γ_2 при $y < 0$ и $\bar{\Gamma}_2$, фазовые траектории, идущие из бесконечности, впервые попадают на плоскость XY . Отсюда следует, что C — такая кривая, которая отображением T_1 переводится в часть границы Σ , состоящую из L и части \tilde{J}^+ от точки \tilde{M} до \tilde{M}_1 . Следовательно, для отыскания C надо применить отображение T_1^{-1} к этой части Σ . Фазовые траектории, проходящие через Σ , ограничивают в фазовом пространстве область притяжения отрезка состояний равновесия.

Перейдем к рассмотрению разбиения фазового пространства на траектории при $(n, m) \in \omega_1$. В случае A отрезок состояний равновесия асимптотически устойчив. В фазовом пространстве существует неустойчивый цикл, соответствующий неподвижной точке отображения $T_1 S T_2$ (рис. 3). Этот предельный цикл симметричен и состоит из четырех кусков траекторий, два из которых проходят в R_+ и R_- , а два других принадлежат полоске скользящих движений. Область Σ притяжения отрезка состояний равновесия в плоскости XY в этом случае ограничена траекторией скользящего движения L^* , проходящей через неподвижную точку, и кривой $C = T_1^{-1} L^*$. Поверхность, образованная траекториями, проходящими через границу Σ , является областью притяжения отрезка состояний равновесия в фазовом пространстве. Траектории, расположенные вне

области притяжения отрезка $O_- O_+$, асимптотически приближаясь к S^- , стремятся в бесконечность при $t \rightarrow \infty$.

В случае B отрезок состояний равновесия неустойчив, и все фазовые траектории, асимптотически приближаясь к S^- , уходят в бесконечность.

При бифуркационном значении параметров в случае A предельный цикл не содержит отрезков скользящего режима и проходит через границу полоски скользящих движений. В случае B при бифуркационном значении параметров отрезок состояний равновесия нейтрален, а через каждую точку границы полоски скользящих движений от O_+ до P проходит замкнутая траектория, состоящая из участков траекторий в R_+ и R_- и двух отрезков траекторий скользящего режима.

Рассмотрим характер приближения фазовых траекторий к отрезку состояний равновесия. В случае A изображающая точка либо попадает в область \tilde{G} притяжения состояний равновесия и затем по траекториям скользящих движений приближается к состоянию равновесия (рис. 4 а), либо приближается к O_+ (O_-), двигаясь по траекториям в R_+ (R_-) и G (рис. 4 б). Точки O_- и O_+ соответствуют вращениям рамы гиростабилизатора с максимальной скоростью. Отрезок состояний равновесия представляет собой асимптотически устойчивое многообразие динамической системы.

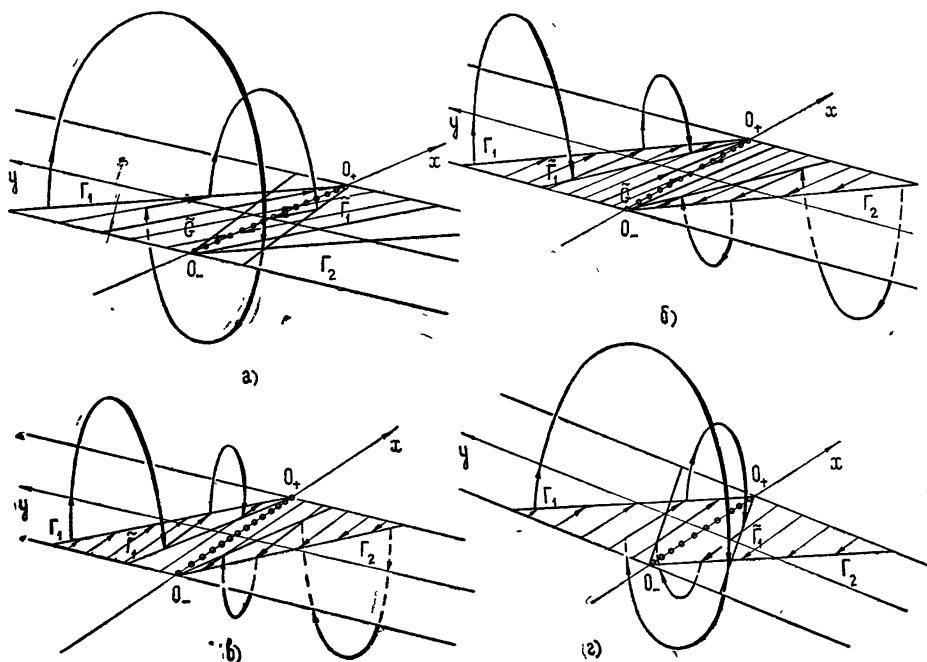


Рис. 4.

В случае B могут быть две возможности.

1) При $(n, m) \in \omega_3$ фазовые траектории асимптотически приближаются к концевым точкам отрезка $O_- O_+$ (рис. 4в). В этом случае отдельные точки отрезка $O_- O_+$ неустойчивы по отношению к начальным возмущениям, но в то же время совокупность точек O_- и O_+ составляет асимптотически устойчивое многообразие [6]. Точки O_- и O_+ соответствуют положению покоя гиростабилизатора, при котором кинетическая

ось гироскопа отклонена от перпендикуляра к плоскости рамы на величину, при которой момент, развивающий разгрузочным двигателем, равен моменту сил трения.

2) При $n, m \in \omega_3$ фазовые траектории асимптотически приближаются ко всему отрезку O_-O_+ , а не к отдельным его точкам (рис. 4 г). Отдельные точки неустойчивы по отношению к начальным возмущениям. Рама гиростабилизатора в этом случае совершает колебания около состояния покоя с уменьшающимися размахами, а гироскоп колеблется около перпендикуляра к плоскости рамы, причем размах колебаний гироскопа стремится к удвоенному значению отклонения, при котором момент, развивающий разгрузочным двигателем, равен максимальному значению момента трения.

Заметим, что случаи A и B при $k = 2\rho$ совпадают.

ПРИЛОЖЕНИЕ 1

Выражая β через τ из уравнений $F(\beta, \tau) = 0$ и используя выражения для $K(\beta, \tau)$, $\frac{dK}{d\beta}$ и $\frac{d\tau}{d\beta}$ [1], получим для случая A

$$p(\tau) = e^{-k\tau} \Phi^{-1}(\tau) \left\{ -\frac{k(k^2 + \sigma^2 - \rho^2)}{2\rho\sigma} \sin^2(\sigma\tau) + \right. \\ \left. + \frac{k}{2\sigma} \sin(2\sigma\tau) + 1 + e^{(k+\rho)\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] \right\}, \quad (17)$$

$$q(\tau) = e^{-k\tau} \Phi_{(\tau)}^{-1} \left\{ -\frac{k(k^2 + \sigma^2 - \rho^2)}{2\rho\sigma} \sin^2(\sigma\tau) - \frac{k}{2\sigma} \sin(2\sigma\tau) + \right. \\ \left. + 1 - e^{-(k+\rho)\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right] \right\}; \quad (18)$$

Для случая B

$$p(\tau) = e^{\rho\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right], \quad (19)$$

$$q(\tau) = e^{-k\tau} \Phi_{(\tau)}^{-1} \left[e^{-(k+\rho)\tau} + \frac{k+\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] \times \\ \times \left[\frac{\rho}{\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right], \quad (20)$$

где

$$\Phi(\tau) = -e^{-(k+\rho)\tau} \left[\frac{k+\rho}{\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right] + 1.$$

Выражение $p(\tau) - q(\tau)$ получается из (17) — (20) и имеет для обоих случаев следующий вид:

$$p - q = e^{-k\tau} \Phi_{(\tau)}^{-1} \left\{ \frac{k}{\sigma} \sin(2\sigma\tau) + e^{(k+\rho)\tau} \times \right. \\ \left. \times \left[\frac{\rho}{\sigma} \sin(\sigma\tau) - \cos(\sigma\tau) \right] + e^{-(k+\rho)\tau} \left[\frac{\rho}{\sigma} \sin(\sigma\tau) + \cos(\sigma\tau) \right] \right\}.$$

При $\pi/\sigma \leq \tau \leq \tau_r$ $\Phi(\tau) > 0$, следовательно, знак $p - q$ определяется знаком числителя. Обозначим функцию, стоящую в фигурных скобках,

через $f(\tau)$, и покажем, что $f(\tau)$ при $\pi/\sigma < \tau < \tau_r$ может обращаться в нуль лишь в одной точке. Преобразуем $f(\tau)$. Так как $\sigma\tau = \pi$ и $\sigma\tau = 3\pi/2$ нулями $f(\tau)$ быть не могут, то, разделив $f(\tau)$ на $\sin(2\sigma\tau)$, получим

$$\frac{\rho}{\sigma} \frac{\operatorname{ch}(k + \rho)\tau}{\cos(\sigma\tau)} - \frac{\operatorname{sh}(k + \rho)\tau}{\sin(\sigma\tau)} + \frac{k}{\sigma} = 0. \quad (21)$$

Легко видеть, что при $3\pi/2\sigma < \tau < 2\pi$ $f(\tau) > 0$. Следовательно, остается исследовать это уравнение при $\pi/\sigma \leq \tau < 3\pi/2$. Графическое решение этого уравнения показано на рис. 5. Следовательно, функция $f(\tau)$ может обращаться в нуль лишь в одной точке.

Используя этот факт, легко проанализировать знак $p - q$ для точек области D_1 при любых значениях параметров $m > n$. Для этого запишем

$$p_A\left(\frac{\pi}{\sigma}\right) = \exp\left(\frac{\rho}{\sigma}\pi\right), \quad q_A\left(\frac{\pi}{\sigma}\right) = \exp\left(-\frac{k}{\sigma}\pi\right),$$

$$p_B\left(\frac{\pi}{\sigma}\right) = \exp\left(\frac{\rho}{\sigma}\pi\right), \quad q_B\left(\frac{\pi}{\sigma}\right) = \exp\left(-\frac{k}{\sigma}\pi\right), \quad (22)$$

$$p_A(\tau_r) = \frac{k}{2\rho} K_r, \quad q_A(\tau_r) = -\frac{k - 2\rho}{2\rho} \exp(-k\tau_r),$$

$$p_B(\tau_r) = K_r, \quad q_B(\tau_r) = 0.$$

Так как при $n, m \in \omega_1$ и $\omega_2 K_r \geq 0$, то из (22) следует, что $p - q > 0$ при всех $\pi/\sigma < \tau < 2\pi/\sigma$. При $n, m \in \omega_3 K_r < 0$. В этом случае несложно показать, что $p - q > 0$ для $\tau = \underline{\tau}$, где τ соответствует $\underline{\beta}$, для которого $K(\underline{\beta}, \underline{\tau}) = 0$. Так как для всех точек D_1 $\beta < \underline{\beta}$, то для всех $x, y \in D_1$ при любых $n, m \in \omega_3$ выполняется неравенство $p - q \geq 0$.

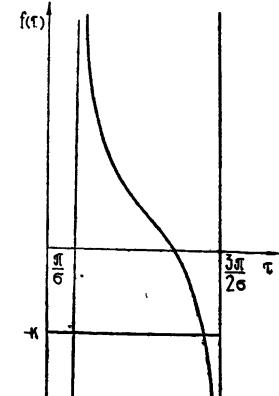


Рис. 5.

ПРИЛОЖЕНИЕ 2

Пусть $T(x, y) = (\bar{x}, \bar{y})$ — отображение, имеющее неподвижную точку (x^*, y^*) седлового типа. Очевидно, что в окрестности (x^*, y^*) T — несжимающее отображение. Пусть по координате x — сжатие, а по координате y — растяжение в некоторой метрике. Рассмотрим произвольную точку в окрестности (x^*, y^*) , лежащую, например, на α^+ сепаратрисе. Очевидно, что $T^k(x_0, y_0) = (x_k, y_k) \rightarrow (x^*, y^*)$ при $k \rightarrow \infty$. Запишем значения итераций точки (x_0, y_0) при последовательном применении отображения T в виде таблицы

$$\begin{array}{ccccccc} x_0 & x_1 & x_2 & \dots & x_k \\ y_0 & y_1 & y_2 & \dots & y_k \end{array} \quad (23)$$

Поставим задачу: определить значения x_k и y_0 по известным значениям x_0 и y_k . Очевидно, что эта задача будет решена, если удастся построить отображение \tilde{T}^k , вспомогательное к T^k [5]. Так как $T^k(x_0, y_0) = (x_k, y_k)$, то $\tilde{T}^k(x_0, y_k) = (x_k, y_0)$. Будем искать отображение \tilde{T}^k методом итераций. В таблице (23) зададим нулевые приближения $x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)}$ и, применяя последовательно отображение \tilde{T} , идя справа налево, определим $y_{k-1}^{(1)}, y_{k-2}^{(1)}, \dots, y_1^{(1)}, y_0^{(1)}$. Затем, повторяя вычисления, но уже слева

направо, определим $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \dots, x_k^{(1)}$ и т. д. Будем считать за один пересчет таблицы (23) пересчет справа налево и слева направо. После N пересчетов получим таблицу

$$\begin{array}{c} x_0 \quad x_1^{(N)} \quad x_2^{(N)} \quad \dots \quad x_k^{(N)} \\ y_0^N \quad y_1^{(N)} \quad y_2^{(N)} \quad \dots \quad y_k \end{array}$$

В силу сжимаемости отображения \widetilde{T} [5] итерационный процесс устойчив и при $N \rightarrow \infty$ сходится к некоторой инвариантной таблице, которой, очевидно, будет таблица (23), т. е. $x_k^{(N)}$ и $y_0^{(N)}$ будут некоторыми приближениями к значениям x_k и y_0 . Если в качестве y_k взять y^* и увеличивать число членов таблицы, то при $k \rightarrow \infty$, $N \rightarrow \infty$ $y_0^{(N)} \rightarrow y_0$, $x_k^{(N)} \rightarrow x_k$. Отсюда легко получить алгоритм определения точки на сепаратрисе с любой заданной степенью точности. Пусть требуется найти точку на сепаратрисе a^+ с точностью ε . Выбираем такое x_0 , чтобы прямая $x = x_0$ пересекала сепаратрису в некоторой точке (x_0, y_0) . Требуется найти y_0 . Составляем таблицу с числом членов $k = k_0 + l\Delta_k$, где k_0 — число членов исходной таблицы, $l = 1, 2, \dots, \Delta_k$ — приращение числа членов таблицы. За y_k берем y^* и выбираем нулевое приближение, например, $x_j = x^* + jh$, где $h = (x_0 - x^*)/k$ ($j = 1, 2, \dots, k$). Производим пересчеты таблицы до тех пор, пока $|y_0^{(N)} - y_0^{(N-1)}| < \varepsilon_1$, где $\varepsilon_1 < \varepsilon$. После этого, полагая $l = 1, 2, \dots$, повторяем вычисления. Процесс заканчивается, когда $|y_0^l - y_0^{l-1}| < \varepsilon$. В качестве искомого значения y_0 берем y_0^{l+1} . При вычислении по приведенному алгоритму в настоящей работе было выбрано $\varepsilon = 10^{-5}$, $k_0 = 5$, $\Delta_k = 5$, и уже при $l = 2$ была получена заданная точность. Приведенный алгоритм с успехом может быть использован для нахождения неподвижной точки седлового типа любой кратности.

В заключение автор выражает глубокую благодарность Ю. И. Неймарку за руководство и помощь при выполнении работы.

ЛИТЕРАТУРА

1. Н. Я. Коган, Изв. высш. уч. зав — Радиофизика, 14, № 3, (1971).
2. Ю. И. Неймарк, Изв. высш. уч. зав.—Радиофизика, 10, № 3, 311 (1967).
3. В. А. Брусины, Ю. И. Неймарк, М. И. Фейгин, Изв. высш. уч. зав.—Радиофизика, 6, № 4, 785 (1963).
4. Ю. И. Неймарк, Изв. высш. уч. зав.—Радиофизика, 1, № 2, 95 (1958).
5. Ю. И. Неймарк, Тр. 2 Всесоюзного съезда по теоретической и прикладной механике, вып. 2, изд. Наука, М., 1965.
6. Ю. И. Неймарк, Н. А. Фуфайев, Динамика неголономных систем, изд. Наука, М., 1967.

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
13 июля 1970 г.

TO THE THEORY OF A SINGLE-AXIS GYROSTABILIZER II

N. Ya. Kogan

Dynamics of a single-axis gyrostabilizer with dry friction in the precession axis or in the stabilization axis is investigated by the point-mapping method.

УДК 62—506

ОБ ЭКСТРЕМАЛЬНЫХ СИСТЕМАХ С ПРЕИМУЩЕСТВЕННЫМ ПОИСКОМ

Н. Н. Леонов

Исследуется характер движений в экстремальных системах преимущественного поиска для устойчивых объектов с инерционным входом.

1. В работах [1, 2] рассматривались два способа экстремального управления инерционными устойчивыми объектами, осуществляющие преимущественный поиск, на примере объекта с инерционностью первого порядка на выходе. Исследования экстремальных систем автоколебательного типа с аналогичными объектами, не подчиненных условию преимущественности направления поиска, обнаружили существование специфических различий в поведении объектов, обладающих инерционностью на выходе, и объектов, обладающих инерционностью на входе [3—5]. Поэтому, при выяснении возможностей преимущественного поиска нельзя ограничиться исследованием преимущественного экстремального управления лишь на объектах с инерционностью на выходе. Настоящая работа содержит исследование преимущественного поиска на устойчивом объекте, имеющем инерционность первого порядка на входе (рис. 1).

2. Уравнения движения указанного объекта и исполнительного элемента могут быть записаны так:

$$\varphi = -u^2, \dot{u} + u = z, \dot{z}(t) = \eta(t - \theta). \quad (1)$$

Рассмотрим управление, осуществляющее по схеме, предложенной в [2]

$$\eta = \begin{cases} -z & \text{при } \dot{\varphi} < -\Delta \\ +1 & \text{при } \dot{\varphi} > -\Delta, \dot{z} > 0 \\ -1 & \text{при } t \in (t_k, t_k + \tau), \text{ если } \dot{\varphi} > -\Delta, \dot{z} < 0 \end{cases} \quad (2)$$

где t_k — момент начала движения, либо момент изменения знака \dot{z} с $+1$ на -1 . При этом управлении система (1), (2) описывается двумя приведенными параметрами τ , Δ .

Фазовое пространство системы (1), (2) — двумерное, в координатах u , v ($v = \dot{u}$), двулипистое, образованное двумя плоскостями Φ_{+1} , Φ_{-1} , на которых $\eta = +1$, $\eta = -1$ соответственно. В переменных u , v уравнения (1), (2) записываются так:

$$\dot{u} = v, \dot{v} = \eta - v,$$

$$\eta(t_k + \theta) = \begin{cases} -\eta(t) & \text{при } 2uv > \Delta \\ +1 & \text{при } 2uv < \Delta, \eta(t) > 0 \\ -1 & \text{при } t \in (t_k, t_k + \tau), \text{ если } 2uv < \Delta, \eta(t) < 0 \end{cases}. \quad (3)$$

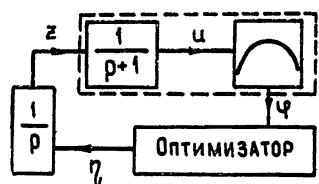


Рис. 1.

Величина θ в дальнейшем полагается равной нулю, она необходима для непротиворечивости записи системы (3) при описании движения рассматриваемой системы в области $\varphi < -\Delta$, т. е. в области $2uv > \Delta$.

3. Рассмотрим характерные черты поведения фазовой точки при движении ее в фазовом пространстве системы (3).

Уравнения движения фазовой точки вдоль траекторий листов Φ_{+1} , Φ_{-1} могут быть записаны следующим образом:

$$u = u_0 + (1 - e^{-t})v_0 - \eta(1 - t - e^{-t}), \quad v = e^{-t}v_0 + \eta(1 - e^{-t}), \quad (4)$$

где $u_0 = u(0)$, $v_0 = v(0)$.

Траектории листа Φ_{+1} изображены на рис. 2а. Траектории листа Φ_{-1} симметричны им относительно точки $(0,0)$. Движение фазовой точки, согласно (3), происходит по этим траекториям в области $2uv < \Delta$, где $\varphi > -\Delta$. В области $2uv > \Delta$ имеет место неравенство $\varphi < -\Delta$ и в этой области траектории системы соответствуют режимам скольжения системы и представляют собой параллельные прямые [4]

$$u + v = \text{const}. \quad (5)$$

Фазовые траектории системы (3) изображены на рис. 2б.

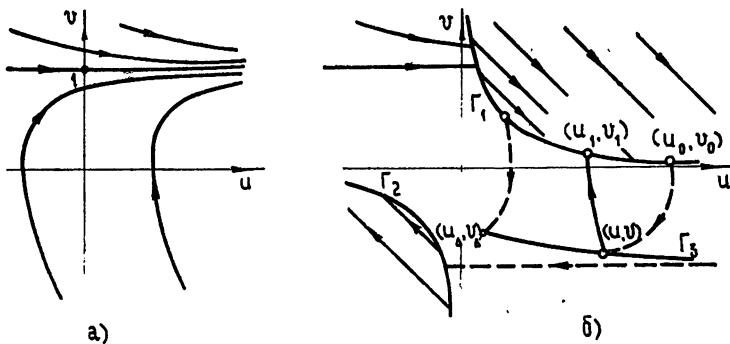


Рис. 2.

Обозначим линии, разделяющие области гладких движений от движений, соответствующих режимам скольжения, через Γ_1 для $u > 0$ и через Γ_2 для $u < 0$. Уравнение этих линий $2uv = \Delta$.

Кривая Γ_1 имеет по одной точке контакта с траекториями режимов скольжения и траекториями листа Φ_{-1} . Первая из них имеет абсциссу $u = \sqrt{0,5\Delta}$, а абсцисса точки контакта Γ_1 с траекторией листа Φ_{-1} определяется равенством

$$4u^3 + 2\Delta u^2 - \Delta^2 = 0. \quad (6)$$

Нетрудно показать, что уравнение (6) имеет единственный положительный корень $u = u_c$, и что $u_c < \sqrt{0,5\Delta}$.

Если каждую точку кривой Γ_1 с $u > \sqrt{0,5\Delta}$ переместить за время τ вдоль соответствующей траектории листа Φ_{-1} , то в результате получим некоторую непрерывную кривую, которую обозначим через Γ_3 .

Теперь можно проследить последовательность движений фазовой точки в фазовом пространстве.

Из любого начального состояния в области $2uv > \Delta$ фазовая точка по траекториям режимов скольжения приходит либо на Γ_1 , либо на Γ_2 .

Из области $2uv < \Delta$ фазовая точка по листу Φ_{+1} приходит либо на Γ_1 , либо на участок Γ_2 с $u \in (-u_c, 0)$. Начав движение по листу Φ_{-1} в области $2uv < \Delta$, фазовая точка либо попадает на Γ_1 с $u \in (0, u_c)$,

либо на Γ_2 , либо через время τ с начала движения по Φ_{-1} переходит на Φ_{+1} , если за это время не попадает ни на Γ_1 , ни на Γ_2 . Обозначим участок линии Γ_1 с $u > \sqrt{0,5\Delta}$ через Γ_1^1 , а с $u < -\sqrt{0,5\Delta}$ — через Γ_1^2 . Соответственно участок Γ_2 с $u < -\sqrt{0,5\Delta}$ обозначим через Γ_2^1 , а с $u > \sqrt{0,5\Delta}$ — через Γ_2^2 .

Если фазовая точка попадает из области $2uv < \Delta$ либо на Γ_1^2 , либо на Γ_2^2 , то далее ее движение происходит по траектории режима скольжения [6] и она приходит соответственно либо на Γ_1^1 , либо на Γ_2^1 .

Из начального состояния, принадлежащего Γ_1^1 , фазовая точка может совершать движение лишь по листу Φ_{+1} , и в результате она приходит на Γ_1^1 . Из начального же состояния, принадлежащего Γ_1^1 , фазовая точка может совершать движение лишь по листу Φ_{-1} , в результате чего она либо приходит на Γ_2 , либо на Γ_3 , а затем в обоих случаях вновь возвращается на Γ_1^1 , как было описано выше.

Таким образом, исследование движений фазовой точки может быть сведено к изучению отображения линии Γ_1^1 в себя, т. е. к изучению соответствующего отображения прямой в себя, которое обозначим через T .

Согласно изложенному выше, представляются возможными следующие элементарные последовательности движения фазовой точки:

$$\begin{aligned} \Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^1, \quad \Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1, \quad \Gamma_1^1 \xrightarrow{-1} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1, \\ \Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1, \quad \Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1, \quad (7) \\ \Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1, \quad \Gamma_1^1 \xrightarrow{-1} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1, \quad \Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1. \end{aligned}$$

Здесь стрелкой указан переход с одной кривой Γ_i на другую Γ_j по траектории пространства, индекс которого стоит над стрелкой (нуль-индекс пространства траекторий режимов скольжения).

Каждое движение фазовой точки, начинающееся с Γ_1^1 , может быть описано последовательностью переходов этой точки по кривым Γ_i , которая образована элементарными последовательностями (7).

4. Установим, какие элементарные последовательности из перечисленных в (7) и в каких сочетаниях имеют место в системе (3). Рассмотрим для этого плоскость параметров τ , Δ и выделим в ней области существования этих последовательностей.

Прежде всего необходимо отметить существование последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^1$ для любых $\tau > 0$, $\Delta > 0$. Иными словами, для любых τ , Δ существуют такие точки линии Γ_1^1 , для которых определена эта последовательность.

Выясним теперь, какие еще из перечисленных в (7) последовательностей могут существовать в системе (3).

Обозначим через Γ_4 траекторию листа Φ_{-1} , проходящую через точку $(\sqrt{0,5\Delta}, \sqrt{0,5\Delta})$ линии Γ_1 . Ее уравнения

$$u = (2 - e^{-t})\sqrt{0,5\Delta} + 1 - t - e^{-t}, \quad v = e^{-t}\sqrt{0,5\Delta} - 1 + e^{-t}. \quad (8)$$

Из (8) следует, что Γ_4 проходит через точку $(-\sqrt{0,5\Delta}, -\sqrt{0,5\Delta})$ линии Γ_2 при $\Delta = \Delta^*$, где $\Delta^* = 2th^2\sqrt{2\Delta}$. Если $\Delta < \Delta^*$, то Γ_4 пересекает Γ_2^2 , а если $\Delta > \Delta^*$, то Γ_2^1 . Это говорит о том, что в соответствующей части плоскости $\tau\Delta$, кроме $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \dots$, существуют при $\Delta < \Delta^*$ только последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{+1} \dots$, а при $\Delta > \Delta^*$ существуют последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^1 \xrightarrow{+1} \dots$ и может быть $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{+1} \dots$. Для упомянутой вы-

ше части плоскости $\tau\Delta$ кривые Γ_2 , Γ_3 должны иметь по крайней мере одну общую точку, ибо в противном случае возможны лишь последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \rightarrow \dots$. Уравнение Γ_3

$$u = 0,5 \Delta(e^{-\tau}v + e^{-\tau} - 1)^{-1} + (e^{-\tau} - 1)v + e^{-\tau} - \tau - 1. \quad (9)$$

Обозначим граничную точку линии Γ_3 , соответствующую точке $(\sqrt{0,5\Delta}, \sqrt{0,5\Delta})$ линии Γ_1 , через (u_Δ, v_Δ) . Ее координаты, согласно (8)

$$u_\Delta = (2 - e^{-\tau})\sqrt{0,5\Delta} + 1 - \tau - e^{-\tau}, \quad v_\Delta = e^{-\tau}\sqrt{0,5\Delta} - 1 + e^{-\tau}. \quad (10)$$

Из (9) можно получить, что $-1 < \frac{du}{dv} < 0$, $\left.\frac{du}{dv}\right|_{(u_\Delta, v_\Delta)} = -1$. А так как коэффициент наклона кривой Γ_2^1 по модулю меньше единицы, а кривой Γ_2^2 — больше единицы, то пересечение Γ_2 , Γ_3 наступает при $\Delta < \Delta^*$, когда точка (u_Δ, v_Δ) попадает на Γ_2^2 , а при $\Delta > \Delta^*$, когда появляется точка контакта кривых Γ_2^1 , Γ_3 .

Точка (u_Δ, v_Δ) принадлежит Γ_2 , если выполняется условие

$$\left(\frac{\Delta}{2}\right)^{1/2} = \left\{ \left[\frac{\tau e^{-\tau}}{2(1 - e^{-\tau})^2} \right]^2 + \frac{\tau}{(1 - e^{-\tau})^2} \right\}^{1/2} - 1 - \frac{\tau e^{-\tau}}{2(1 - e^{-\tau})^2}. \quad (11)$$

В плоскости $\tau\Delta$ уравнение (11) описывает некоторую кривую, которую обозначим γ_1 .

Точки пересечения Γ_2 , Γ_3 определяются равенством $f(u) = 0$.

$$4e^{-\tau}(1 - e^{-\tau})u^3 + 2(1 - e^{-\tau})[\Delta e^{-\tau} - 2(1 - e^{-\tau} - \tau e^{-\tau})]u^2 - 2\Delta(2 - 3e^{-\tau} + e^{-2\tau} - \tau e^{-\tau})u - \Delta^2(1 - e^{-\tau}) = 0, \quad (12)$$

а точки контакта Γ_2 , Γ_3 — равенствами $f(u) = 0$ и $f'(u) = 0$

$$6e^{-\tau}(1 - e^{-\tau})u^2 + 2(1 - e^{-\tau})[\Delta e^{-\tau} - 2(1 - e^{-\tau} - \tau e^{-\tau})]u - \Delta(2 - 3e^{-\tau} + e^{-2\tau} - \tau e^{-\tau})\Delta = 0. \quad (13)$$

Нас будет интересовать левая точка контакта, для которой

$$6e^{-\tau}u = 2(1 - e^{-\tau} - \tau e^{-\tau}) - \Delta e^{-\tau} - \left\{ [\Delta e^{-\tau} - 2(1 - e^{-\tau} - \tau e^{-\tau})]^2 + 6 \left(2 - e^{-\tau} - \frac{\tau \exp(-\tau)}{1 - \exp(-\tau)} \right) \Delta e^{-\tau} \right\}^{1/2}. \quad (14)$$

Точка контакта кривых Γ_2^1 , Γ_3 существует при тех τ , Δ , которые удовлетворяют уравнениям (12), (14). Соответствующую кривую в плоскости $\tau\Delta$ обозначим γ_2 .

Таким образом, линии γ_1 при $\Delta < \Delta^*$ γ_2 при $\Delta > \Delta^*$ делят первый квадрант плоскости $\tau\Delta$ на две части: для одной из них, расположенной слева от этой линии, линия Γ_3 не пересекается с Γ_2 и расположена целиком в области $2uv < \Delta$ фазового пространства (для этой области последовательности переходов имеют вид $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \rightarrow \dots$). Для другой части плоскости $\tau\Delta$ линии Γ_3 , Γ_2 имеют общие точки, часть линии Γ_3 расположена в области $2uv > \Delta$ фазового пространства, и последовательности переходов имеют вид $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \rightarrow \dots$, либо $\Gamma_1^1 \xrightarrow{-1} \Gamma_2 \rightarrow \dots$.

Пусть Γ_3 не имеет общих точек с Γ_2 . Рассмотрим при этом поведение траектории листа Φ_{+1} , проходящей через точку (u_Δ, v_Δ) линии Γ_3 . Обозначим эту траекторию через Γ_5 . Ее уравнение

$$u = u_\Delta + (1 - e^{-\tau})v_\Delta - 1 + t + e^{-\tau}, \quad v = e^{-\tau}v_\Delta + 1 - e^{-\tau}. \quad (15)$$

Траектория Γ_5 проходит через точку $(\sqrt{0,5\Delta}, \sqrt{0,5\Delta})$ линии Γ_1 , если

$$\sqrt{0,5\Delta} = \operatorname{th} \frac{\tau}{2}. \quad (16)$$

Линию, описываемую равенством (16), обозначим γ_3 . Траектория Γ_5 касается линии Γ_2^2 в точке $(-u_c, -v_c)$, если

$$e^{-\tau} = \{1 \pm [1 - (1 + \sqrt{0,5\Delta})(1 + v_c) \exp(-u_e - v_c - \sqrt{2\Delta})]^{1/2}\} (1 + \sqrt{0,5\Delta})^{-1} \quad (17)$$

и $\tau > u_c + v_c + \sqrt{2\Delta}$ (эту линию обозначим γ_4).

Таким образом, та часть первого квадранта плоскости $\tau\Delta$, которая расположена слева от линий γ_1, γ_2 , разбивается линиями γ_3, γ_4 на области G_1, G_2, G_3 (рис. 3). Для $(\tau, \Delta) \in G_1$ существует лишь последовательность переходов $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^1$. Для $(\tau, \Delta) \in G_2$ существует еще последовательность $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1$, а для $(\tau, \Delta) \in G_3$, кроме указанных двух, третья $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1$.

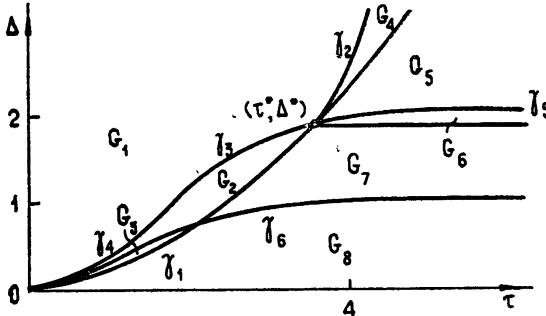


Рис. 3.

Для разбиения части квадранта, расположенной справа от γ_1, γ_2 , на подобные же области, найдем условия пересечения Γ_3 с Γ_2 в точках $(-\sqrt{0,5\Delta}, -\sqrt{0,5\Delta}), (-u_e, -v_c)$.

Первое условие выполняется на линии γ_5

$$\left(\frac{\Delta}{2}\right)^{1/2} = 1 - \frac{\tau e^{-\tau}}{2(1 - e^{-\tau})^2} + \left\{ \left[\frac{\tau e^{-\tau}}{2(1 - e^{-\tau})^2} \right]^2 - \frac{\tau e^{-2\tau}}{(1 - e^{-\tau})^2} \right\}^{1/2}, \quad (18)$$

$$\left(\frac{\Delta}{2}\right)^{1/2} > 1 - \frac{\tau}{2(e^\tau - 1)},$$

а второе — на линии γ_6 :

$$4u_c(1 - e^{-\tau})(\Delta e^{-\tau} - 1 + e^{-\tau} + \tau e^{-\tau}) - \Delta(2 - 3e^{-\tau} + e^{-2\tau} - \tau e^{-\tau}) + \quad (19)$$

$$+\sqrt{\tau^2 e^{-2\tau} + (1 - e^{-\tau})(2\tau e^{-2\tau}(1 - 2e^{-\tau}) + (1 - e^{-\tau})[4\Delta^3(1 + e^{-\tau}) - 4 + 5e^{-\tau}])} = 0,$$

$$6u_c^2(e^\tau - 1) - 2u_c(e^\tau - 1)[\Delta - 2(e^\tau - 1 - \tau)] - \Delta(2e^{2\tau} - 3e^\tau + 1 - \tau e^\tau) < 0.$$

Линии $\gamma_1, \gamma_2, \gamma_5, \gamma_6, \Delta = \Delta^*$ разбивают указанную часть первого квадранта плоскости $\tau\Delta$ на области G_4, G_5, G_6, G_7, G_8 . Для $(\tau, \Delta) \in G_1$ и $(\tau, \Delta) \in G_5$ существуют последовательности переходов $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^1$ и $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1$, для $(\tau, \Delta) \in G_6$ к ним добавляются $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1$.

$\Gamma_1^{1-1} \rightarrow \Gamma_2^{20} \rightarrow \Gamma_2^{1+1} \rightarrow \Gamma_1^{20} \rightarrow \Gamma_1^1$. Если $(\tau, \Delta) \in G_7$, то существуют последовательности $\Gamma_1^{1-1} \rightarrow \Gamma_3^{+1} \rightarrow \Gamma_1^1$, $\Gamma_1^{1-1} \rightarrow \Gamma_2^{20} \rightarrow \Gamma_2^{1+1} \rightarrow \Gamma_1^{20} \rightarrow \Gamma_1^1$ и может существовать последовательность $\Gamma_1^{1-1} \rightarrow \Gamma_3^{+1} \rightarrow \Gamma_1^{20} \rightarrow \Gamma_1^1$. Если $(\tau, \Delta) \in G_8$, то к последним последовательностям добавляется $\Gamma_1^{1-1} \rightarrow \Gamma_3^{+1} \rightarrow \Gamma_2^{20} \rightarrow \Gamma_2^{1+1} \rightarrow \Gamma_1^{20} \rightarrow \Gamma_1^1$.

Итак, из последовательностей переходов, перечисленных в (7), последние две не имеют места в системе (3).

5. Отображение T линии Γ_1^1 в себя определено при $u_0 \in (\sqrt{0,5\Delta}, \infty)$ и образовано элементарными отображениями T_i ($i = 1, 2, \dots, 6$), соответствующими первым шести из указанных в (7) элементарным последовательностям перехода фазовой точки. Вначале исследуем эти элементарные отображения, а затем перейдем к изучению отображения T в целом.

а) Отображение, соответствующее элементарной последовательности $\Gamma_1^{1-1} \rightarrow \Gamma_3^{+1} \rightarrow \Gamma_1^1$, обозначим через T_1 . Пусть (u_0, v_0) — начальная точка Γ_1^1 , (u, v) — соответствующая ей точка Γ_3 , и (u_1, v_1) — конечная точка Γ_1^1 . Тогда, согласно (4),

$$u = u_0 + (1 - e^{-\tau})v_0 + 1 - \tau - e^{-\tau}, \quad v = v_0 e^{-\tau} - 1 + e^{-\tau}, \quad 2u_0 v_0 = \Delta, \quad (20)$$

$$u_1 = u_0 + (1 - e^{-t})v - 1 + t + e^{-t}, \quad v_1 = v e^{-t} + 1 - e^{-t}, \quad 2u_1 v_1 = \Delta.$$

Согласно (20), равенство, описывающее отображение T_1 ($u_1 = T_1 u_0$), может быть записано в виде

$$\left[2 - \exp(-\tau) - \exp(-\tau) \frac{\Delta}{2u_0} \right] \exp \left(u_0 + \frac{\Delta}{2u_0} - \tau \right) = \left(1 - \frac{\Delta}{2u_1} \right) \times \\ \times \exp \left(u_1 + \frac{\Delta}{2u_1} \right). \quad (21)$$

Из анализа поведения фазовых траекторий системы (3) следует, что отображение T_1 имеет монотонно возрастающую непрерывную функцию последовательности. Поэтому оно может иметь лишь простые неподвижные точки. Из (8) получаем, что T_1 имеет единственную простую неподвижную точку

$$u_1^* = 0,5 \Delta \operatorname{cth} \frac{\tau}{2}. \quad (22)$$

Нетрудно обнаружить, что для больших u_0 ($u_0 \gg \sqrt{0,5\Delta}$) имеет место неравенство $T_1 u_0 < u_0$. А так как $u_1 = T_1 u_0$ — непрерывная возрастающая функция, то u_1^* устойчива.

Неподвижной точке u_1^* соответствует предельный цикл системы (3), образованный двумя отрезками траекторий листов Φ_{+1} , Φ_{-1} (рис. 4 а). Он существует в том случае, если u_1^* принадлежит области определения отображения T_1 (т. е. $u_1^* > \sqrt{0,5\Delta}$) и если точка (u, v) линии Γ_3 , принадлежащая указанному циклу, принадлежит области

$2uv < \Delta$. Первое условие выполняется при $\Delta > 2\operatorname{th}^2 \frac{\tau}{2}$, второе — при

$\Delta > \tau \operatorname{th} \frac{\tau}{2} - 2\operatorname{th}^2 \frac{\tau}{2}$. Соответствующие граничные линии $\Delta = 2\operatorname{th}^2 \frac{\tau}{2}$ (пусть

это линия l_1 , она совпадает с γ_3) и $\Delta = \tau \operatorname{th} \frac{\tau}{2} - 2\operatorname{th}^2 \frac{\tau}{2}$ (а это — l_2) пере-

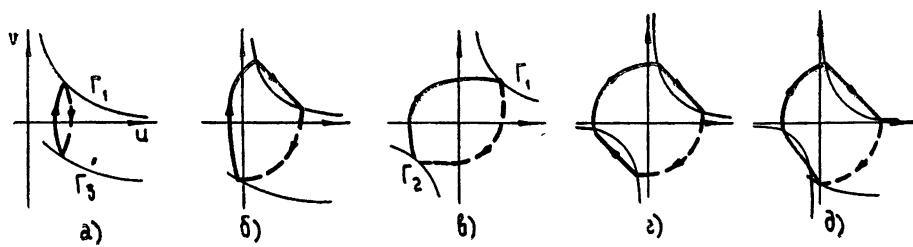


Рис. 4.

секаются в точке (τ^*, Δ^*) где $\tau^* = 4\operatorname{th} \frac{\tau^*}{2}$. Поэтому область O_1 существования неподвижной точки u_1^* в плоскости параметров $\tau\Delta$ определяется неравенствами (рис. 5)

$$\Delta > \begin{cases} 2\operatorname{th}^2 \frac{\tau}{2} & \text{при } \tau < \tau^* \\ \tau \operatorname{th} \frac{\tau}{2} - 2\operatorname{th}^2 \frac{\tau}{2} & \text{при } \tau > \tau^* \end{cases} \quad (23)$$

б) Рассмотрим теперь отображение, соответствующее элементарной последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1$. Обозначим его T_2 . Пусть (u_0, v_0) , (u, v) , (u_1, v_1) , (u_2, v_2) — точки соответствующей фазовой траектории, принадлежащие Γ_1^1 , Γ_3 , Γ_1^2 и Γ_1^1 . Координаты этих точек согласно (4), (5) связаны соотношениями (20) и $u_2 = v_1$, $v_2 = u_1$. Из этих соотношений получаем равенство, описывающее отображение $T_2(u_2 = T_2 u_0)$

$$\left(2 - e^{-\tau} - e^{-\tau} \frac{\Delta}{2u_0} \right) \exp \left(u_0 + \frac{\Delta}{2u_0} - \tau \right) = (1 - u_2) \times \times \exp \left(u_2 + \frac{\Delta}{2u_2} \right). \quad (24)$$

Отображение T_2 имеет единственную простую неподвижную точку

$$u_2^* = 0,5[(1 - e^{-\tau})^2 + \sqrt{(1 - e^{-\tau})^4 + 2\Delta e^{-2\tau}}]. \quad (25)$$

Ей соответствует предельный цикл, образованный двумя отрезками траекторий листов Φ_{+1} , Φ_{-1} и отрезком прямолинейной траектории режима скольжения (рис. 4 б). Он существует, если u_2^* принадлежит области определения отображения $T_2(u_2^* > \sqrt{0,5\Delta})$, если точка (u, v) линии Γ_3 принадлежит области $2uv < \Delta$ и если все точки отрезка траектории листа Φ_{+1} , входящего в указанный предельный цикл, принадлежат области $2uv < \Delta$. Указанные условия выделяют в плоскости $\tau\Delta$ область O_2 существования неподвижной точки u_2^* , ограниченную линиями l_1, l_3, l_4 соответственно (рис. 5). Для линии l_3 точка (u, v) является точкой пересечения Γ_3 и Γ_2 , и уравнение l_3 имеет вид

$$(1 - e^{-\tau})^2(\Delta e^{-2\tau} + 1 - 2e^{-\tau} - 2e^{-3\tau} + e^{-4\tau}) + \tau e^{-2\tau}(1 - e^{-2\tau}) = \\ = [\tau e^{-2\tau} + (1 + e^{-\tau})(1 - e^{-\tau})^3] \sqrt{(1 - e^{-\tau})^4 + 2\Delta e^{-2\tau}}. \quad (26)$$

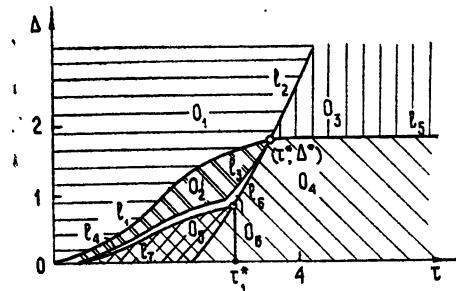


Рис. 5.

Для l_4 точка (u, v) принадлежит области $2uv < \Delta$, а отрезок траектории листа Φ_{+1} , входящий в предельный цикл, касается в точке $(-u_c, -v_c)$ линии Γ_2^2 . При этом u_2^* и u_c связаны между собой соотношением

$$u_2^* = 1 - \left(1 + \frac{\Delta}{2u_c}\right) \exp\left(-u_c - \frac{\Delta}{2u_c} - u_2^* - \frac{\Delta}{2u_2^*}\right). \quad (27)$$

Координаты точек линии l_4 могут быть вычислены следующим образом:

$$\Delta \xrightarrow{(6)} u_c \xrightarrow{(27)} u_2^{*(25)} \xrightarrow{} \tau. \quad (28)$$

Здесь цифра над стрелкой означает номер равенства, по которому вычисляется искомая величина по известной исходной.

Для определения устойчивости неподвижной точки u_2^* необходимо рассмотреть получающееся из (24) выражение

$$\frac{du_2}{du_0} = \frac{\frac{e^{-\tau}\Delta}{2u_0^2} \exp\left(u_0 + \frac{\Delta}{2u_0} - \tau\right) + (1-u_2)\left(1 - \frac{\Delta}{2u_0^2}\right) \exp\left(u_2 + \frac{\Delta}{2u_2}\right)}{\left[(1-u_2)\left(1 - \frac{\Delta}{2u_2^2}\right) - 1\right] \exp\left(u_2 + \frac{\Delta}{2u_2}\right)}. \quad (29)$$

Так как $u_2 \in (\sqrt{0,5\Delta}, 1)$, $u_0 > \sqrt{0,5\Delta}$, то $\frac{du_2}{du_0} < 0$. Следовательно, для устойчивости u_2^* необходимо выполнение условия $\left.\frac{du_2}{du_0}\right|_{u_2^*} > -1$. Условие $\left.\frac{du_2}{du_0}\right|_{u_2^*} = -1$, которое может быть записано следующим образом

$$e^{-2\tau} = 2u_2^{*2} \Delta^{-1} (2u_2^* - 1) + 1 - u_2^*. \quad (30)$$

выполняется на некоторой кривой λ плоскости $\tau\Delta$, координаты точек которой вычисляются с помощью равенств (25), (30). Численное построение этой кривой показало, что она не пересекает область O_2 . А так как для точки $(2; 1,1)$ области O_2 $\left.\frac{du_2}{du_0}\right|_{u_2^*} = -0,03$, то неподвижная

точка u_2^* устойчива в области O_2 .

в) Отображение, соответствующее элементарной последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1$, обозначим через T_3 . Изучение таких движений проводилось в [4], где оно было сведено к исследованию отображения G , соответствующего переходу $\Gamma_2^1 \xrightarrow{+1} \Gamma_1^1$ с последующим симметричным относительно начала координат переносом фазовой точки. Нетрудно обнаружить, что $T_3 = SG^2S$, где $Su = -u$. В [4] было установлено, что G имеет единственную простую абсолютную устойчивую неподвижную точку u_3^*

$$u_3^* = 0,5\Delta \operatorname{ctg}\left(u_3^* + \frac{\Delta}{2u_3^*}\right). \quad (31)$$

Этой неподвижной точке соответствует симметричный относительно начала координат устойчивый предельный цикл, образованный двумя отрезками траекторий листов Φ_{+1} , Φ_{-1} (рис. 4 в).

Область существования неподвижной точки u_3^* определяется

условиями $u_3^* > \sqrt{0.5\Delta}$ (условие принадлежности точки u_3^* области определения T_3) и $t \leq \tau$ (время движения фазовой точки по предельному циклу от Γ_1^1 до Γ_2^1 должно быть не больше τ). Эти условия выделяют в плоскости $\tau\Delta$ область O_3 (рис. 5), для которой

$$\Delta^* < \Delta < \tau \operatorname{th} \frac{\tau}{2} - 2 \operatorname{th}^2 \frac{\tau}{2} \quad (32)$$

и которая ограничена линиями l_1 и l_5 (для линии l_5 $\Delta = \Delta^*$).

г) Отображение, соответствующее элементарной последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1$, обозначим через T_4 . В [4] изучалось отображение H , соответствующее переходу $\Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1$ с последующим симметричным относительно начала координат переносом фазовой точки. Отображения T_4 , H связаны соотношением $T_4 = SH^2S$. В [4] установлено, что H имеет единственную простую абсолютно устойчивую неподвижную точку — u_4^* . Поэтому T_4 имеет также единственную простую абсолютно устойчивую неподвижную точку u_4^* :

$$1 + \frac{\Delta}{2u_4^*} = (1 - u_4^*) \exp \left(2u_4^* + \frac{\Delta}{u_4^*} \right). \quad (33)$$

Этой неподвижной точке соответствует симметричный относительно начала координат устойчивый предельный цикл, образованный двумя отрезками траекторий листов Φ_{+1} , Φ_{-1} и двумя отрезками прямолинейных траекторий режимов скольжения (рис. 4 г).

Область существования неподвижной точки u_4^* в плоскости $\tau\Delta$ определяется условием принадлежности u_4^* области определения отображения T_4 и тем условием, что время t движения фазовой точки по предельному циклу от Γ_1^1 до Γ_2^2 должно быть не больше τ . Первое условие выполняется в той части плоскости $\tau\Delta$ где $\Delta < \Delta^*$. А так как (33) может быть записано следующим образом через параметр t

$$u_4^* = \frac{2 - 2e^{-t} - te^{-t}}{2(1 - e^{-t})}, \quad \Delta = \frac{(2 - 2e^{-t} - te^{-t})(2e^{-t} - 2 + t)}{2(1 - e^{-t})^2}, \quad (34)$$

то второе условие выполняется при выполнении неравенства $\Delta < (2 - 2e^{-\tau} - \tau e^{-\tau})(2e^{-\tau} - 2 + \tau) 0,5(1 - e^{-\tau})^{-2}$. Эти два условия выделяют в плоскости $\tau\Delta$ область O_4 (рис. 5), ограниченную линиями l_5 , l_6 . Линия l_6 описывается равенством

$$\Delta = (2 - 2e^{-\tau} - \tau e^{-\tau})(2e^{-\tau} - 2 + \tau) 0,5(1 - e^{-\tau})^2, \quad (35)$$

а область O_4 — неравенствами:

$$\Delta < \begin{cases} \Delta^* & \text{при } \tau > \tau^* \\ \frac{(2 - 2e^{-\tau} - \tau e^{-\tau})(2e^{-\tau} - 2 + \tau)}{2(1 - e^{-\tau})^2} & \text{при } \tau < \tau^* \end{cases}. \quad (36)$$

д) Пусть T_5 — отображение, соответствующее элементарной последовательности $\Gamma_1^1 \xrightarrow{-1} \Gamma_3 \xrightarrow{+1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^2 \xrightarrow{0} \Gamma_1^1$, и пусть (u_0, v_0) , (u, v) , (u_1, v_1) , (u_2, v_2) , (u_3, v_3) , (u_4, v_4) — точки соответствующей траектории, принадлежащие линиям Γ_1^1 , Γ_3 , Γ_2^2 , Γ_2^1 , Γ_1^2 , Γ_1^1 . Согласно (4), (5) отображение T_5 описывается равенствами

$$u = u_0 + (1 - e^{-\tau})v_0 + 1 - \tau - e^{-\tau}, \quad v = e^{-\tau}v_0 - 1 + e^{-\tau}, \quad 2u_0v_0 = \Delta,$$

$$\begin{aligned}
 u_1 &= u + (1 - e^{-t}) v - 1 + t + e^{-t}, \quad v_1 = e^{-t} v + 1 - e^{-t}, \quad 2u_1 v_1 = \Delta, \\
 u_2 &= v_1, \quad v_2 = u_1, \\
 u_3 &= u_2 + [1 - \exp(-t_1)] v_2 - 1 + t_1 + \exp(-t_1), \\
 v_3 &= \exp(-t_1) v_2 + 1 - \exp(-t_1), \\
 u_4 &= v_3, \quad v_4 = u_3.
 \end{aligned} \tag{37}$$

Из анализа траекторий системы (3) следует, что отображение T_5 описывается непрерывной монотонно убывающей функцией последования. Следовательно, отображение T_5 может иметь не более одной простой неподвижной точки. Неподвижная точка u_5^* отображения T_5 определяется из (37) постановкой $u_0 = u_4 = u_5^*$, $t + t_1 = \tau$. В результате этой подстановки получаем два равенства, позволяющие по параметрам τ , Δ найти t_1 и u_5^*

$$\begin{aligned}
 (1 - \exp(-t_1)) [2\exp(-\tau) - 2\exp(-2\tau) + \exp(t_1 - 2\tau) - \exp(-t_1) - \\
 - t_1 \exp(-2\tau)] &= \Delta/2 \{(\exp(-t_1) - \exp(-2\tau))(\exp(-t_1) - \exp(t_1 - 2\tau)) - \\
 - (1 - \exp(-\tau))^2 [1 + \exp(-\tau) - 2\exp(t_1 - \tau) - (1 - \exp(-\tau)) \times \\
 \times \exp(2t_1 - 2\tau)]^2 [(1 - \exp(-t_1))(\exp(-t_1) - \exp(t_1 - 2\tau)) 0,5 \Delta - \\
 - (1 - 2\exp(t_1 - \tau) + \exp(t_1 - 2\tau))(2\exp(-\tau) - \exp(-2\tau) - \\
 - t_1 \exp(-t_1) - \exp(-2t_1))]^{-1}\}, \tag{38}
 \end{aligned}$$

$$\begin{aligned}
 u_5^* &= \{(1 - \exp(-t_1))[\exp(-2t_1) - 2\exp(-\tau - t_1) - \\
 - \exp(-2\tau) + 2\exp(-2\tau - t_1) + t_1 \exp(-2\tau - t_1)] + 0,5 \Delta \times \\
 \times (\exp(-t_1) - \exp(-2\tau))(\exp(-2t_1) - \exp(-2\tau))\} \times \\
 \times \{(1 - \exp(-\tau))[\exp(-2t_1)(1 - \exp(-\tau)) - 2\exp(-\tau - t_1) - \\
 - \exp(-2\tau)(1 - \exp(-\tau))]^{-1}\}. \tag{39}
 \end{aligned}$$

Существование неподвижной точки u_5^* обусловлено двумя неравенствами $t \geq 0$ (или $t_1 \leq \tau$) и $u_1(u_5^*) \geq -u_c$. Эти условия выделяют в плоскости $\tau\Delta$ область O_5 (рис. 5), ограниченную линиями l_6 , l_7 . Равенство $u_1(u_5^*) = -u_c$ может быть записано так:

$$\begin{aligned}
 [1 - \exp(-\tau)][\exp(-t_1) - u_c + 1 - \exp(-t_1)]\{\exp(-2t_1) \times \\
 \times [1 + \exp(-\tau) - 2\exp(-\tau - t_1) - \exp(-2\tau)[1 - \exp(-\tau)]\} = \\
 = [\exp(-2t_1) - 2\exp(-\tau - t_1) - \exp(-2\tau) + 2\exp(-2\tau - t_1) + \tag{40} \\
 + t_1 \exp(-2\tau - t_1)] [1 - \exp(-t_1)] + \frac{\Delta}{2} [\exp(-t_1) - \exp(-2\tau)] \times \\
 \times [\exp(-2t_1) - \exp(-2\tau)].
 \end{aligned}$$

Линия l_7 описывается равенствами (6), (38), (40). По Δ находится u_c , затем по (38), (40) вычисляются t_1 , τ . Линия l_7 проходит через точку $(0, 0)$ плоскости $\tau\Delta$ и пересекает линию l_6 в точке с абсциссой $\tau = \tau^*$, являющейся корнем уравнения

$$\begin{aligned}
 \tau^2 [1 + \exp(-\tau)] \exp(-\tau) + 12[1 - \exp(-\tau)]^2 - \\
 - 2\tau [1 - \exp(-\tau)] [2 + 3\exp(-\tau)] = 0. \tag{41}
 \end{aligned}$$

Соответствующий u_3^* предельный цикл системы (3) образован одним отрезком траектории листа Φ_{+1} , двумя отрезками траекторий листа Φ_{+1} и двумя отрезками прямолинейных траекторий режимов скольжения (рис. 4 д). Аналитическое исследование устойчивости этого движения не проводилось; эксперимент на электронной модели показывает, что оно устойчиво.

е) Согласно [4] отображение T_6 , соответствующее элементарной последовательности переходов $\Gamma_1^1 \xrightarrow{-1} \Gamma_2^2 \xrightarrow{0} \Gamma_2^1 \xrightarrow{+1} \Gamma_1^1$, инвариантного множества не имеет.

6. Переходим к изучению отображения T в целом. Результаты исследований, описанных в разделе 4, позволяют установить, какие отображения T_i ($i = 1, 2, \dots, 6$) образуют T при фиксированных τ, Δ . В разделе [5] исследованы инвариантные множества элементарных отображений T_i . Теперь необходимо установить характер «сшивания» отображений T_i , образующих T , в граничных точках их областей определения, а именно — непрерывно это «сшивание» или разрывно, и если разрывно, то каков характер разрыва, так как характер разрыва существенно влияет на структуру инвариантных множеств отображений [6, 7].

Используя изложенные выше результаты, можно показать, что общая граничная точка областей определения существует для отображений T_2 и T_1 , T_3 и T_1 , T_6 и T_4 , T_4 и T_5 , T_3 и T_6 , а также для отображений T_4 и T_1 , T_6 и T_1 , T_4 и T_2 , T_5 и T_2 . При этом «сшивание» первых пяти пар отображений происходит непрерывно, а остальные четыре пары «сшиваются» с разрывом непрерывности, таким, что предел функции последовательности в точке разрыва слева больше предела справа.

Теперь можно определить вид отображения T в различных областях первого квадранта плоскости $\tau\Delta$.

Для $(\tau, \Delta) \in O_1$ отображение T может быть представлено одним из рис. 6а; бб; для $(\tau, \Delta) \in O_3$ — рис. 6в, г, д; для $(\tau, \Delta) \in O_4$ — рис. 6е, ж, з; Не-трудно убедиться, что в этих случаях неподвижные точки u_1^*, u_3^*, u_4^* абсолютно устойчивы.

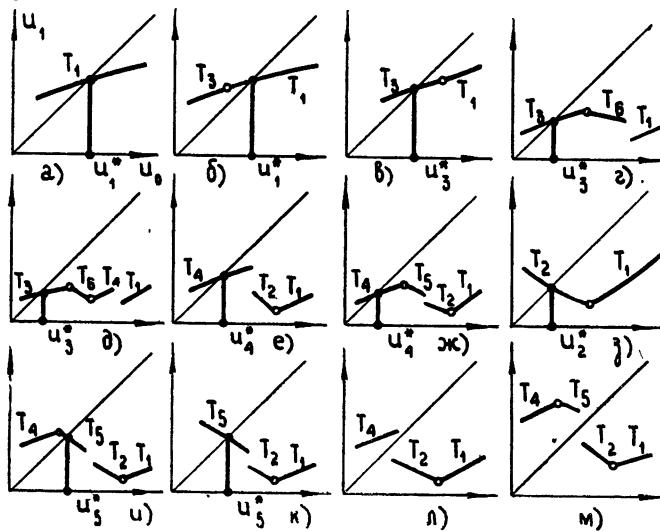


Рис. 6.

Для $(\tau, \Delta) \in O_2$ отображение T может быть представлено рис. 6з; для $(\tau, \Delta) \in O_5$ — рис. 6и, к. В этих случаях аналитически показать абсолютную устойчивость неподвижных точек u_2^*, u_5^* не удается. Однако,

эксперимент не обнаружил стационарных режимов, отличных от режимов, соответствующих неподвижным точкам u_2^* , u_5^* .

Для $(\tau, \Delta) \in O_6$ отображение T простых неподвижных точек не имеет и может быть представлено рис. 6 л. м. Согласно [6, 7] для отображений такого вида имеет место весьма сложное разбиение области O_6 на счетное множество областей существования кратных неподвижных точек отображения T различных кратностей до сколь угодно высоких порядков кратности.

ЛИТЕРАТУРА

1. Н. Н. Леонов, Изв. высш. уч. зав. — Радиофизика, 13, № 11, 1630 (1970).
2. Н. Н. Леонов, Изв. высш. уч. зав. — Радиофизика, 13, № 11, 1636 (1970).
3. Г. А. Новинская, Изв. высш. уч. зав. — Радиофизика, 4, № 3, 566 (1961).
4. Н. Н. Леонов, Изв. высш. уч. зав. — Радиофизика, 10, № 3, 393 (1967).
5. Н. Н. Леонов, Автоматика и телемеханика, № 12, 156 (1969).
6. Н. Н. Леонов, ДАН СССР, 143, № 5, 1038 (1962).
7. Н. Н. Леонов, О разрывном точечном отображении прямой в прямую, Кандидатская диссертация, Горький, 1962.

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
24 апреля 1970 г

ON EXTREMAL SYSTEMS WITH PREFERENCE SEARCH

N. N. Leonov

The article presents an investigation of character of motions in preference search extremal systems for stable plants with inertial input.

УДК 62 — 506

ОБ ОПТИМАЛЬНОМ УПРАВЛЕНИИ В ЭКСТРЕМАЛЬНЫХ СИСТЕМАХ

Н. Н. Леонов

Обсуждаются вопросы выбора критерия качества и выясняются условия оптимизации движения одного класса экстремальных систем

1. Экстремальные системы управления производственными процессами являются системами автоматической оптимизации определенных критериев качества этих процессов. Сами процессы автоматической оптимизации являются динамическими процессами, и поэтому можно ставить, в свою очередь, задачу оптимизации этих процессов, что сделано для экстремальных систем достаточно общего вида [1]. Эта задача, как и любая задача оптимизации движения, включает в себя два вопроса—выбор критерия качества и отыскание способа достижения и поддержания экстремального значения этого критерия. Ниже эти вопросы рассматриваются для экстремальных систем непрерывного, неэпизодического действия в том смысле, что экстремальное управление не прекращается в течение всего рабочего процесса.

Общепринятые критерии качества работы экстремальных систем совпадают с критериями, используемыми для оценки работы обычных следящих систем. Так, качество переходного процесса оценивается временем установления, быстродействием, а качество установившегося процесса слежения—некоторой усредненной оценкой рассогласования регулируемой и задающей величин. Однако основные цели следящих и экстремальных систем различны. Следящие системы призваны уменьшать рассогласование регулируемой и задающей величин, независимо от знака рассогласования. Поэтому естественно оценивать качество переходных и установившихся процессов в них среднеквадратичными показателями, и естественно требование быстродействия в переходном процессе. Экстремальные системы призваны уничтожать рассогласование регулируемой и задающей величин лишь одного знака. Поэтому использование для оценки их работы среднеквадратичных критериев качества может привести к ошибочным выводам, ибо среднеквадратичные критерии не различают рассогласования по знаку. Эта специфика экстремальных систем должна находить отражение при выборе критерии качества их работы.

Если в следящих системах естественно требование быстродействия при оценке переходных процессов, то в экстремальных системах непрерывного действия качество переходного процесса было бы естественнее оценивать теми же средними критериями, что и качество установившихся процессов, исходя из основной задачи экстремального управления—поддерживать на максимально (или минимально) возможном уровне регулируемую величину в течение всего процесса работы системы. С точки зрения сравнения двух критериев переходного процесса—быстродействия и среднего рассогласования—рассмотрим простейшую экстремальную систему. Влияние помех при этом не учитывается.

2. Уравнения движения экстремальной системы с объектом, обладающим инерционностью первого порядка на выходе, могут быть записаны так:

$$\dot{\varphi} + \varphi = -u^2, \quad \dot{u} = \eta, \quad \eta = G(u, \varphi, t). \quad (1)$$

Здесь u — вход объекта (рис. 1), φ — его выход, $G(u, \varphi, t)$ — оператор автоматического оптимизатора. В данном случае характеристика нелинейного звена объекта имеет экстремум — максимум. Случай экстремума — минимума рассматривается аналогично.

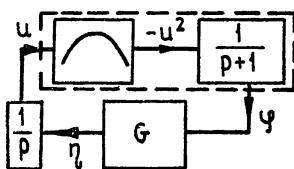


Рис. 1.

Естественно сделать предположение, что управление ограничено $|G(u, \varphi, t)| < h$. Примем за критерий качества работы рассматриваемой системы среднюю за любой промежуток времени длительности τ величину выхода объекта

$$\tau^{-1} \int_0^\tau \varphi(t) dt. \quad (2)$$

Найдем управление, максимизирующее величину (2). Решение системы (1) можно записать так

$$\varphi(t) = \varphi_0 e^{-t} - e^{-t} \int_0^t u^2(\tau) e^\tau d\tau, \quad (3)$$

$$u(t) = u_0 + \int_0^t G(u, \varphi, t) dt, \quad (4)$$

где $u_0 = u(0)$, $\varphi_0 = \varphi(0)$.

Очевидно, управление, максимизирующее в любой момент времени величину $\varphi(t)$ выхода объекта, максимизирует и критерий (2). Если такое управление существует, то из (3) следует, что максимизация $\varphi(t)$ происходит при минимизации $u^2(t)$, т. е. $|u(t)|$.

Таким образом, в рассматриваемом случае задача оптимизации качества работы системы свелась к задаче минимизации в каждый момент времени лишь величины координаты входа объекта $|u(t)|$. Управление, решающее эту задачу, как нетрудно проверить, может быть записано следующим образом

$$G(u, \varphi, t) = \begin{cases} -h \operatorname{sgn} u & (u \neq 0) \\ 0 & (u = 0) \end{cases}. \quad (5)$$

Фазовые траектории системы (1) с управлением (5) на плоскости $u\varphi$ имеют вид, изображенный на рис. 2. Каждую фазовую траекторию можно в этом случае разбить на два участка.

Для первого из них скорость изменения координаты входа объекта постоянна и равна h по величине, а для второго, совпадающего с частью оси φ , вход объекта u не меняется.

3. Найдем теперь управление, удовлетворяющее требованию быстродействия переходного процесса в системе (1).

Очевидно, управление, минимизирующее время движения фазовой точки из некоторой об-

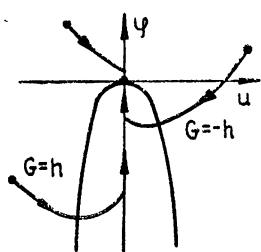


Рис. 2.

ласти фазового пространства до некоторой точки A , существует, если для любой точки этой области существует, по крайней мере, одно управление, при котором фазовая точка приходит в точку A за конечный промежуток времени.

В рассматриваемой задаче было бы естественно за точку A выбрать точку $(0, 0)$ плоскости $\psi\varphi$, соответствующую экстремальной точке нелинейной статической характеристики объекта управления. Однако при этом область Q конечного времени поиска ограничена в плоскости $\psi\varphi$ полутраекториями, проходящими через точку $A(0, 0)$ (рис. 3).

С помощью принципа максимума [2] можно найти управление, минимизирующее время перехода фазовой точки из любой точки области Q в точку A . Фазовые траектории, соответствующие этому управлению, изображены на рис. 3. Каждая фазовая траектория состоит здесь из двух отрезков траекторий с $G = +h$ и $G = -h$, проходящих через начальную точку и точку A .

Остальная часть плоскости $\psi\varphi$ не является для точки $A(0, 0)$ областью конечного времени поиска. Для любого начального состояния, не принадлежащего области Q , время достижения фазовой точкой $A(0, 0)$ не ограничено.

Подходящим выбором конечной точки A можно добиться, чтобы не только область Q , но и вся плоскость $\psi\varphi$ стала областью конечного времени поиска. Выбор новой конечной точки A_ϵ может быть сделан на основе того, что в реальных экстремальных системах слежение за экстремальным значением выхода объекта происходит с некоторой ошибкой. В случае экстремума—максимума при работе системы в установившемся режиме слежения среднее значение выхода объекта, поддерживаемое оптимизатором, отличается от максимально возможного значения выхода объекта на величину средней ошибки слежения, которую обозначим через $-\epsilon$ ($\epsilon > 0$). Поэтому за новую конечную точку A_ϵ естественно выбрать точку $(0, -\epsilon)$.

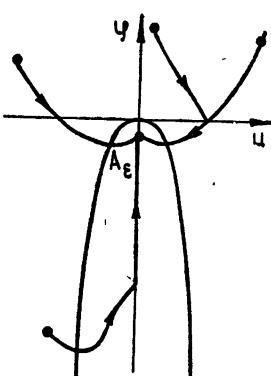


Рис. 4.

Для конечной точки A_ϵ область Q переходит в область Q_ϵ (рис. 4). Для любого начального состояния, не принадлежащего области Q_ϵ , расположенного в ограниченной части плоскости $\psi\varphi$, управление (5) обеспечивает конечное время поиска точки A_ϵ .

Управление, минимизирующее время поиска в области Q_ϵ , имеет тот же вид, что и в области Q .

Обозначим часть плоскости $\psi\varphi$, являющуюся дополнением области Q_ϵ до всей плоскости, через P . Покажем, что управление (5) минимизирует время поиска в P .

Если из множества допустимых управлений выбрать управление $G_1(u_1, \varphi_1, t)$, отличное от управления $G(u, \varphi, t)$, определяемого выражением (5), то согласно (3) для движений, начинающихся из общего начального состояния,

$$\varphi - \varphi_1 = e^{-t} \int_0^t e^{\tau} (u_1^2 - u^2) d\tau. \quad (6)$$

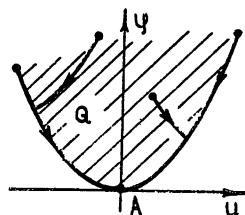


Рис. 3.

Если управление G_1 отлично от G хотя бы на некотором интервале (t_1, t_2) , то на этом интервале $u^2 < u_1^2$. Следовательно, на основании (6), $\varphi(t) > \varphi_1(t)$ для $t > t_1$. Это говорит о том, что управление (5) максимизирует координату $\varphi(t)$ в любой момент времени, и при этом управлении точка $A_1(0, -\epsilon)$ достигается фазовой точкой за минимально возможное время для любого начального состояния из P .

При $\epsilon \rightarrow 0$ $A_\epsilon \rightarrow A$, область Q_ϵ сужается до области Q , а время поиска для начальных состояний из области P неограниченно возрастает, ибо точка $A(0, 0)$ является состоянием равновесия для траектории $u = 0$.

4. Итак, установлено, что управление (5) максимизирует среднее отклонение выхода объекта φ от его экстремального значения в статике φ_0 для любого начального состояния системы (1).

Управление, минимизирующее время перехода фазовой точки из некоторого начального состояния системы в конечную точку $A_\epsilon(0, -\epsilon)$, соответствующую экстремальному значению выхода объекта φ_0 , существенно зависит от начального состояния. Фазовое пространство в этом случае разбивается на две области. Для одной из них управление, решающее задачу быстродействия, совпадает с управлением (5), максимизирующим величину (2). Для другой области управление, решающее задачу быстродействия в переходном процессе, существенно отличается от управления (5). Управление, минимизирующее в этой области время поиска, не решает основную задачу, стоящую перед автоматическими оптимизаторами, ибо не является максимизирующим выходную координату объекта.

5. Изучение оптимального поведения системы (1) проводилось выше в предположении наличия полной информации о поведении этой системы. Однако экстремальные системы характерны возможностью получения лишь существенно неполной информации об их поведении. Поэтому возникает вопрос о возможности реализации найденных оптимальных управлений.

Управление (5) может быть реализовано лишь в том случае, если в каждый момент времени будет известен $\operatorname{sgn} u$. Для определения знака отклонения выхода объекта от его значения, соответствующего экстремуму выхода, в цепь управления необходимо ввести звено с оператором, обратным оператору линейной части объекта управления, поместив его после объекта перед оптимизатором [3]. Введение в цепь управления экстремальных систем звена с оператором, обратным оператору линейной части объекта управления, с целью улучшения качества работы экстремальных систем предлагалось в [4, 5].

Введение указанного звена в цепь управления вместе с дифференцирующим звеном позволяет получить на входе оптимизатора сигнал $\sigma = \dot{\varphi} + \varphi$, который согласно (1) равен $-2u\eta$. Величина η — скорость изменения входа объекта — доступна измерению, и $\operatorname{sgn} \sigma = -2 \operatorname{sgn} \eta \operatorname{sgn} u$ характеризует направление отклонения входа объекта от его значения, соответствующего экстремуму выхода.

Таким образом, управление (5), оптимизирующее по неполной информации о поведении объекта, показатель (2) качества работы системы (1), реализуемо при условии реализуемости с достаточной степенью точности звена с оператором $r^2 + r$.

Управление, решающее задачу быстродействия в переходном процессе на всем множестве начальных состояний, требует для своей реализации знания величин u , φ в любой момент времени, т. е. требует наличия полной информации о поведении системы. Это говорит о невозможности реализации подобного управления в экстремальных системах.

6. В заключение следует отметить, что в экстремальных системах с объектами, отличными от рассмотренного в (1), оптимальные процессы не реализуемы. Так, в экстремальной системе с объектом, обладающим инерционностью, расположенной перед нелинейностью, обладающей для реализации оптимальности переходного процесса наличие полной информации о поведении системы. Для объектов типа, рассмотренного выше, но с более высоким порядком инерционности, переходный процесс уже не является апериодическим, и в этом случае даже возможность реализации звена с оператором, обратным оператору линейной части объекта, не позволит реализовать оптимальное управление из-за недостатка информации о поведении системы.

При этом для систем, объекты которых отличны от описанного в (1), не существует управления, оптимизирующего (2) для любого τ . Для таких систем величина τ не должна быть меньше времени установления.

ЛИТЕРАТУРА

1. А. А. Фельдбаум, Вычислительные устройства в автоматических системах, Физматгиз, 1959.
2. Л. С. Понtryгин, В. Г. Болтянский, Р. В. Гамкrelidze, Е. Ф. Мышленко, Математическая теория оптимальных процессов, Физматгиз, 1961.
3. Н. Н. Леонов, Автоматика и телемеханика, 26, № 10, 1720 (1965).
4. О. М. Крыжановский, В. М. Кунцевич, Автоматика, № 3 (1958).
5. В. В. Казакевич, ДАН СССР, 133, № 4, 756 (1960).

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
13 апреля 1970 г

ON OPTIMAL CONTROL IN EXTREMAL SYSTEMS

N. N. Leonov

Problems of quality criterion choice are discussed and conditions for motion optimization in one class of extremal systems are found.

УДК 62 — 506

**ПРИМЕНЕНИЕ РАЗНОСТНЫХ УРАВНЕНИЙ
ДЛЯ ИССЛЕДОВАНИЯ УСТАНОВИВШЕГОСЯ РЕЖИМА РАБОТЫ
ГРАДИЕНТНОЙ ЭКСТРЕМАЛЬНОЙ СИСТЕМЫ
С ИНЕРЦИОННЫМ ОБЪЕКТОМ**

A. M. Кориков

Дискретная экстремальная система градиентного типа при наличии помех на входе и выходе объекта исследуется при помощи разностных уравнений. Рассматриваются две разновидности инерционных объектов. Доказаны теоремы, позволяющие исследовать установившийся режим работы системы при помощи линейного разностного уравнения, являющегося линейной частью уравнения, описывающего процесс отслеживания экстремума. Исследуется влияние инерционности объекта на работу системы.

ВВЕДЕНИЕ

Процесс отслеживания экстремума в дискретных экстремальных системах адекватно описывается разностными уравнениями. Разностные уравнения, непосредственно описывающие процесс отслеживания экстремума, успешно применялись для исследования дискретных экстремальных систем градиентного типа в ряде работ [1—4]. В этих работах объект управления считался безынерционным. В статье [5] проведен анализ экстремальной системы с многомерным инерционным объектом при предположении, отмеченном ниже. В настоящей статье экстремальная система с инерционным объектом, как и в работах [1—5], исследуется при помощи разностных уравнений.

Функциональная схема экстремальной системы показана на рис. 1 а. Инерционный экстремальный объект управления обычно представляют

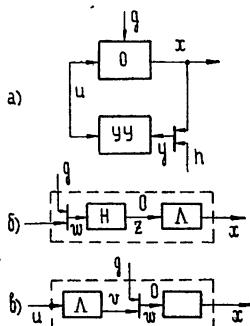


Рис. 1. а) — функциональная схема экстремальной системы: O — объект управления, UU — управляющее устройство; б, в) структура инерционного объекта управления: H — экстремальное безынерционное звено, L — линейная инерционная часть.

как последовательное соединение экстремального безынерционного звена и линейной инерционной части. Различают следующие случаи: 1) инерционная часть расположена после экстремального звена (рис. 1 б); 2) инерционная часть расположена до экстремального звена (рис. 1 в); 3) инерционные части расположены до и после экстремального звена. Ограничимся исследованием первых двух случаев. Экстрем-

мальную систему с объектом, показанным на рис. 1 б, будем называть кратко: система А, а экстремальную систему с объектом, показанным на рис. 1 в,—система Б. Экстремальную характеристику считаем квадратичной, т. е. $z = w^2 = (u + g)^2$ для системы А и $x = w^2 = (v + g)^2$ для системы Б. Относительно помехи h на выходе объекта предполагаем, что она аддитивная и ее приращения являются независимыми случайными величинами с математическим ожиданием (МО), равным нулю, и конечной дисперсией σ^2 . Помеха g вызывает дрейф экстремальной характеристики с некоторой постоянной скоростью γ , т. е. $g(t) = \gamma t$.

Управляющие воздействия вырабатываются управляющим устройством в моменты времени $t = n\tau$, где n —номер цикла, τ —длительность цикла (время определения градиента). Разобьем каждый цикл на два интервала длительности $\tau/2$. Через $L[2n+j]$ будем обозначать величину L в момент времени $(2n+j)\tau/2$. Алгоритм работы управляющего устройства следующий: в начале первого интервала n -го цикла на вход объекта подается пробное возмущение a_0 , в начале второго интервала n -го цикла—пробное возмущение $-a_0$; в конце первого и второго интервала n -го цикла измеряется выход объекта управления и управляющее устройство образует разность

$$\begin{aligned} \Delta y[2n] &= y[2n-1] - y[2n] = x[2n-1] - x[2n] + \\ &+ h[2n-1] - h[2n] = \Delta x[2n] + \Delta h[2n]; \end{aligned} \quad (1)$$

в начале следующего $(n+1)$ -го цикла на вход объекта подается управляющее воздействие $u[2n]$, такое, что его приращение

$$\Delta u[2n] = u[2n] - u[2(n-1)] = -a\Delta y[2n], \quad (2)$$

где a —коэффициент рабочего шага (управляющего воздействия). Легко проверить, что (2) можно получить и при помощи синхронного детектора, если поисковое колебание является прямоугольным периодическим воздействием.

Будем предполагать, что помеха g изменяется на величину $\gamma\tau/2$ только в моменты времени $(2n+j)\tau/2$ ($j = 0, 1; n = 0, 1, 2, \dots$), тогда все переменные в экстремальной системе изменяются в моменты времени $(2n+j)\tau/2$, поэтому удобным является следующее описание линейных инерционных частей объекта:

$$v[2n+j] = \sum_{s=0}^{2n+j-1} b[2n+j-s] u[s] = \sum_{s=1}^{2n+j} b[s] u[2n+j-s], \quad (3)$$

$$x[2n+j] = \sum_{s=0}^{2n+j-1} c[2n+j-s] z[s] = \sum_{s=1}^{2n+j} c[s] z[2n+j-s] \quad (4)$$

$$(j = 0, 1; n = 0, 1, 2, \dots).$$

Здесь $b[s]$, $c[s]$ —весовые коэффициенты инерционных частей.

Считаем, что система находилась в покое до момента $t = 0$. В работе [5] предполагалось следующее описание инерционной части объекта:

$$\begin{aligned} v[2n+j] &= \sum_{s=0}^{\alpha} b'[2s] v[2n+j-2s] \\ (j &= 0, 1; n = 0, 1, 2, \dots). \end{aligned}$$

Величина α равна длительности переходного процесса в инерционном объекте.

Эффективность работы экстремальной системы в установившемся режиме ($n \rightarrow \infty$) будем характеризовать МО выхода объекта. В силу

квадратичной экстремальной характеристики МО выхода объекта определяется МО и дисперсией входной переменной экстремального безынерционного звена.

1. ИССЛЕДОВАНИЕ СИСТЕМЫ А

Для входной переменной объекта запишем цепочку равенств

$$\begin{aligned} w[2(n+1)] &= w[2n] + \gamma\tau + \Delta u[2(n+1)] = \\ &= w[2n] + \gamma\tau - a(\Delta x[2(n+1)] + \Delta h[2(n+1)]). \end{aligned} \quad (5)$$

При помощи (4) получим

$$\Delta x[2(n+1)] = \sum_{s=0}^{2n} c[2n+1-s] \Delta z[s+1] - z[0] c[2(n+1)], \quad (6)$$

где $\Delta z[s+1] = z[s] - z[s+1]$, причем

$$\Delta z[s+1] = \left(w[s] + \frac{\gamma\tau}{2} + a_0 \right)^2 - (w[s] + \gamma\tau - a_0)^2,$$

если $s = 0, 2, \dots, 2n$, и

$$\Delta z[s+1] = (w[s-1] + \gamma\tau - a_0)^2 - \left(w[s+1] + \frac{\gamma\tau}{2} + a_0 \right)^2,$$

если $s = 1, 3, \dots, 2n-1$. Введем обозначения:

$$\Theta = 4a_0 - \gamma\tau, \quad \eta = 2a_0 + \gamma\tau, \quad x = 2(\gamma\tau - a_0); \quad (7)$$

$$\alpha'_{n,s} = c[2(n-s)] - c[2(n-s+1)]; \quad (8)$$

$$\beta'_{n,s} = \Theta c[2(n-s)+1] - \eta c[2(n-s+1)] + xc[2(n-s)]; \quad (9)$$

$$\begin{aligned} \lambda'_n &= \frac{3}{4} \gamma\tau \Theta \left(\sum_{s=0}^n c[2(n-s)+1] - \sum_{s=0}^{n-1} c[2(n-s)] \right) - \\ &\quad - \frac{\eta^2}{4} c[2(n+1)]. \end{aligned} \quad (10)$$

При использовании обозначений (8), (9) следует помнить, что $c[0] = 0$. В обозначениях (8)–(10) формула (6) примет следующий вид:

$$\Delta x[2(n+1)] = \sum_{s=0}^n (\alpha'_{n,s} w^2[2s] + \beta'_{n,s} w[2s]) + \lambda'_n.$$

Подставляя это выражение в (5) и вводя обозначения

$$\alpha_{n,s} = -a\alpha'_{n,s}, \quad \beta_{n,s} = \delta_{0,n-s} - a\beta'_{n,s},$$

где $\delta_{n,s}$ — символ Кронекера, и $\lambda_n = \gamma\tau - a\lambda'_n$, получим следующее разностное уравнение:

$$w[2(n+1)] = \sum_{s=0}^n (\alpha_{n,s} w^2[2s] + \beta_{n,s} w[2s]) + \lambda_n - a\Delta h[2(n+1)]. \quad (11)$$

Сделаем дополнительное предположение об инерционной части объекта управления. Пусть $c[s] = c_0 c^s$. Если инерционная часть является апериодическим звеном с передаточной функцией

$$C(p) = \frac{k}{Tp + 1},$$

где p — произвольный комплексный параметр, T , k — постоянная времени и коэффициент усиления апериодического звена соответственно, то

$$c_0 = k [\exp(\tau/2T) - 1], \quad c = \exp(-\tau/2T). \quad (12)$$

Вычитая из (11) выражение для $c^2 w[2n]$, после несложных преобразований получим

$$\begin{aligned} w[2(n+1)] &= \sum_{s=0}^1 (\rho_s w^2[2(n-s)] + \nu_s w[2(n-s)]) + \lambda + \\ &+ \delta_{0,n} (\lambda_0 - \lambda - c^2 w[0]) - a(\Delta h[2(n+1)] - (1 - \delta_{0,n}) c^2 \Delta h[2n]) \\ &\quad (n = 0, 1, 2, \dots). \end{aligned} \quad (13)$$

Здесь

$$\begin{aligned} \rho_0 &= -\rho_1 = ac_0 c^2, \\ \nu_0 &= 1 + c^2 - ac_0 c (\Theta - \eta c), \quad \nu_1 = -(1 + ac_0 c) c^2, \\ \lambda &= \gamma \tau (1 - c^2) [1 - (3/4) a k \Theta (1 - c) (1 + c)^{-1}]. \end{aligned}$$

Введем в рассмотрение разностное уравнение

$$\begin{aligned} w^*[2(n+1)] &= \sum_{s=0}^1 \nu_s w^*[2(n-s)] + \lambda + \delta_{0,n} (\lambda_0 - \lambda - c^2 w[0]) - \\ &- a(\Delta h[2(n+1)] - (1 - \delta_{0,n}) c^2 \Delta h[2n]), \end{aligned} \quad (14)$$

являющееся линейной частью уравнения (13). При помощи (13) и (14) образуем последовательности случайных величин (ПСВ)

$$w[0], w[2], \dots, w[2n], w[2(n+1)], \dots; \quad (15)$$

$$w^*[0], w^*[2], \dots, w^*[2n], w^*[2(n+1)], \dots. \quad (16)$$

Будем называть ПСВ $\{\xi[s]\}$ ($s = 0, 1, 2, \dots$) устойчивой, если ее предельные МО и дисперсия ограничены, т. е.

$$m_\xi = \lim_{s \rightarrow \infty} m_\xi[s] = \lim_{s \rightarrow \infty} M\xi[s] < \infty,$$

$$D_\xi = \lim_{s \rightarrow \infty} D_\xi[s] = \lim_{s \rightarrow \infty} M(\xi[s] - m_\xi[s])^2 < \infty.$$

Здесь M означает математическое ожидание (МО). Используя (13), (14), как рекуррентные соотношения, получим

$$\begin{aligned} w[2(n+1)] &= \rho_0 \sum_{s=0}^{n-1} \rho_s (w^2[2(n-s)] - w^2[2(n-1-s)]) + \\ &+ \rho_0 p_n w^2[0] + p_{n+1} w[0] + p_n (\lambda_0 - \lambda - c^2 w[0]) + q_n \lambda - a \sum_{s=0}^n r_s \times \\ &\times \Delta h[2(n+1-s)]; \end{aligned} \quad (17)$$

$$\begin{aligned} w^*[2(n+1)] = & p_{n+1}w^*[0] + p_n(\lambda_0 - \lambda - c^2w[0]) + q_n\lambda - \\ & - a \sum_{s=0}^n r_s \Delta h [2(n+1-s)]; \end{aligned} \quad (18)$$

где

$$\begin{aligned} p_{n+1} = & v_0 p_n + v_1 p_{n-1}, \quad p_{-n} = \delta_{0,-n} \\ (n = 0, 1, 2, \dots); \end{aligned} \quad (19)$$

$$q_n = \sum_{s=0}^n p_s; \quad (20)$$

$$r_s = p_s - c^2 p_{s-1}, \quad r_0 = 1. \quad (21)$$

Теорема 1. Если ПСВ (15), (16) устойчивые, то $m_w = m_{w^*}$, а $D_w \rightarrow D_{w^*}$ при $\tau/T \rightarrow \infty$ или $\tau/T \rightarrow 0$.

Теорема 2. Экстремальная система устойчива в том и только том случае, если ПСВ (16) является устойчивой.

Доказательства теорем содержатся в Приложении.

Из теоремы 2 следует, что область устойчивости системы А совпадает с областью сходимости ряда (20). Используя неравенство (П.5), приведенное в Приложении, получим: система А устойчива, если ее параметры удовлетворяют неравенствам

$$0 < a(4a_0 - \gamma\tau) < 2k(1 - e^{-\gamma/2T})^{-2}. \quad (22)$$

Отсюда следует, что экстремальная система с инерционным объектом более устойчива и менее «критична» к выбору параметров управляющего устройства, чем система с безынерционным объектом.

Теорема 1 позволяет исследовать установившийся режим работы системы А с помощью (18), если инерционностью объекта можно пренебречь и если инерционность объекта значительна. В этом случае $m_x = \lim_{n \rightarrow \infty} Mx[n] = k(m_{w^*}^2 + D_{w^*})$. Из выражения (П.8) (см. Приложение) следует, что максимальное значение ρ_v^2 , определяющее верхнюю оценку для D_w , мало и равно 0,0625 $a^2 k^2$. Следовательно, если интенсивность помехи h мала, то выражения

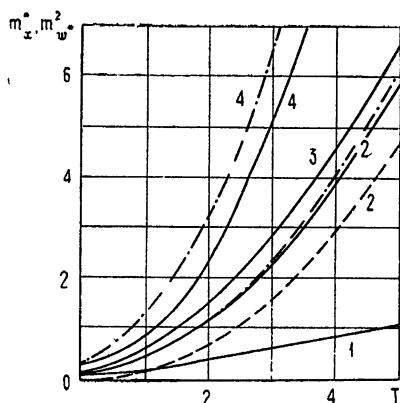
$$\begin{aligned} m_x^* = & k(m_{w^*}^2 + D_{w^*}), \quad m_{w^*} = \frac{\lambda}{1 - v_0 - v_1}, \\ D_{w^*} = & a^2 \sigma^2 \frac{(1 - v_1)(1 + e^{-2\gamma/T}) - 2v_0 e^{-\gamma/T}}{(1 + v_1)[(1 - v_1)^2 - v_0^2]} \end{aligned}$$

позволяют оценить МО выхода объекта и влияние инерционности объекта на эффективность работы системы А при любом значении τ/T .

Постоянная времени T апериодического звена определяет длительность переходных процессов в объекте управления и может служить мерой инерционности системы. На рис. 2 сплошными линиями показаны зависимости $m_x^*(T)$ при различных параметрах системы. Из сравнения сплошных линий 1, 2 следует, что наличие дрейфа экстремума инерционного объекта значительно снижает эффективность работы системы. При этом в формировании величины m_x^* наибольший вклад вносит $m_{w^*}^2$ (штриховая линия 2 показывает зависимость $m_{w^*}^2(T)$). Из сравнения сплошных линий 2, 3, 4 следует наличие оптимального интервала τ между рабочими шагами. Дополнительные расчеты показывают нали-

чие оптимального значения коэффициента рабочего шага a . Зависимости $m_x^*(a)$ и $m_w^*(\tau)$ имеют такой же вид, что и в работах [1, 2, 5].

Рис. 2. Влияние инерционности объекта на эффективность работы экстремальной системы в установившемся режиме при $k=1$, $\sigma^2=0,1$, $a=1$, $a_0=0,25$; 1) $\gamma=0$, $\tau=1$; 2) $\gamma=0,1$, $\tau=1$; 3) $\gamma=0,1$, $\tau=0,5$; 4) $\gamma=0,1$, $\tau=5$;
 m_x^* для системы А; m_w^2 для системы А; m_x^* для системы Б.



2. АНАЛИЗ СИСТЕМЫ Б

Функциональная схема системы соответствует рис. 1 а, в. Для входной переменной экстремального безынерционного звена справедливо равенство

$$w[2(n+1)] = w[2n] + \gamma\tau + \Delta v[2(n+1)], \quad (23)$$

где $\Delta v[2(n+1)] = v[2(n+1)] - v[2n]$. При помощи (3) получим

$$\Delta v[2(n+1)] = \sum_{s=2}^{2n+1} b[2(n+1)-s] \Delta u[s] + \tilde{u}[0], \quad (24)$$

где $\Delta u[s] = u[s] - u[s-2]$, а величина

$$\tilde{u}[0] = \sum_{s=0}^1 b[2(n+1)-s] u[s]$$

определяется работой системы в начальном цикле. Определим приращение управляющего воздействия. Пусть $b[s] = b_0 b^s$, тогда

$$w[2s+1] = bw[2s] + (1-b)\gamma\tau s + b_0 b(u[2s] + a_0) + \frac{\gamma\tau}{2},$$

$$w[2(s+1)] = b^2 w[2s] + (1+b)\{(1-b)\gamma\tau s + b_0 bu[2s]\} - b_0 b(1-b)a_0 + \gamma\tau,$$

отсюда

$$w[2s+1] = \frac{1}{1+b} \left\{ w[2(s+1)] + bw[2s] + 2b_0 ba_0 - \frac{1-b}{2} \gamma\tau \right\}.$$

Возводя это выражение в квадрат и вычитая из него $w^2[2(s+1)]$, получим

$$\Delta x[2(s+1)] = w^2[2s+1] - w^2[2(s+1)].$$

Подставляя последнее выражение в (1), а затем (1) в (2), найдем приращение управляющего воздействия в виде

$$\begin{aligned} \Delta u[2s] = & \frac{-a}{(1+b)^2} \left\{ b^2(w^2[2(s-1)] - w^2[2s]) + \right. \\ & + 2bw[2s](w[2(s-1)] - w[2s]) + \left(2a_0b_0b - \frac{1-b}{2}\gamma\tau \right)^2 + \\ & \left. + (4a_0b_0b - \gamma\tau(1-b))w[2s] + bw[2(s-1)] \right\} - a\Delta h[2s]. \end{aligned}$$

Подставляя это выражение в (24), учитывая, что $\Delta u[2s+1] = \Delta u[2s]$, а затем подставляя (24) в (23) и производя те же преобразования, что и в разделе 1, получим

$$\begin{aligned} w[2(n+1)] = & \sum_{s=0}^1 \pi_s (w^2[2n] - w[2(n-1)]w[2(n-1+s)]) + \\ & + \sum_{s=0}^1 \mu_s w[2(n-s)] + \delta - ab_0b(1+b)\Delta h[2n] \quad (n=1, 2, \dots). \end{aligned} \quad (25)$$

Здесь

$$\begin{aligned} \pi_0 &= \frac{ab_0b^3}{1+b}, \quad \pi_1 = \frac{2ab_0b^2}{1+b}, \\ \mu_0 &= 1 + b^2 - \frac{ab_0b}{1+b}[4a_0b_0b - \gamma\tau(1-b)], \\ \mu_1 &= -b^2 \left[1 + \frac{ab_0}{1+b}(4a_0b_0b - \gamma\tau(1-b)) \right], \\ \delta &= \gamma\tau(1-b^2) - \frac{ab_0b}{1+b} \left(2a_0b_0b - \frac{1-b}{2}\gamma\tau \right)^2, \\ w[2] &= \mu_0 w[0] + \delta + \epsilon, \end{aligned} \quad (26)$$

где

$$\epsilon = (1 - \mu_0)w[0] + \gamma\tau - \delta + \tilde{u}[0].$$

Для ПСВ $\{w[2n]\}$ ($n=0, 1, 2, \dots$), полученной при помощи (25), (26), и ПСВ $\{w^*[2n]\}$ ($n=0, 1, 2, \dots$), образованной при помощи (26) и разностного уравнения

$$\begin{aligned} w^*[2(n+1)] = & \sum_{s=0}^1 \mu_s w^*[2(n-s)] + \delta - ab_0b(1+b)\Delta h[2n] \\ & (n=1, 2, \dots), \end{aligned} \quad (27)$$

справедливы утверждения, аналогичные теоремам 1, 2 раздела 1, которые позволяют исследовать установившийся режим работы системы Б при помощи уравнения (27). Решение уравнения (27) при начальных условиях (26) имеет вид

$$\begin{aligned} w^*[2(n+1)] = & p_{n+1}w^*[0] + p_n\epsilon + q_n\delta - \\ & - ab_0b(1+b) \sum_{s=0}^{n-1} p_s \Delta h[2(n-s)], \end{aligned} \quad (28)$$

где p_n и q_n по-прежнему определяются формулами (19), (20) с заменой ν_i на μ_i ($i = 0, 1$). Область устойчивости системы Б совпадает с областью сходимости ряда (20). Так как инерционная часть является апериодическим звеном, то

$$b_0 = k(e^{\gamma/2T} - 1), \quad b = e^{-\gamma/2T}; \quad (12a)$$

система Б устойчива, если ее параметры удовлетворяют неравенствам

$$0 < a(4a_0k - \gamma\tau) < \frac{2}{k}(1 - e^{-\gamma/2T})^{-2}. \quad (29)$$

При $k = 1$ или $a_0^A = ka_0^B$ (верхний индекс указывает на принадлежность параметра системе А или системе Б) области устойчивости систем А и Б совпадают. Заметим, что при $k = 1$ и $T \rightarrow 0$ неравенства (22), (29) определяют условие устойчивости экстремальной системы с безынерционным объектом, полученное в работе [2].

В области устойчивости системы математическое ожидание выхода объекта при $n \rightarrow \infty$ определяется следующими формулами:

$$m_x^* = m_{w^*}^2 + D_{w^*}, \quad (30)$$

$$m_{w^*} = \frac{\delta}{1 - \mu_0 - \mu_1}, \quad (31)$$

$$D_{w^*} = \frac{a^2 \sigma^2 k^2 (1 - \mu_1) (1 - e^{-\gamma/T})^2}{(1 + \mu_1) [(1 - \mu_1)^2 - \mu_0^2]}. \quad (32)$$

Численные расчеты, проведенные по формулам (30) — (32), показали, что зависимости $m_x^*(T)$, $m_x^*(\tau)$, $m_x^*(a)$ системы Б качественно не отличаются от соответствующих зависимостей системы А. Штрих-пунктирные линии 2, 4 на рис. 2 иллюстрируют зависимость $m_x^*(T)$ системы Б. Штрих-пунктирная линия 1 на рис. 2 слилась со сплошной линией 1, т. е. при отсутствии дрейфа экстремума математическое ожидание выхода систем А и Б отличается незначительно. Отличие штрих-пунктирных линий 2, 4 от соответствующих сплошных линий отражает различное влияние дрейфа экстремума на эффективность работы систем А и Б.

3. ОСНОВНЫЕ РЕЗУЛЬТАТЫ

1) Показано, что процесс отслеживания экстремума в градиентной экстремальной системе с инерционным объектом описывается нелинейным разностным уравнением.

2) Доказаны теоремы, позволяющие исследовать установившийся режим работы системы при помощи линейного разностного уравнения с постоянными кэофициентами, являющимся линейной частью уравнения, описывающего процесс отслеживания экстремума.

3) Показано, что инерционность объекта понижает качество работы системы и повышает устойчивость экстремальной системы.

4) Выяснено, что дрейф экстремума значительно снижает эффективность работы экстремальной системы с инерционным объектом.

В заключение автор выражает глубокую благодарность Г. А. Медведеву за внимание к данной работе и ценные замечания.

ПРИЛОЖЕНИЕ

1) Доказательство теоремы 1. Введем в рассмотрение случайные величины

$$w^* = q\lambda - a \sum_{s=0}^{\infty} r_s \Delta h_s, \quad w = w^* + \nabla w,$$

где

$$q = \lim_{n \rightarrow \infty} q_n, \quad \Delta h_s = \lim_{n \rightarrow \infty} \Delta h [2(n+1-s)],$$

$$\nabla w = p_0 \sum_{s=0}^{\infty} p_s \lim_{n \rightarrow \infty} (w^2 [2(n-s)] - w^2 [2(n-s-1)]).$$

Из (17), (18) следует, что если ПСВ (15), (16) устойчивые, то ряд (20) сходится. Тогда

$$\lim_{n \rightarrow \infty} Mw [2n] = Mw = \lim_{n \rightarrow \infty} Mw^* [2n] = Mw^* = q\lambda, \quad (\text{П.1})$$

так как $M\nabla w = 0$. Первое утверждение теоремы доказано.

В свою очередь,

$$\lim_{n \rightarrow \infty} D_{w^*} [2n] = D_{w^*} = a^2 \sigma^2 \sum_{s=0}^{\infty} r_s^2, \quad (\text{П.2})$$

$$\lim_{n \rightarrow \infty} D_w [2n] = D_w = D_{w^*} + D_{\nabla w} + 2K_{w^*, \nabla w},$$

где $K_{w^*, \nabla w}$ — ковариация случайных величин w^* и ∇w . Так как

$$-(D_{w^*} D_{\nabla w})^{1/2} \leq K_{w^*, \nabla w} \leq (D_{w^*} D_{\nabla w})^{1/2},$$

то

$$(\sqrt{D_{w^*}} - \sqrt{D_{\nabla w}})^2 \leq D_w \leq (\sqrt{D_{w^*}} + \sqrt{D_{\nabla w}})^2 \quad (\text{П.3})$$

Введем обозначение

$$K(s-j) = M \{ (w^2 [2(n-s)] - Mw^2 [2(n-s)]) \times \\ \times (w^2 [2(n-j)] - Mw^2 [2(n-j)]) \}.$$

Тогда

$$D_{\nabla w} = p_0^2 \sum_{s,j=0}^{\infty} p_s p_j (2K(s-j) - K(s-j-1) - K(s-j+1)) = \\ = 2p_0^2 \left\{ (K(0) - K(1)) \sum_{s=0}^{\infty} p_s^2 + \sum_{i=1}^{\infty} (2K(i) - K(i-1) - K(i+1)) \times \right. \\ \left. \times \sum_{s=0}^{\infty} p_s p_{s+i} \right\}. \quad (\text{П.4})$$

Подставляя многократно (19) в (20), получим

$$q = \sum_{s=0}^{\infty} p_s = \sum_{s=0}^{\infty} (\nu_0 + \nu_1)^s.$$

Отсюда следует, что ряд (20) сходится, если

$$|\nu_0 + \nu_1| < 1. \quad (\text{П.5})$$

С помощью (П.5) и (19) получим

$$\sum_{s=0}^{\infty} p_s p_{s+i} \geq \sum_{s=0}^{\infty} p_s p_{s+i+1} \quad (i = 0, 1, 2, \dots). \quad (\text{П.6})$$

Рассмотрим выражение (П.4). Если

$$2K(i) - K(i-1) - K(i+1) \geq 0,$$

то, используя (П.4), (П.6), найдем, что $D_{vw} < 0$. Так как всегда $D_{vw} \geq 0$, то $D_{vw} = 0$. Следовательно, $D_w = D_{w*}$. Это соотношение доказывает более общее положение, чем второе утверждение теоремы. Если

$$2K(i) - K(i-1) - K(i+1) < 0,$$

то из (П.4), (П.6) следует, что

$$0 \leq D_{vw} \leq 2\rho_0^2(K(0) - K(1)) \sum_{s=0}^{\infty} p_s^2. \quad (\text{П.7})$$

Если ПСВ (15) устойчивая, то

$$(K(0) - K(1)) \sum_{s=0}^{\infty} p_s^2 < \infty.$$

Так как

$$\rho_0^2 = a^2 k^2 (1 - e^{-\tau/2T})^2 e^{-\tau/T}, \quad (\text{П.8})$$

то $\rho_0^2 \rightarrow 0$, если $\tau/T \rightarrow \infty$ или $\tau/T \rightarrow 0$. Это доказывает второе утверждение теоремы. Теорема доказана.

2) Доказательство теоремы 2. Экстремальная система устойчива, если $m_w^2 < \infty$, $D_w < \infty$ при любых ограниченных по абсолютной величине входных возмущениях. При помощи (19) и (21) получим

$$\sum_{s=0}^{\infty} r_s^2 = \left(1 + c^4 - \frac{2v_0c^2}{1-v_1} \right) \sum_{s=0}^{\infty} p_s^2. \quad (\text{П.9})$$

Выражения (17), (18), (П.1) — (П.3), (П.7), (П.9) делают очевидными следующие утверждения: 1) экстремальная система устойчива в том и только в том случае, если сходится ряд (20); 2) из устойчивости ПСВ (16) следует устойчивость ПСВ (15) и наоборот. Эти утверждения доказывают теорему.

Л И Т Е Р А Т У Р А

1. А. А. Фельдбаум, Автоматика и телемеханика, 21, № 2, 167 (1960).
2. Т. И. Товстуха, Автоматика и телемеханика, 22, № 8, 1027 (1961).
3. А. А. Первозванский, Случайные процессы в нелинейных автоматических системах, Физматгиз, М., 1962.
4. Н. М. Александровский, Н. Е. Маджаров, Тр. МЭИ, вып. 59, 175 (1965).
5. А. М. Кориков, Автоматика и вычислительная техника, № 4, 43 (1970).

**DIFFERENCE EQUATIONS APPLICATION TO INVESTIGATION OF
STATIONARY MODE OF GRADIENT EXTREMAL SYSTEM WITH
INERTIAL PLANT***A. M. Korikov*

A gradient discrete extremal system with noise at the input and output of the plant is investigated with the aid of difference equations. Two types of inertial plants are considered. Theorems are proved enabling to investigate the system's stationary mode by means of a linear difference equation which is a linear part of the equation describing the extremum follow-up process. The influence of the plant's inertia upon the operation of the system is studied.

УДК 517.9

О КОНЕЧНЫХ РЕШЕНИЯХ СИСТЕМЫ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ПЕРИОДИЧЕСКИМИ КОЭФФИЦИЕНТАМИ

К. Г. Валеев, В. С. Яковенко

Даны необходимые и достаточные условия существования решения в виде конечного тригонометрического многочлена с экспоненциальным множителем для системы линейных дифференциальных уравнений с периодическими коэффициентами. Наиболее простые результаты получены для уравнений с простыми гармоническими коэффициентами. Для уравнения второго порядка существование конечных решений связано с вырождением областей устойчивости

1. Рассматривается система линейных дифференциальных уравнений

$$\sum_{s=0}^n \sum_{r=-\infty}^{\infty} A_{sr} e^{irt} \frac{d^s Y}{dt^s} = 0. \quad (1.1)$$

Здесь Y — m -мерный вектор. A_{sr} — $m \times m$ постоянные матрицы. Матричные ряды предполагаются сходящимися. Предполагаем, что

$$A_{n0} = E, \quad \sum_{r=-\infty, r \neq \infty}^{\infty} |A_{nr}| < 1, \quad (1.2)$$

что обеспечивает разрешимость относительно старших производных. E — единичная $m \times m$ -матрица. Предполагаем, что система (1.1) имеет решение вида

$$Y(t) = e^{\rho t} \sum_{k=-\beta}^{\alpha} Y_k e^{ikt}, \quad Y_\alpha \neq 0, \quad Y_{-\beta} \neq 0, \quad (1.3)$$

где α, β — некоторые числа. Подстановка ряда для $Y(t)$ в (1.1) приводит к системе линейных алгебраических уравнений

$$\sum_{k=-\beta}^{\alpha} L_{v-k}(p + ik) Y_k = 0 \quad (v = 0, \pm 1, \pm 2, \dots). \quad (1.4)$$

Здесь использовано обозначение для полиномиальных матриц

$$L_r(p) \equiv \sum_{s=0}^n A_{sr} p^s. \quad (1.5)$$

Очевидно, что существование ненулевого решения у системы (1.4) является необходимым и достаточным условием существования решения вида (1.3). Если в коэффициентах системы (1.1) имеется лишь конечное число гармоник, то в системе (1.4) будет конечное число уравнений.

2. Рассмотрим систему с простыми гармоническими коэффициентами

$$\sum_{s=0}^n (A_s + B_s e^{it} + C_s e^{-it}) \frac{d^s Y}{dt^s} = 0. \quad (2.1)$$

Уравнения (1.4) примут вид

$$L_1(p + i\alpha)Y_\alpha = 0, \quad L_{-1}(p - i\beta)Y_{-\beta} = 0, \quad (2.2)$$

$$\sum_{k=-\beta}^{\alpha} L_{v-k}(p + ik)Y_k = 0 \quad (v = -\beta, -\beta + 1, \dots, \alpha), \quad (2.3)$$

где используются обозначения

$$L_0(p) = \sum_{s=0}^n A_s p^s, \quad L_1(p) = \sum_{s=0}^n B_s p^s, \quad L_{-1}(p) = \sum_{s=0}^n C_s p^s, \quad (2.4)$$

$$L_r(p) \equiv 0 \quad (r = \pm 2, \pm 3, \pm 4, \dots).$$

Необходимым условием существования решения вида (1.3) является выполнение следующих уравнений

$$\det L_1(p + i\alpha) = 0, \quad \det L_{-1}(p - i\beta) = 0, \quad (2.5)$$

$$\det \begin{pmatrix} L_0(p + i\alpha) & L_1(p + i\alpha - i) & \dots & 0 \\ L_{-1}(p + i\alpha) & L_0(p + i\alpha - i) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & L_0(p - i\beta) \end{pmatrix} = 0. \quad (2.6)$$

Пример 1. Рассмотрим уравнение Маттье

$$\frac{d^2y}{dt^2} + (\lambda + 2\mu \cos t) y = 0.$$

При $\mu \neq 0$ условие (2.5) не выполнено, поэтому уравнение Маттье не имеет конечного решения вида (1.3).

Пример 2. Система дифференциальных уравнений

$$\frac{d^2y}{dt^2} + \omega_1^2 y_1 + a \cos t y_2 = 0, \quad \frac{d^2y_2}{dt^2} + \omega_2^2 y_2 + b \cos t y_1 = 0.$$

Условия (2.5) не выполнены, если $ab \neq 0$. При этом система уравнений не имеет конечного решения вида (1.3).

Пример 3. Рассматривается система уравнений

$$(E + 2A \cos t) \frac{d^2Y}{dt^2} + 2A \sin t \frac{dY}{dt} + EY = 0.$$

При $p = 0$ условия (2.5), (2.6) выполняются тождественно. Находим конечное решение системы

$$Y = Y_1 e^{it} + 2A(Y_1 + Y_{-1}) + Y_{-1} e^{-it}.$$

Векторы Y_1 , Y_{-1} произвольны.

При построении периодических решений периода 2π для системы с вещественными коэффициентами можно положить в системе (1.4) $p = 0$, $\alpha = -\beta$.

Пример 4. Для дифференциального уравнения

$$2\mu \frac{d^4y}{dt^4} + 2\mu \cos t \frac{d^2y}{dt^2} + (\lambda + 8\mu \cos t)y = 0$$

условие (2.5) при $p = 0$ дает $\alpha = -\beta = 2$. Уравнение (2.6) принимает вид

$$(\lambda + 32\mu)^2(\lambda + 2\mu)(\lambda - 4\mu)(\lambda + 6\mu) = 0.$$

Приравнивая каждый множитель нулю, получим периодические решения периода 2π

1. $\lambda = -32\mu$, $y_1 = \cos 2t$, $y_2 = \sin 2t$;
2. $\lambda = -2\mu$, $y_3 = 10 \sin t - \sin 2t$;
3. $\lambda = 4\mu$, $y_4 = 9 - 12 \cos t + \cos 2t$;
4. $\lambda = -6\mu$, $y_5 = 13 + 26 \cos t - 3 \cos 2t$.

3. Рассмотрим двупараметрическое дифференциальное уравнение второго порядка [1]

$$(1 + 2\mu a \cos t + 2\mu b \sin t) \frac{d^2y}{dt^2} + \mu(c + 2d \cos t + 2e \sin t) \times \\ \times \frac{dy}{dt} + (\lambda + 2\mu f \cos t + 2\mu g \sin t)y = 0 \quad (\mu > 0, c \geq 0). \quad (3.1)$$

Известно, [2], что на границе области неустойчивости существует периодическое или полупериодическое решение периода 2π . Необходимое условие (2.5) существования конечного периодического или полупериодического решения имеет при $p = 0,5\gamma i$ вид

$$4(f + gi) - 2(di - e)\gamma - (a + bi)\gamma^2 = 0 \quad (\gamma = 0, 1, 2, \dots). \quad (3.2)$$

При выполнении добавочного условия (2.6) при $p = 0,5\gamma i$, $\alpha = 0$, $\beta = -\gamma$ будет действительно существовать конечное решение. Получим теорему.

Теорема 1. Для того, чтобы уравнение (3.1) имело конечное периодическое или полупериодическое решение вида (1.3), необходимо, чтобы квадратное уравнение

$$4(f + gi) - 2(di - e)\gamma - (a + bi)\gamma^2 = 0$$

имело целый неотрицательный корень.

Сопоставляя эту теорему с теоремой (4.1) [1], приходим к интересной теореме.

Теорема 2. Если существует периодическое или полупериодическое решение периода 2π уравнения (3.1) вида

$$y(t) = \exp\left(\frac{\gamma_i}{2}t\right)(y_1 + e^{-it}y_{1-i} + \dots + e^{-it}y_0),$$

то на плоскости параметров λ , γ области неустойчивости с номерами $\gamma+2$, $\gamma+4$, $\gamma+6$, ... или не существуют, или вырождаются в линию. Границы областей неустойчивости с номерами γ , $\gamma-2$, ... определяются алгебраической кривой с уравнением $\gamma+1$ степени.

Пример 5. Для дифференциального уравнения

$$\frac{d^2y}{dt^2} - 2\mu \sin t \frac{dy}{dt} + (\lambda + 2\mu \cos t)y = 0$$

уравнение (3.2) имеет корень $\gamma = 2$. Уравнение границ нулевой и второй областей неустойчивости определяются уравнением

$$(\lambda - 1)(\lambda^2 - \lambda - 4\mu^2) = 0.$$

При $\lambda = 1$ имеется периодическое решение

$$y = \sin t.$$

Другое периодическое решение имеет вид

$$\lambda = 0,5 \pm \sqrt{0,25 + 4\mu^2}, \quad y = 4\mu + (-1 \pm \sqrt{1 + 16\mu^2}) \cos t.$$

В заключение отметим, что ряд уравнений общего вида (1.1) заменой могут быть иногда приведены к форме уравнений (2.1), что облегчает исследование. Для одного уравнения Хилла некоторые сходные результаты были получены в работе [3].

ЛИТЕРАТУРА

- 1 К. Г. Валеев, Изв. высш. уч. зав — Радиофизика, 5, № 4, 766 (1962).
- 2 М. Д. Стретт, О функциях Ляме, Матье и родственных им в физике и технике, ОНТИ, Харьков, 1955.
3. Л. Я Якупов, ДАН СССР, 181, № 2, 296 (1968).

Киевский институт инженеров
гражданской авиации

Поступила в редакцию
28 апреля 1970 г.

ON FINITE SOLUTIONS FOR A SYSTEM OF LINEAR DIFFERENTIAL EQUATIONS WITH PERIODIC COEFFICIENTS

K. G. Valeev, V. S. Yakovenko

Necessary and sufficient conditions for the existence of a linear differential system with periodic coefficients solution in the form of a finite trigonometric polynomial with an exponential multiplier are given. The simplest results are obtained for equations with simple harmonic coefficients. For a second order equation the existence of finite solutions is connected with degeneration of stability regions.

УДК 62 — 501

О СТОХАСТИЧЕСКОЙ УСТОЙЧИВОСТИ ЛИНЕЙНЫХ СИСТЕМ С НЕСТАЦИОНАРНО ФЛУКТУИРУЮЩИМИ ПАРАМЕТРАМИ

Г. Н. Бочков, О. В. Музичук

Рассматривается линейная нестационарная стохастическая система первого порядка и анализируются различные типы стохастической устойчивости ее движения. Исследуется система с флуктуациями параметров фликкерного типа. Вводится понятие временной стабильности системы.

1. Вопрос об устойчивости системы со случайно изменяющимися параметрами появился в литературе, по-видимому, с работы Розенблюма [1]. Розенблюм указал на «странный факт», что, хотя среднее значение и второй момент решения уравнения

$$\dot{X} + a(t) X = a(t) E(t), \quad X(0) = 0,$$

где $E(t)$ есть единичная функция, и $a(t) = a_0 + \alpha(t)$ — случайный стационарный нормально распределенный процесс с $\langle \alpha(t) \rangle = 0$, становится неограниченным при $t \rightarrow \infty$, решение приближается к единице по вероятности. Анализ показал, что к неустойчивости моментов ведет наличие «нулевых» компонент в спектре флуктуаций параметра. При условии, что спектральная плотность $S_\alpha(\Omega = 0) > 2\pi a_0$, система становится асимптотически неустойчивой в среднеквадратичном смысле.

К настоящему времени опубликовано уже немало работ, исследующих поведение стохастических систем — динамических систем, обладающих флуктуациями параметров (см., например, библиографию в [2, 3]). Во всех этих работах рассматриваются системы со стационарно флуктуирующими параметрами. Однако такая идеализация часто становится совершенно неприемлемой, например, при асимптотических оценках поведения реальных систем. Большинство систем, как это следует из многочисленных экспериментальных исследований, в действительности обладают нестационарно флуктуирующими параметрами «фликкерного» типа (с неинтегрируемыми спектрами мощности вблизи частоты $\Omega = 0$).

В настоящей работе исследуется простейшая линейная нестационарная стохастическая система первого порядка и анализируются статистические характеристики ее движения. Полученные результаты иллюстрируются примером усилителя с нестационарно флуктуирующим коэффициентом усиления, охваченного цепью инерционной положительной обратной связи.

2. Рассмотрим однородное линейное уравнение

$$\frac{dX}{dt} + a(t) X = 0 \quad (X(0) = X_0) \quad (1)$$

с флуктующим параметром

$$a(t) = a_0 + \alpha(t) \quad (a_0 = \langle a(t) \rangle),$$

где $\alpha(t)$ — случайный нормально распределенный процесс с нулевым средним значением $\langle \alpha \rangle = 0$.

Будем полагать, что флюктуации $\alpha(t)$ в общем случае нестационарны и обладают неинтегрируемым спектром мощности*. Для таких флюктуаций разработан аппарат структурных функций [4, 5], который используется в настоящей статье. Итак, считаем заданной структурную функцию флюктуаций параметра

$$d_\alpha[t, t + \theta; \tau] = \frac{1}{2} \langle [\alpha(t + \tau) - \alpha(t)] [\alpha(t + \theta + \tau) - \alpha(t + \theta)] \rangle.$$

Предположим также, что $\alpha(t)$ обладает стационарными привращениями. Для таких флюктуаций $d_\alpha[t, t + \theta; \tau] = d_\alpha[\theta; \tau]$; ей соответствует спектральная плотность $S_\alpha(\Omega)$.

Задачей настоящей работы является исследование статистических характеристик движения $X(t)$ системы (1). В качестве таких характеристик рассматриваются n -й момент $M_{xn}(t) \equiv \langle x^n(t) \rangle$ и вероятностное распределение $W_x(x; t)$.

3. В отсутствие флюктуаций параметра ($\alpha(t) \equiv 0$) «невозмущенное» движение системы

$$X^0(t) = X_0 \exp(-a_0 t). \quad (2)$$

«Возмущенное» движение $X(t)$, являющееся решением уравнения (1), равно

$$X(t) = X_0 \exp \left[- \int_0^t \alpha(u) du \right] = X^0(t) \exp \left[- \int_0^t \alpha(u) du \right]. \quad (3)$$

Вычислим n -й момент движения

$$M_{xn}(t) = X_0^n \exp(-na_0 t) \langle \exp[-n\lambda(t)] \rangle. \quad (4)$$

Здесь введено обозначение $\lambda(t) = \int_0^t \alpha(u) du$. Поскольку $\alpha(t)$ нормально, нормально распределенным будет и $\lambda(t)$ ($\langle \lambda(t) \rangle = 0$), поэтому

$$\langle \exp[-n\lambda(t)] \rangle = \exp \left[\frac{n^2}{2} \langle \lambda^2(t) \rangle \right] = \exp \{ n^2 d_\lambda[0, 0; t] \}, \quad (5)$$

$$M_{xn}(t) = X_0^n \exp(-na_0 t) \exp \{ n^2 d_\lambda[0, 0; t] \},$$

где $d_\lambda[t_1, t_2; \tau]$ — статистическая структурная функция процесса $\lambda(t)$. Из определения структурной функции $d_\lambda[t_1, t_2; \tau]$ можно получить связь d_λ со структурной функцией флюктуаций параметра $\alpha(t)$:

$$d_\lambda[0, 0; t] = \int_0^t u d_\alpha[0; u] du. \quad (6)$$

Таким образом, n -й момент движения $X(t)$ может быть записан как

$$M_{xn}(t) = X_0^n \exp(-na_0 t) \exp \left\{ n^2 \int_0^t u d_\alpha[0; u] du \right\}. \quad (7)$$

Если же флюктуации параметра $\alpha(t)$ таковы, что существует функция корреляции $\Phi_\alpha(\tau)$ почти для всех τ , то, записывая связь d_λ и Φ_α ,

* Отметим также, что в дальнейшем существенно используется условие «нулевой привязки» [6] нестационарно флюктуирующего параметра: $\alpha(t=0) = \alpha(0) = 0$. Такое предположение является вполне оправданным в целом ряде практических важных задач.

нетрудно получить в этом случае из (5) следующее выражение для момента $\langle X^n(t) \rangle$ (ср., например, с [6, 7])^{*}:

$$M_{xn}(t) = X_0^n \exp(-na_0 t) \exp\left[n^2 \int_0^t (t-u) \Phi_\alpha(u) du\right]. \quad (8)$$

4. Остановимся теперь на анализе полученных выражений (7) и (8). Введем в рассмотрение некоторые «характеристические показатели» поведения решения $X(t)$: $F_n^0 = (1/t) \ln \{[X^0(t)]^n\}$ — характеристический показатель «невозмущенного» движения, и $F_n^1 = (1/t) \times \ln \{\langle X^n(t) \rangle\}$ — характеристический (статистический) показатель «воздушенного» движения. Легко получить, что

$$F_n^0 = -na_0, \quad (9)$$

$$F_n^1 = F_n^0 + F_n^+, \quad (10)$$

где

$$F_n^+ = n^2 \int_0^t \left(1 - \frac{u}{t}\right) \Phi_\alpha(u) du, \quad (11)$$

если α — стационарный случайный процесс с заданной функцией $\Phi_\alpha(u)$, и

$$F_n^+ = n^2 \int_0^t \frac{u}{t} d_\alpha[0; u] du, \quad (12)$$

если α — случайный процесс со стационарными приращениями с заданной функцией $d_\alpha[0, u]$ и $\alpha(0) = 0$; F_n^+ назовем «воздушением» характеристического показателя F_n^0 .

Показатель $\chi_n^0 = \lim_{t \rightarrow \infty} F_n^0$ есть не что иное, как характеристический показатель Ляпунова [8, 9] для невоздушенной системы. Показатель $\chi_n^1 = \lim_{t \rightarrow \infty} F_n^1$ (характеризующий движение при $t \rightarrow \infty$) представляет собой обобщение показателя Ляпунова на случай стохастической системы.

Обращаясь к формулам (7) и (8) и замечая, что $F_n^+(t)$ — монотонно возрастающая функция времени**, можно понять, что характеристический показатель F_n^+ является «фактором расходности» движения $X(t)$ и, соответственно, F_n^0 — «фактором затухания» движения (в случае $a_0 > 0$). Этот вывод говорит о «дестабилизирующем» действии произвольных флюктуаций параметра системы.

Нетрудно показать также, что движение будет статистически p -устойчивым (асимптотически)*** $|M_{xn}(t \rightarrow \infty) \rightarrow 0|$ для всех $n \leq p$ лишь в случае $\chi_p^1 < 0$.

Вместе с тем представляет практический интерес рассмотрение не только асимптотического поведения решения, но и исследование харак-

* Отметим, что формулы (7) и (8) соответствуют двум качественно различным по постановке задачам (при выводе формулы (7) существенно использовано условие «нулевой привязки» флюктуаций параметра), поэтому не следует думать, что (8) может быть получено из (7).

** Так как $d_\alpha[0; u]$ всегда больше нуля

*** Определения различных типов стохастической устойчивости см., например, в [3].

тера движения на конечном временном интервале. Можно, например, поставить задачу отыскания временного интервала $[0 \leq t \leq T_0]$, в течение которого решение $X(t)$ является статистически ограниченным [7] некоторым вполне определенным значением $R | \langle X^n(t) \rangle < R^n |$, и, в частности, интервала T_0 , на котором $\langle X(t) \rangle$ не превзойдет начального отклонения X_0 ($\langle X^2(t) \rangle$ не превзойдет квадрата начального отклонения X_0^2 и т. д.). Величина T_0 характеризует среднее время пребывания системы в области, заданной начальным отклонением $[0, X_0]$. Совершенно ясно, что в общем случае T_0 может принимать значения, лежащие в пределах от нуля до бесконечности* (рис. 1).

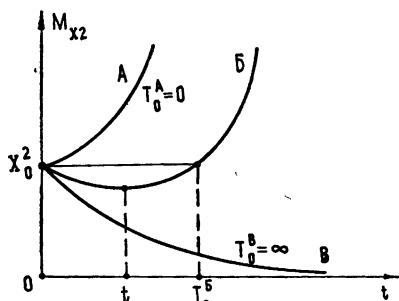


Рис. 1.

функций $M_{xn}(t)$ (представленных выражениями (7) и (8))**

$$\int_0^{t_0} \Phi_\alpha(u) du = a_0/n, \quad (13)$$

$$t_0 d_\alpha [0; t_0] = a_0/n. \quad (14)$$

Величина t_0 , которая вместе с T_0 характеризует движение, является важной характеристикой системы (1). Для всех $t < t_0$ движение $X(t)$ является статистически затухающим (система, которой сообщено начальное отклонение X_0 , ведет себя «квазистабильно»). При $t > t_0$ «фактор расходности» начинает преобладать над «фактором затухания» и моменты решения начинают экспоненциально расти. Временной интервал $[0, t_0]$ назовем областью временной стабильности системы.

Заметим, что из выражения (14) следует, что величина t_0 для $(n+1)$ -го момента движения всегда меньше, чем t_0 для n -го момента. Следовательно, область стабильности уменьшается с ростом момента.

Из соотношения (14) нетрудно понять также, что с «ростом расходности» среднего квадрата $\langle a^2(t) \rangle$ «область стабильности» системы (1) сокращается.

5. Перейдем теперь к исследованию одного практически важного случая рассмотренной задачи. Будем полагать, что флуктуации $a(t)$ носят фликкерный характер, т. е. спектральная плотность флукутаций параметра имеет вид

$$S_\alpha(\Omega) = \frac{A}{2\pi} |\Omega|^{-\gamma} \quad (0 < \gamma < 3). \quad (15)$$

Анализируя поведение моментов движения в этом случае, целесообразно выделить два качественно различных интервала значений показателя спектра γ :

* Последний случай соответствует статистической (асимптотической) устойчивости движения $X(t)$.

** Легко показать, что значения t_0 заведомо существуют и действительно дают минимум функции $M_{xn}(t)$, если $d_\alpha [0; t] = (1/2) \langle a^2(t) \rangle$ — монотонно возрастающая функция, а $\Phi_\alpha(t)$ всюду положительна.

1) $0 < \gamma < 1$ — область стационарных флюктуаций параметра $a(t)$, где почти для всех τ существует функция корреляции флюктуаций параметра [5]:

$$\Phi_\alpha(\tau) = B_1(\gamma) |\tau|^{\gamma-1}, \quad B_1(\gamma) = \frac{A}{\pi} \Gamma(1-\gamma) \sin \frac{\pi\gamma}{2}; \quad (16)$$

2) $1 < \gamma < 3$ — область «существенно нестационарных» флюктуаций параметра, где функция $\Phi_\alpha(\tau)$ не существует, но существует структурная функция $d_\alpha[0; \tau]$:

$$d_\alpha[0; \tau] = B_2(\gamma) |\tau|^{\gamma-1}, \quad B_2(\gamma) = \frac{A}{2\Gamma(\gamma) \sin(\pi/2)(\gamma-1)}. \quad (17)$$

В первом случае из выражений (8) и (16) можно получить

$$M_{xn}(t) = X_0^n \exp(-na_0 t) \exp\left[\frac{n^2 B_1(\gamma)}{\gamma(\gamma+1)} t^{\gamma+1}\right]. \quad (18)$$

Для границы области временной стабильности в данном случае будем иметь

$$t_0 = \left[\frac{\gamma a_0}{n B_1(\gamma)} \right]^{1/\gamma}. \quad (19)$$

Интересно осуществить предельный переход в выражении (19) при $\gamma \rightarrow 0$ к случаю «белых» флюктуаций параметра со спектральной плотностью

$$S_\alpha(\Omega) = A/2\pi = S_0 = \text{const} \quad (20)$$

или, иначе говоря, к случаю стационарной стохастической системы, столь подробно исследованной в литературе.

Выполняя вычисления, получим

$$\lim_{\gamma \rightarrow 0} t_0(a_0, n, \gamma) = \lim_{\gamma \rightarrow 0} \left[\frac{2a_0}{nA} \right]^{1/\gamma} = \begin{cases} \infty & (a_0 > nA/2) \\ 0 & (a_0 < nA/2) \end{cases}. \quad (21)$$

Значение $t_0 = \infty$ соответствует асимптотически устойчивому движению, а это имеет место при условии (21):

$$a_0 > nA/2 \quad \text{или} \quad a_0 > \pi n S_\alpha(0). \quad (22)$$

В частности, если $n = 2$, тогда имеем $a_0 > 2\pi S_\alpha(0)$. Получили хорошо известное условие [1-7] среднеквадратичной (статистической) устойчивости движения системы (1).

Во втором случае ($1 < \gamma < 3$) из (7), (14) и (17) получим

$$M_{xn}(t) = X_0^n \exp(-na_0 t) \exp\left[\frac{n^2 B_2(\gamma)}{\gamma+1} t^{\gamma+1}\right], \quad (23)$$

$$t_0 = \left[\frac{a_0}{n B_2(\gamma)} \right]^{1/\gamma}. \quad (24)$$

Таким образом, в данном случае при любом A , a_0 и γ движение системы (1) при $t \rightarrow \infty$ является статистически неограниченным. (Характеристический показатель движения $\chi^1(t) = \lim_{t \rightarrow \infty} F^1(t) = +\infty$.) Но

в то же время на временном интервале $0 \leq t < t_0 = \sqrt[n]{\frac{a_0}{n} B_2(\gamma)}$ момент $M_{xn}(t)$ является затухающей функцией, достигающей при $t = t_0$ минимума, равного

$$\langle X^n(t) \rangle_{\min} = X_0^n \exp \left(-\frac{\gamma}{\gamma + 1} n a_0 t_0 \right). \quad (25)$$

В частности, при $\gamma = 2 \langle X^n(t) \rangle_{\min} = X_0^n \exp \left(-\frac{2}{3} a_0 \sqrt{\frac{n a_0}{A}} \right)$. Поведение $M_{xn}(t)$ в этом случае ($\gamma=2$) показано на рис. 2 для различных значений a_0 и в предположении $a_0/A = \text{const}$.

Ограничивааясь рассмотрением среднего квадрата решения, нетрудно получить также, что на временном интервале $0 < t < T_0$, где

$$T_0 = \sqrt{\gamma + 1} t_0, \quad (26)$$

решение остается статистически ограниченным, причем на указанном интервале $\langle X^2(t) \rangle$ не превосходит квадрата начального отклонения X_0^2 .

На рис. 3 показано «движение» границы области временной стабильности системы t_0 при изменении показателя спектра γ^* .

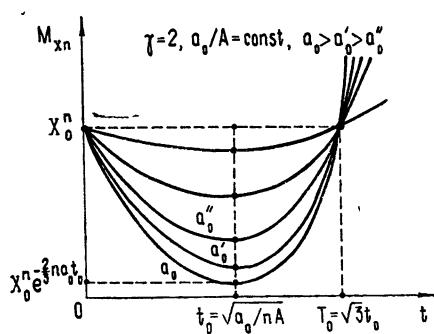


Рис. 2.

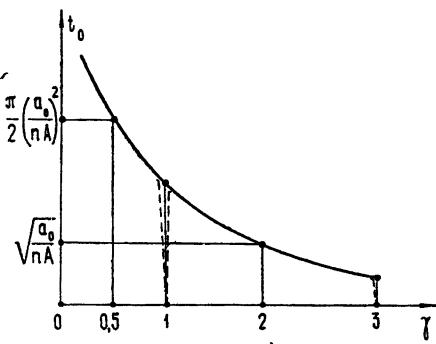


Рис. 3.

6. Перейдем к анализу вероятностного (условного) распределения $W_x(x, t | x_0, t_0)$ ($X(t_0) = X_0 > 0$ — начальное состояние системы) и, в частности, исследуем вероятность превышения движения $X(t)$ определенного значения X^*

$$P\{X(t) > x^*\} = \int_{x^*}^{\infty} W_x(x, t | x_0, t_0) dx. \quad (27)$$

Интегрируя (1) и вводя обозначение $Z(t) = -[a_0 t + \lambda(t)]$, будем иметь

$$\ln \frac{X(t)}{X_0} = Z(t). \quad (28)$$

* На рис. 3 вблизи точек $\gamma=1$ и $\gamma=3$ пунктиром показана кривая, полученная при непосредственных вычислениях по формулам (19) и (24). Особенность в точке $\gamma=1$ связана с расходимостью $\langle \alpha^2(t) \rangle$ из-за неинтегрируемости спектра $S_\alpha(\Omega) \sim 1/\Omega$ при $\Omega \rightarrow \infty$. Если учесть, что любой реальный спектр всегда спадает быстрее, чем $1/\Omega$, и ввести при $\gamma=1$ спектральную плотность соотношением

$$S_\alpha(\Omega) = \frac{A\Omega_B}{2\pi |\Omega| (\Omega_B + |\Omega|)} \quad (S_\alpha(\Omega) \sim 1/\Omega^2 \text{ при } \Omega \gg \Omega_B)$$

(как это сделано, например, в [5], § 17), то в целом получим кривую, показанную на рис. 3 сплошной линией. Случай $\gamma=3$ также является «особенным», так как при этом «перестает работать» используемый нами аппарат структурных функций (флуктуации параметра уже не обладают стационарными приращениями).

Замечая, что $Z(t)$ (как и $\lambda(t)$) имеет нормальное распределение

$$W_z(z, t) = \frac{1}{\sqrt{2\pi \langle \lambda^2(t) \rangle}} \exp \left[-\frac{(z + a_0 t)^2}{2 \langle \lambda^2(t) \rangle} \right], \quad (29)$$

нетрудно получить из (28) и (29), что

$$W_x(x, t | x_0, t_0) = \begin{cases} \frac{1}{X \sqrt{2\pi \langle \lambda^2(t) \rangle}} \exp \left[-\frac{\left(\ln \frac{x}{x_0} + a_0 t \right)^2}{2 \langle \lambda^2(t) \rangle} \right] & (x > 0) \\ 0 & (x < 0) \end{cases} \quad (30)$$

При вычислении (27) необходимо различать два случая в зависимости от соотношения X^* и X_0 : а) $X^* \geq X_0$ и б) $X^* < X_0$.

В случае а) интегрирование (27) с учетом (30) дает*

$$P\{x(t) > x^*\} = \frac{1}{2} \left\{ 1 - F \left[\frac{1}{2} \left(\ln \frac{x^*}{x_0} + a_0 t \right) / \sqrt{d_\lambda[0, 0; t]} \right] \right\}; \quad (31)$$

для случая «фликкерных» флюктуаций со спектром $S_\alpha(\Omega) = (A/2\pi) |\Omega|^{-\gamma}$ из (16), (17) и (31) получим, что

$$P\{x(t) > x^*\} = \left\{ \begin{array}{l} \frac{1}{2} \left\{ 1 - F \left[\frac{1}{2} \left| \ln \frac{x^*}{x_0} + a_0 t \right| / \sqrt{\frac{B_1(\gamma)}{\gamma(\gamma+1)} t^{\gamma+1}} \right] \right\} (\gamma < 1) \\ \frac{1}{2} \left\{ 1 - F \left[\frac{1}{2} \left| \ln \frac{x^*}{x_0} + a_0 t \right| / \sqrt{\frac{B_2(\gamma)}{\gamma+1} t^{\gamma+1}} \right] \right\} (\gamma > 1) \end{array} \right. \quad (32)$$

$$P\{x(t) > x^*\} = \left\{ \begin{array}{l} \frac{1}{2} \left\{ 1 - F \left[\frac{1}{2} \left| \ln \frac{x^*}{x_0} + a_0 t \right| / \sqrt{\frac{B_1(\gamma)}{\gamma(\gamma+1)} t^{\gamma+1}} \right] \right\} (\gamma < 1) \\ \frac{1}{2} \left\{ 1 - F \left[\frac{1}{2} \left| \ln \frac{x^*}{x_0} + a_0 t \right| / \sqrt{\frac{B_2(\gamma)}{\gamma+1} t^{\gamma+1}} \right] \right\} (\gamma > 1) \end{array} \right. \quad (33)$$

Интересно заметить, что при $t \rightarrow 0$ $P\{x > x^*\} \rightarrow 0$ (рис. 4 а, в) и что

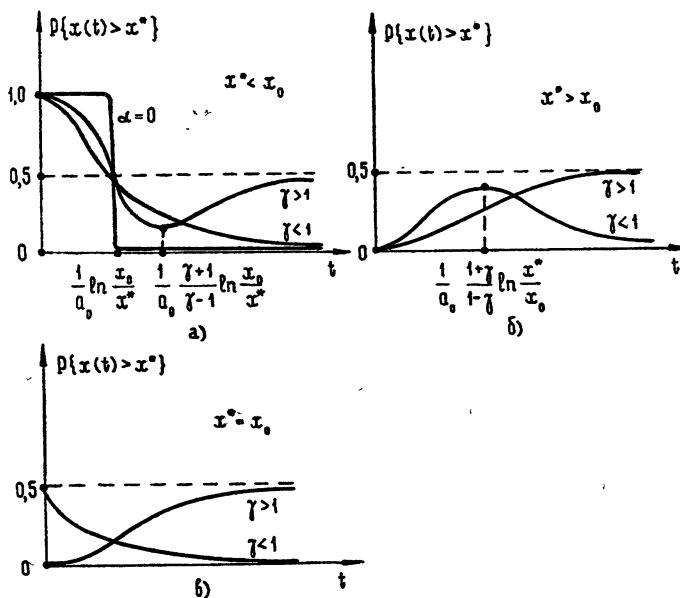


Рис. 4.

* Здесь $F(u) = \frac{2}{\sqrt{\pi}} \int_0^u \exp(-v^2) dv$ — «интеграл вероятности» ($F(0)=0$, $F(\infty)=1$).

при $t \rightarrow \infty$

$$P\{x(t) > x^*\} = \begin{cases} 0 & (\gamma < 1) \\ 1/2 & (\gamma > 1) \end{cases}$$

В случае б) ($X^* < X_0$) результат зависит от того, превышает или не превышает t значение $t^* = \frac{1}{a_0} \ln \frac{x^*}{x_0}$.

Если $t < t^*$, то

$$P\{x(t) > x^*\} = \begin{cases} \frac{1}{2} \{1 + F[Q_1]\} & (\gamma < 1) \\ \frac{1}{2} \{1 + F[Q_2]\} & (\gamma > 1) \end{cases}, \quad (34)$$

(где Q_1 и Q_2 совпадают с соответствующими выражениями в квадратных скобках формул (32) и (33)).

Если $t > t^*$, то вычисления $P\{x(t) > x^*\}$ в случае б) также приводят к формулам (32) и (33) и, следовательно, при $t \rightarrow \infty$ по-прежнему

$$P\{x(t) > x^*\} \rightarrow \begin{cases} 0 & (\gamma < 1) \\ 1/2 & (\gamma > 1) \end{cases}$$

и, конечно, при $t \rightarrow 0$ $P\{x > x^*\} \rightarrow 1$ (рис. 4, б).

Отметим также, что из полученной формулы (32) предельным переходом при $\gamma \rightarrow 0$ $|S_a(\Omega) \rightarrow S_0 = A/2\pi|$ легко получить результаты, справедливые для случая стационарных «белых» флюктуаций параметра a . Эти результаты оказываются в точности совпадающими с соответствующими формулами в [6], полученными способом, отличным от нашего.

Таким образом, показано, что при $t \rightarrow \infty$ вероятность превышения любого определенного значения стремится к нулю при $\gamma < 1$ и приближается к 1/2 при $\gamma > 1$. Этот вывод говорит о двух качественно различных ситуациях в движении системы в зависимости от величины показателя спектра флюктуаций γ . Несмотря на то, что моменты движения в обоих случаях со временем стремятся к бесконечности, вероятность обнаружения «большого перемещения» $X(t)$ в любом члене ансамбля систем в случае $\gamma < 1$ пренебрежимо мала. Условие $\gamma < 1$ является условием асимптотической устойчивости по вероятности ($P\{|X(x_0; t \rightarrow \infty)| > X^*\} \rightarrow 0$). В случае «существенно нестационарных» флюктуаций параметра ($\gamma > 1$) расходимость моментов движения (q — неустойчивость) сочетается с асимптотической неустойчивостью по вероятности).

7. Рассмотрим в качестве примера поведение «широкополосного» усилителя постоянного тока с нестационарно флюктуирующим коэффициентом усиления, охваченного цепью положительной обратной связи, представляющей собой инерционную RC цепь (НЧ фильтр)*. Структурная схема такой системы показана на рис. 5. Будем полагать флюктуации усиления $K(t) = K_0[1 + \delta k(t)]$ фликкерными со спектром $S_{\delta k}(\Omega) = (A/2\pi) |\Omega|^{-\gamma}$ и нормально распределенными ($\langle \delta k \rangle = 0$).

В отсутствие флюктуаций усиления ($\delta k(t) \equiv 0$) система в целом обладает следующим коэффициентом передачи (рис. 6):

$$H_0 = \frac{Y(p)}{E(p)} = \frac{(1 + j\Omega T) K_0}{(1 - K_0) + j\Omega T}; \quad |H_0|^2 = \frac{K_0^2}{(1 - K_0)^2} \frac{1 + \Omega^2 T^2}{1 + \Omega^2 T^2 / (1 - K_0)^2}, \quad (35)$$

* Параметры цепи обратной связи для простоты рассмотрения считаем нефлюктуирующими.

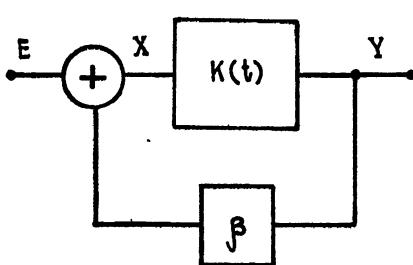


Рис. 5.

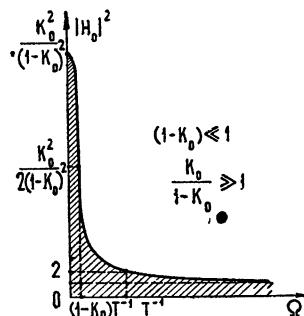


Рис. 6.

где T — постоянная времени цепи обратной связи. Устойчивость системы (в отсутствие флуктуаций) определяется условием $K_0 < 1$.

При флуктуациях коэффициента усиления (и в отсутствие входного сигнала $E(t) \equiv 0$) поведение системы будет описываться следующим уравнением (X — напряжение на выходе усилителя):

$$\frac{dX}{dt} + \frac{1}{T} [1 - K_0 - \Delta k(t)] X = 0. \quad (36)$$

Сравнивая (36) и (1), имеем

$$\alpha_0 = \frac{1}{T} (1 - K_0), \quad \alpha(t) = -\frac{1}{T} \Delta k(t) = -\frac{K_0}{T} \delta k(t). \quad (37)$$

На основании проведенных выше расчетов не представляет труда проанализировать поведение данной системы при различных по характеру флуктуациях коэффициента усиления. Приведем лишь некоторые, представляющие практический интерес, количественные оценки. Для определенности зафиксируем постоянную времени цепи обратной связи, полагая $T = 10^{-3}$ сек. Начнем исследование со случая, когда усилитель постоянного тока обладает сравнительно малыми относительными флуктуациями коэффициента усиления $S_{\delta k} = \frac{10^{-10}}{\Omega}$ (гц^{-1})*. В этом случае из (19), (24) и (37), задавшись коэффициентом усиления K_0 , полосой системы $\Pi = \alpha_0 = \frac{1}{T} (1 - K_0)$ или величиной усиления системы по постоянному току $H_0(0) = K_0/(1 - K_0)$, нетрудно найти для «интервала среднеквадратичной стабильности» рассматриваемой системы оценки, приведенные в табл. 1 (где для сравнения приведены расчеты и в случае существенно нестационарных флуктуаций коэффициента усиления с показателем спектра $\gamma = 2$). Следует отметить, что все расчеты, приведенные в табл. 1, соответствуют усилителям постоянного тока с «достаточно стабильными» коэффициентами усиления. Тем не менее в последнем случае область стабильности составляет всего ~ 20 мин. При снижении требований на стабильность коэффициента усиления K резко сокращается и «область стабильности» рассматриваемой системы t_0 . Так, при $S_{\delta k} = \frac{10^{-5}}{2\pi\Omega^2}$ (гц^{-1}) и $\Pi_1 = 1$ гц, $\Pi_2 = 10^{-3}$ гц будем иметь со-

* Здесь и в дальнейшем оценки спектров флуктуаций усиления выбраны на основании данных по флуктуациям параметров элементов радиосхем из [6].

ответственно $t_{01} \simeq 2 \cdot 10^3$ сек и $t_{02} \simeq 10$ (сек), т. е. уже весьма малую область стабильности, что свидетельствует в данном случае о невозможности длительного практического использования подобной системы, например, для целей НЧ фильтрации.

Таблица 1

T (сек)	$H_0 = \frac{K_0}{1 - K_0}$	$\Pi = a_0 (2\pi)$	$A (2\pi)^{-1}$	γ	t_0 (сек)
10^{-3}	10^3	1	10^{-10}	1	10^9
	10^6	10^{-3}		1	$2 \cdot 10^6$
	10^3	1		2	$7 \cdot 10^4$
	10^6	10^{-3}		2	$2 \cdot 10^3$

Рассмотренная схема (рис. 5) встречается весьма часто при аналоговом моделировании различных систем*. В любом реальном усилителе с высоким коэффициентом усиления также всегда есть возможность для возникновения «паразитной» положительной обратной связи, и поэтому анализ такого усилителя по-прежнему сводится к рассмотренной структурной схеме.

Следует иметь в виду, что такая модель реальной системы (нормальность плюс нестационарный характер флуктуаций параметров) всегда является стохастически неустойчивой. Но тем не менее, как показывают только что приведенные оценки, практически такой «усилитель» можно считать устойчивым для весьма большого интервала времени при соответствующих требованиях на стабильность блоков усиления. Здесь совершенно естественно говорить о долговременной стабильности системы. Поэтому введенное нами ранее понятие временной стабильности в отдельных случаях может трактоваться как долговременная или же кратковременная стабильность в зависимости от характера флуктуаций параметров конкретной системы.

Авторы благодарны А. Н. Малахову за постановку задачи, обсуждение работы и ценные замечания.

ЛИТЕРАТУРА

1. A. Rosenblom, Analysis of Randomly Time-Varying Linear Systems, Ph. D. Thesis, Univ. of California, Los Angeles, 1954.
2. H. Kushner, Stochastic Stability and Control Academic Press., N. Y.—London 1967.
3. Р. З. Хасьминский, Устойчивость систем дифференциальных уравнений при случайных возмущениях их параметров, изд. Наука, М., 1969.
4. А. Н. Малахов, Изв. высш. уч. зав.—Радиофизика, 9, № 3, 595 (1966).
5. А. Н. Малахов, Флуктуации в автоколебательных системах, изд. Наука, М., 1968.
6. T. K. Saughey, J. K. Dienes, J. Math. and Phys., 41, 300 (1962).
7. А. Н. Малахов, Изв. высш. уч. зав.—Радиофизика, 6, № 1, 42 (1963).
8. А. М. Ляпунов, Общая задача об устойчивости движения, Гостехиздат, М., 1960.
9. Б. Ф. Былов, Р. Э. Виноград, Д. М. Гробман, В. В. Немецкий, Теория показателей Ляпунова, изд. Наука, М., 1966.

Горьковский государственный
университет

Поступила в редакцию
24 апреля 1970 г.

* К исследованию почти аналогичных эквивалентных схем сводится анализ «операционных усилителей» (интеграторов, дифференцирующих усилителей и т. д.) с учетом нестабильности блоков усиления.

ON STOCHASTIC STABILITY OF LINEAR SYSTEMS WITH
NONSTATIONARILY FLUCTUATING PARAMETERS

G. N. Bochkov, O. V. Musychuk

A linear nonstationary stochastic system of the first order is considered and various types of its motion stochastic stability are analyzed. A system with flicker-type fluctuations of parameters is studied. The concept of temporary stability of a system is introduced.

УДК 531.391

НЕКОТОРЫЕ ВОПРОСЫ ВЗАИМОДЕЙСТВИЯ КВАЗИГАРМОНИЧЕСКИХ КОЛЕБАНИЙ В СЛАБОНЕЛИНЕЙНЫХ СИСТЕМАХ

Л. В. Постников

Рассматривается автономная динамическая система, описываемая дифференциальным уравнением N -го порядка с одной нелинейностью. Формулируются условия, при которых в системе возможны колебания, состоящие из суммы квазигармонических составляющих. Написана система укороченных уравнений. Показано, что при аппроксимации нелинейности кубическим полиномом и отсутствии синхронного взаимодействия в автоколебательной системе возможен только одночастотный режим.

Квазигармонические колебания в слаболинейных системах уже давно привлекают к себе внимание специалистов. Математическое исследование таких явлений, как автоколебания в ламповом генераторе [1], явление захватывания [2], принудительная синхронизация на кратных частотах [3], явление затягивания [4], взаимная синхронизация генераторов [5] и т. д. знаменовало собой рождение нелинейной теории колебаний.

С развитием метода точечных преобразований и качественной теории дифференциальных уравнений интерес к квазигармоническим колебаниям несколько ослаб. В шестидесятые годы эти колебания вновь стали подвергаться детальному изучению. Для этого периода характерен интерес к прикладным задачам [6–9]. Большинство работ посвящено исследованию неавтономных систем с одной степенью свободы. Есть работы и по исследованию дискретных систем с большим числом степеней свободы, систем с запаздыванием в обратной связи [10–13], однако в этом случае рассмотрение ведется в одночастотном приближении.

Развитие нелинейной электродинамики, нелинейной оптики, физики плазмы поставило перед теорией ряд новых задач по исследованию квазигармонических колебаний и волн [21–23]. Решение их может считаться началом нелинейной теории колебаний распределенных систем [14–16]. В этих системах оказались актуальными задачи, связанные с вопросами многочастотного взаимодействия.

Объектом исследования данной работы служат многочастотные квазигармонические колебания в дискретных системах*.

1. УСЛОВИЯ КВАЗИГАРМОНИЧНОСТИ. СИСТЕМА УКОРОЧЕННЫХ УРАВНЕНИЙ

Рассмотрим динамическую систему, описываемую уравнением

$$a(p)x = -\mu b(p)f(x), \quad (1)$$

где $p = d/dt$, $\mu \ll 1$, $a(p)$ и $b(p)$ — некоторые функции оператора p , не имеющие в пространстве изображений полюсов при конечных значениях p ; $f(x)$ — нелинейная функция, разлагаемая в степенной ряд, начи-

* В данном случае система считается дискретной, если она не содержит распределенных нелинейностей.

нающийся членом, содержащим x^2 . К уравнению (1) может привести, например, исследование динамики системы, показанной на рис. 1. Здесь

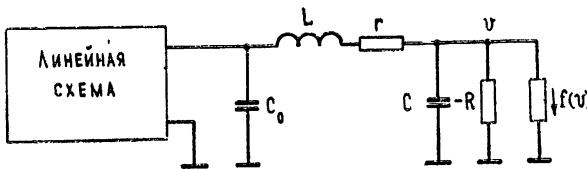


Рис. 1.

прямоугольником обозначена произвольная линейная схема; C_0 , L , r , C и R — соответственно монтажная емкость, индуктивность выводов, сопротивление растекания, емкость перехода и абсолютная величина дифференциального сопротивления туннельного диода; $f(v)$ — ток через его нелинейную часть. В этом случае величина $b(p)/a(p)$ является входной операторной проводимостью всей линейной части системы, включающей и линейные элементы туннельного диода.

Элементарные выкладки показывают, что

$$\frac{b(p)}{a(p)} = \frac{R [(Lp + z)(z(p)C_0p + 1) + z(p)]}{(RCp - 1)[(Lp + r)(z(p)C_0p + 1) + z(p)] + R(z(p)C_0p + 1)},$$

где $z(p)$ — входной импеданс линейной схемы. Если линейная схема состоит из дискретных элементов, то ее входной импеданс является дробно-рациональной функцией оператора p , если из распределенных — трансцендентной функцией p .

Пусть уравнение (1) имеет единственное состояние равновесия $x = 0$, характеристическое уравнение

$$a(p) = 0 \quad (2)$$

не имеет кратных или отличающихся друг от друга на величину порядка μ корней, а

$$\lim_{p \rightarrow \infty} \frac{b(p)}{a(p)} = 0.$$

Следуя работе [17], запишем уравнение (1) в виде

$$x = -\mu \frac{b(p)}{a(p)} f(x) \quad (3)$$

и разложим его правую часть на простейшие дроби. Обозначая каждую из простейших дробей через x_v и учитывая, что $p = d/dt$, получим систему уравнений, адекватную, при оговоренных условиях, уравнению (1):

$$\dot{x}_v - p_v x_v = -\mu B_v f(x), \quad (4)$$

$$x = \sum_v x_v,$$

где $B_v = b(p_v)/a'(p_v)$, а p_v — корень характеристического уравнения (2).

Эти уравнения написаны в комплексных координатах и, в отличие от линейной консервативной системы [18], отражают взаимодействие нормальных колебаний.

Пусть N пар комплексно-сопряженных корней характеристического уравнения (2) имеют действительную часть порядка μ^* :

$$\rho_v = \mu\sigma_v + j\Omega_v, \quad (4a)$$

а остальные — большую по величине отрицательную. В этом случае, согласно работе [20], x_v будут квазигармоническими комплексными функциями времени, и решение системы (4) можно искать в виде

$$x_v = z_v \exp(j\omega_v t), \quad (5)$$

где

$$\omega_v = \Omega_v - \mu\xi_v, \quad (5a)$$

— собственные частоты колебаний нелинейной системы (Ω_v — собственная частота ее линейного приближения), а z_v — комплексная амплитуда, являющаяся функцией аргумента $j\mu t$.

Подставляя решение (5) в систему уравнений (4) и приравнивая коэффициенты при $\exp j\omega_v t$, получим систему укороченных уравнений

$$\dot{z}_v = (\sigma_v + j\xi_v) z_v - B_v x_v, \quad (6) \\ (v = 1, 2, \dots, N),$$

где x_v — комплексная амплитуда гармонической составляющей частоты ω_v для функции $f(x)$.

Раскладывая эту функцию в степенной ряд по x и подставляя

$$x = \sum_k [z_k \exp(j\omega_k t) + z_k^* \exp(-j\omega_k t)], \quad (6a)$$

после несложных преобразований получим разложение функции $f(x)$ на гармонические составляющие

$$f(x) = \sum_{k=1}^{\infty} \frac{C_{\sum(n_k + \bar{n}_k)} \prod_k (z_k^{\bar{n}_k} z_k^{n_k}) \exp[j \sum_k (\bar{n}_k - n_k) t]}{\prod_k \bar{n}_k! \prod_k n_k!} \quad (7) \\ (k = 1, 2, \dots, N),$$

где $C_{\sum(n_k + \bar{n}_k)}$ — производная от $f(x)$ при $x = 0$ порядка $\sum_k (n_k + \bar{n}_k)$.

Гармоническая составляющая частоты ω_k соответствует значениям n_k и \bar{n}_k удовлетворяющим соотношению

$$\sum_k \omega_k (\bar{n}_k - n_k - \delta_{vk}) = 0, \quad (8)$$

где δ_{vk} — символ Кронекера.

Пусть собственные частоты ω_k , определяемые с точностью до величин порядка μ характеристическим уравнением (2), удовлетворяют соотношению

$$\sum_{k_1} \alpha_{k_1} \omega_{k_1} - \sum_{k_2} \beta_{k_2} \omega_{k_2} = 0, \quad (8a)$$

где α_{k_1} и β_{k_2} — целые неотрицательные числа, не имеющие общего множителя, отличного от единицы, а значения k_1 и $k_2 \neq k_1$ являются элементами множества $k = 1, 2, \dots, N$.

* Если число корней с малой действительной частью нечетно, то x будет содержать еще медленно меняющуюся «постоянную» составляющую. Взаимодействие ее с квазигармоническими колебаниями может привести к автомодуляции [19].

Тогда, согласно уравнениям (8) и (8 а),

$$\begin{aligned}\bar{n}_{k_1} &= n_{k_1} + \delta_{v k_1} + m\alpha_{k_1}, \\ \bar{n}_{k_2} &= n_{k_2} + \delta_{v k_2} - m\beta_{k_2},\end{aligned}\quad (8б)$$

где m — любое целое число.

Подставляя выражение (8 б) в разложение (7), после несложных преобразований индексов, проведенных для того, чтобы учесть, что \bar{n}_{k_1} и $n_{k_2} \geq 0$, получим выражение для комплексной амплитуды x :

$$x_v = z_v [F_{0v} + \sum_{m=1}^{\infty} \bar{F}_{mv} + \sum_{m=1}^{\infty} \tilde{F}_{mv}], \quad (9)$$

где

$$F_{0v} = \sum_{n_k=0}^{\infty} \frac{C_{1+\sum 2n_k} \prod_{k} (z_{k_1} z_{k_2}^*)^{n_k}}{\prod_{k} n_k! \prod_{k} (n_k + \delta_{v k})!}; \quad (9а)$$

$$\begin{aligned}\bar{F}_{mv} &= \prod_{k_1 k_2} (z_{k_1}^{m\alpha_{k_1}} z_{k_2}^{*\beta_{k_2}}) \times \\ &\times \sum_{n_{k_1} n_{k_2}} \frac{C_{\gamma_v} \prod_{k_1 k_2} [(z_{k_1} z_{k_1}^*)^{n_{k_1}} (z_{k_2} z_{k_2}^*)^{n_{k_2} - \delta_{v k_2}}]}{\prod_{k_1} [n_{k_1}! n_{k_1}]! (n_{k_1} + m\alpha_{k_1} + \delta_{v k_1})! (n_{k_2} + m\beta_{k_2} - \delta_{v k_2})!}, \\ \tilde{F}_{mv} &= \prod_{k_1 k_2} (z_{k_1}^{m\alpha_{k_1}} z_{k_2}^{m\beta_{k_2}}) \times \\ &\times \sum_{n_{k_1} n_{k_2}} \frac{C_{\tilde{\gamma}_v} \prod_{k_1 k_2} [(z_{k_1} z_{k_1}^*)^{n_{k_1} - \delta_{v k_1}} (z_{k_2} z_{k_2}^*)^{n_{k_2}}]}{\prod_{k_1} [n_{k_1}! n_{k_1}]! (n_{k_1} + m\alpha_{k_1} - \delta_{v k_1})! (n_{k_2} + m\beta_{k_2} + \delta_{v k_2})!},\end{aligned}\quad (9б)$$

$$\begin{aligned}\tilde{\gamma}_v &= 1 + \sum_{k_1} (2n_{k_1} + m\alpha_{k_1}) + \sum_{k_2} (2n_{k_2} + m\beta_{k_2} - 2\delta_{v k_2}), \\ \tilde{\gamma}_v &= 1 + \sum_{k_1} (2n_{k_1} + m\alpha_{k_1} - 2\delta_{v k_1}) + \sum_{k_2} (2n_{k_2} + m\beta_{k_2}).\end{aligned}\quad (9в)$$

В выражениях (9 б) подразумевается еще суммирование по всем $\alpha_{k_i} \geq 0$ и $\beta_{k_2} > 0$, удовлетворяющим уравнению (8 а), если они определяются этим уравнением неоднозначно.

Следует учесть, что выражения (9 б) написаны для случая, когда все α_{k_1} и β_{k_2} — целые положительные, отличные от нуля, величины. Если какое-либо из них равно нулю, то суммирование по соответствующему n_{k_1} в выражении для \tilde{F}_{mv} нужно начинать не с нуля, а с $\delta_{v k_1}$. Точно так же при $\beta_{k_2} = 0$ суммирование по n_{k_2} в выражении для \tilde{F}_{mv} нужно начинать с $\delta_{v k_2}$.

Система уравнений (6) с учетом выражений (9), (9а), (9б) и (9в), написанных для произвольного случая, является слишком сложной для дальнейшего исследования, однако, опираясь на нее, можно сделать некоторые общие выводы.

Например, если для собственных частот ω_k выполняется только одно условие типа (8 а), в котором не все α_{k_1} и β_{k_2} равны нулю, то система (9) имеет порядок $N + 1$. Чтобы в этом убедиться, обратим внимание

на то, что F_0 , зависит только от квадратов амплитуд колебаний и является действительной величиной. Это член определяет несинхронное взаимодействие колебаний [14]; \bar{F}_{m_1} и \bar{F}_{m_2} имеют комплексными только множители $\Pi(z_{k_1}^{m_{k_1}} z_{k_2}^{m_{k_2}})$ и $\Pi(z_{k_1}^{*m_{k_1}} z_{k_2}^{*m_{k_2}})$.

Перейдя в системе уравнений (6) к амплитудам $r_v = |z_v|$ и фазам $\varphi_v = \arg z_v$, нетрудно усмотреть, что правая часть системы зависит только от величин r_v и $\theta = \sum_{k_1} \alpha_{k_1} \varphi_{k_1} - \sum_{k_2} \beta_{k_2} \varphi_{k_2}$. Исключив все φ_v , получим систему из $N + 1$ уравнения, написанную для $N + 1$ переменных r_v и θ .

Поскольку по переменной θ правая часть системы периодична (выражение (9) можно считать комплексным рядом Фурье), то в состоянии равновесия фазы колебаний φ_k связаны соотношением

$$\sum_k \alpha_{k_1} \varphi_{k_1} - \sum_{k_2} \beta_{k_2} \varphi_{k_2} = \theta_0 + 2n\pi,$$

где θ_0 — равновесное значение величины θ .

Последнее соотношение говорит о том, что в активной системе все существующие устойчивые квазигармонические колебания, частоты которых удовлетворяют уравнению (8 а) при неравных нулю коэффициентах α_{k_1} и β_{k_2} , взаимно синхронизированы. Колебания, которым соответствуют ω_k с равными нулю коэффициентами в соотношении (8 а), будут автономными (несинхронизированными). Амплитуды их, однако, будут зависеть от амплитуд и взаимной фазы θ всех других колебаний. Эта зависимость осуществляется посредством постоянных составляющих.

Обратим внимание теперь на то, что при $\alpha_{k_1} = 1$ (или $\beta_{k_2} = 1$) состояние равновесия, в котором $z_{k_1} = 0$ (или $z_{k_2} = 0$), не существует. Это означает, что имеет место принудительная синхронизация данной подсистемы со стороны совокупного воздействия других подсистем. При $\alpha_{k_1} > 1$ состояние равновесия $z_{k_1} = 0$ существует. Следовательно, в этом случае механизм синхронизации колебания z_{k_1} можно считать параметрическим.

Наконец, рассмотрим, как зависит величина взаимодействия квазигармонических колебаний и их взаимная синхронизация от характера нелинейности.

Предположим, что ряд, представляющий функцию $f(x)$, абсолютно сходится на всем промежутке $-\infty < x < \infty$; тогда

$$\lim_{n \rightarrow \infty} \frac{C_{n+1}}{C_n} = 0.$$

Это означает, что, начиная с некоторого значения n , выполняется неравенство

$$C_{n+1} \ll C_n.$$

Следовательно, согласно выражениям (9 б) и (9 в), сила синхронного взаимодействия тем больше, чем меньше величина

$$\sum_{k_1} \alpha_{k_1} + \sum_{k_2} \beta_{k_2}, \quad (10)$$

т. е. чем меньше сами коэффициенты α_{k_1} и β_{k_2} и чем меньшее число колебаний синхронно взаимодействуют. Если уравнение (8 а) справедливо только при таких значениях α_{k_1} и β_{k_2} , при которых выражение (10) достаточно велико, то синхронным взаимодействием по сравнению с несинхронным, не зависящим от соотношения частот (8 а), можно пренебречь.

2. НЕСИНХРОННОЕ ВЗАИМОДЕЙСТВИЕ

Рассмотрим подробнее наиболее простой частный случай, когда нелинейность можно аппроксимировать кубическим полиномом, а несинхронным взаимодействием пренебречь.

В этом случае система укороченных уравнений (6) может быть записана в виде

$$\begin{aligned}\dot{z}_v &= \left[\sigma_v + j\xi_v - C_3 B_v \sum_k \frac{z_k z_k^*}{1 + \delta_{vk}} \right] z_v, \\ \dot{z}_v^* &= \left[\sigma_v - j\xi_v - C_3 B_v \sum_k \frac{z_k z_k^*}{1 + \delta_{vk}} \right] z_v^*.\end{aligned}\quad (11)$$

Складывая эти уравнения, умноженные соответственно на z_v^* и z_v , и вычитая из первого, деленного на z_v , второе, деленное на z_v^* , получим

$$\begin{aligned}\dot{\rho}_v &= 2 \left[\sigma_v - C_3 \bar{B}_v \sum_k \frac{\rho_k}{1 + \delta_{vk}} \right] \rho_v, \\ \dot{\varphi}_v &= 2 \left[\xi_v - C_3 \widetilde{B}_v \sum_k \frac{\rho_k}{1 + \delta_{vk}} \right],\end{aligned}\quad (12)$$

где $\bar{B}_v = \operatorname{Re} B_v$, $\widetilde{B}_v = \operatorname{Im} B_v$, $\rho_v = z_v z_v^*$ — квадрат амплитуды колебаний, $\varphi_v = \ln(z_v/z_v^*)$ — удвоенная фаза колебания.

Второе уравнение системы (12), совместно с соотношением (5 а), определяет зависимость собственных частот нелинейной системы от амплитуд автоколебаний в стационарном режиме.

Из первого уравнения системы (12) следует, что существует устойчивое состояние равновесия $\rho_v = 0$ (для всех v), если все $\sigma_v < 0$, и неустойчивое, если хотя бы одно из σ_v больше нуля. Нетрудно усмотреть также, что если хотя бы одно из $C_3 \bar{B}_v$ меньше нуля, то бесконечность фазового пространства ρ_v устойчива.

Пусть все $C_3 \bar{B}_v > 0$, тогда система (12) может иметь состояние равновесия с отличными от нуля фазовыми координатами ρ_v , если соответствующие σ_v положительны.

Будем считать, что $\sigma_v = -\bar{\sigma}_v < 0$, а все остальные σ_v ($v \neq \bar{v}$) положительны. В этом случае состояние равновесия скотя бы одной фазовой координатой ρ_v , отличной от нуля, существовать не может.

Предположим, что существует состояние равновесия

$$\rho_{v_1} = \bar{\rho}_{v_1} \neq 0, \quad \rho_{v_2} = \bar{\rho}_{v_2} = 0,$$

где v_1 , v_2 и \bar{v} все различны и являются элементами множеств.

$$v = 1, 2, \dots, N$$

и что \bar{v} соответствуют отрицательные σ_v , а v_1 и v_2 — положительные.

Чтобы найти условия существования и устойчивости этого состояния равновесия, напишем первое уравнение системы (12) в виде

$$\begin{aligned}\dot{\rho}_{\bar{v}} &= -2 \left\{ \bar{\sigma}_{\bar{v}} + C_3 \bar{B}_{\bar{v}} \left[\sum_{\bar{k}} \frac{\rho_{\bar{k}}}{1 + \delta_{\bar{v}, \bar{k}}} + \sum_{k_1} \rho_{k_1} + \sum_{k_2} \rho_{k_2} \right] \right\} \rho_{\bar{v}}, \\ \dot{\rho}_{v_1} &= 2 \left\{ \sigma_{v_1} - C_3 \bar{B}_{v_1} \left[\sum_{\bar{k}} \rho_{\bar{k}} + \sum_{k_1} \frac{\rho_{k_1}}{1 + \delta_{v_1, k_1}} + \sum_{k_2} \rho_{k_2} \right] \right\} \rho_{v_1}, \\ \dot{\rho}_{v_2} &= 2 \left\{ \sigma_{v_2} - C_3 \bar{B}_{v_2} \left[\sum_{\bar{k}} \rho_{\bar{k}} + \sum_{k_1} \rho_{k_1} + \sum_{k_2} \frac{\rho_{k_2}}{1 + \delta_{v_2, k_2}} \right] \right\} \rho_{v_2}.\end{aligned}\quad (13)$$

Отличные от нуля фазовые координаты состояния равновесия определяются уравнением

$$\sigma_{v_1} - C_3 \bar{B}_{v_1} \sum_{k_1} \frac{\rho_{k_1}}{1 + \delta_{v_1, k_1}} = 0,$$

решение которого имеет вид

$$\bar{\rho}_{v_1} = 2 \left[\frac{\sum_{k_1} S_{k_1}}{N_1 - 1/2} - S_{v_1} \right], \quad (14)$$

где

$$S_{v_1} = \sigma_{v_1} / C_3 \bar{B}_{v_1},$$

а N_1 — число отличных от нуля фазовых координат состояния равновесия. Отсюда следуют условия существования состояния равновесия.

$$S_{v_1} < \frac{\sum_{k_1} S_{k_1}}{N_1 - 1/2},$$

которые должны выполняться для всех v_1 .

Подставляя в систему (13)

$$\rho_{\bar{v}} = R_{\bar{v}} e^{\tilde{p} t}, \quad \rho_{v_1} = \rho_{v_1} + R_{v_1} e^{\tilde{p} t}, \quad \rho_{v_2} = R_{v_2} e^{\tilde{p} t},$$

для достаточно малых $R_{\bar{v}}$, R_{v_1} и R_{v_2} получим систему алгебраических уравнений, определяющих корни характеристического уравнения,

$$\begin{aligned}\{\tilde{p} + 2 [\bar{\sigma}_{\bar{v}} + C_3 \bar{B}_{\bar{v}} \sum_{k_1} \bar{\rho}_{k_1}]\} R_{\bar{v}} &= 0, \\ \{\tilde{p} - 2 [\sigma_{v_1} - C_3 \bar{B}_{v_1} \sum_{k_1} \rho_{k_1}]\} R_{v_1} &= 0,\end{aligned}\quad (15)$$

$$\tilde{p} R_{v_1} + 2\eta_{v_1} \left[\sum_{\bar{k}} R_{\bar{k}} + \sum_{k_1} \frac{R_{k_1}}{1 + \delta_{v_1, k_1}} + \sum_{k_2} R_{k_2} \right] = 0,$$

где $\eta_{v_1} = C_3 \bar{B}_{v_1} \bar{\rho}_{v_1}$. Из первого уравнения этой системы следует, что по координатам $\rho_{\bar{v}}$ состояние равновесия всегда устойчиво. Из второго — по координатам ρ_{v_1} состояние равновесия устойчиво при выполнении неравенств

$$S_{v_1} < \sum_{k_1} \bar{\rho}_{k_1} = \frac{\sum_{k_1} S_{k_1}}{N_1 - 1/2}. \quad (16)$$

Положив в третьем уравнении $R_{\bar{k}} = R_{k_2} = 0$ и исключив R_{v_2} и R_{v_1} , получим характеристическое уравнение, определяющее устойчивость состояния равновесия по координатам r_{v_1} ,

$$2 \sum_{v_1} \frac{\eta_{v_1}}{p - \eta_{v_1}} + 1 = 0. \quad (17)$$

Пусть $\Phi(\tilde{p})$ — полином степени N_1 , имеющий N_1 различных действительных корней η_{v_i} , тогда

$$\frac{p \Phi'(\tilde{p})}{\Phi(\tilde{p})} = N_1 + \sum_{v_1} \frac{\eta_{v_1}}{p - \eta_{v_1}}. \quad (18)$$

Исключив из уравнений (17) и (18) сумму по v_1 , получим характеристическое уравнение в виде

$$(2N_1 - 1) \Phi(\tilde{p}) - 2\tilde{p} \Phi'(\tilde{p}) = 0, \quad (19)$$

из которого следует, что по координатам r_{v_1} состояние равновесия может быть устойчивым в одном единственном случае — при $N_1 = 1$.

Действительно, пусть

$$\Phi(\tilde{p}) = \sum_{n=0}^{N_1} a_n \tilde{p}^n,$$

тогда уравнение (19) имеет вид

$$\sum_n a_n [2(N_1 - n) - 1] \tilde{p}^n = 0. \quad (20)$$

Поскольку $\Phi(\tilde{p})$ имеет только действительные положительные корни, то $(-1)^{N_1-n} a_n > 0$ и, следовательно, при $N_1 > 1$ характеристический полином (20) имеет отрицательные коэффициенты. Это показывает, что состояние равновесия неустойчиво.

Таким образом, рассмотренная здесь система, независимо от числа степеней свободы, может иметь устойчивое колебание только на одной из частот (для системы с двумя степенями свободы это показано в работе [4]).

Из уравнения (20) следует, что по координатам r_{v_1} оно будет устойчивым всегда, по координатам r_{v_2} — при выполнении неравенств

$$S_{v_2} < 2S_{v_1}$$

для всех значений v_2 .

Если эти неравенства выполняются для ряда значений v_1 , то, в зависимости от начальных условий, в системе может установиться любой из соответствующих этим значениям режимов.

Полученный здесь результат физически может быть объяснен на основе уравнений (13). Действительно, наличие колебания на какой-либо одной частоте ω_k ухудшает условие возбуждения колебаний на всех остальных частотах, в том числе и на частоте ω_k . Однако условие возбуждения своего колебания ухудшается меньше, чем все остальные, так как в свое уравнение амплитуда r_v входит с весом $1/2$, тогда как все другие с весом 1.

Полученный здесь результат в какой-то степени оправдывает попытки исследования системы со многими степенями свободы в одиночастотном приближении. Однако, по-видимому, это можно делать только тогда, когда отсутствует синхронное взаимодействие и система содержит только одну нелинейность. Наличие синхронного взаимодействия может привести к возникновению многочастотного режима колебаний [14, 15]. То же самое, видимо, возможно и при отсутствии синхронного взаимодействия, в системе, содержащей две или более нелинейности.

В заключение следует отметить, что данная работа появилась как результат неоднократных обсуждений изложенных выше вопросов с М. И. Рабиновичем. Выражаю ему свою признательность.

ЛИТЕРАТУРА

1. Van der Pol B., Radio Review, 1, 701 (1920).
2. А. А. Андронов, Полное собрание трудов, изд. АН СССР, 1965, стр. 70.
3. Л. И. Мандельштам, Полное собрание трудов, изд. АН СССР, 1947, стр. 13.
4. А. А. Андронов, А. А. Витт, ЖТФ, 4, № 1, 122 (1934).
5. А. Г. Майер, Уч. зап. ГГУ, вып. 2, 3 (1935).
6. Р. М. Кантор, Изв. высш. уч. зав. — Радиотехника, 5, № 3, 356 (1962).
7. Ю. Н. Ефимов, Изв. высш. уч. зав. — Радиотехника, 5, № 3, 347 (1962).
8. Ф. П. Жарков, Тр. Всесоюзного научно-технического совещания, Львов, 1962, стр. 458.
9. А. Е. Каплан, Ю. А. Кравцов, В. А. Рылов, Параметрические генераторы и делители частоты, изд. Сов. радио, М., 1966,
10. Ю. А. Азъян, В. В. Мигулин, Радиотехника и электроника, 1, № 4, 418 (1956).
11. В. П. Рубаник, Изв. высш. уч. зав. — Радиотехника, 6, № 2, 276 (1963).
12. Н. А. Молчанов, Изв. высш. уч. зав. — Радиофизика, 10, № 7, 987 (1967).
13. Ю. И. Марченко, Изв. высш. уч. зав. — Радиофизика, 10, № 11, 1533 (1967).
14. М. И. Рабинович, Е. И. Якубович, Изв. высш. уч. зав. — Радиофизика, 8, № 5, 987 (1966).
15. М. И. Рабинович, ДАН СССР, 191, № 9, 1253 (1970).
16. А. В. Гапонов, Л. А. Островский, М. И. Рабинович, Изв. высш. уч. зав. — Радиофизика, № 2, 163 (1970).
17. Ю. И. Неймарк, Тр. ГИФТИ и радиофакультета ГГУ, 30, 13, серия физическая, 1956.
18. Л. И. Мандельштам, Полное собрание трудов, изд. АН СССР, М., 1955, стр. 234.
19. Н. А. Железцов, ЖТФ, 13, № 4, 495 (1948).
20. Н. Н. Богоявленов, Ю. А. Митропольский, Асимптотические методы теории нелинейных колебаний, издание 3, Физматгиз, М., 1963.
21. С. А. Ахманов, Р. В. Хохлов, Проблемы нелинейной оптики, изд. Наука, М., 1964.
22. Н. Блюмберген, Нелинейная оптика, изд. Мир, М., 1966
23. В. Н. Цытович, Нелинейные эффекты в плазме, изд. Наука, М., 1966.

Горьковский государственный
университет

Поступила в редакцию
20 января 1970 г.

ON INTERACTIONS OF QUASI-HARMONICAL OSCILLATIONS IN WEAKLY NONLINEAR SYSTEMS

L. V. Postnikov

An autonomous dynamic system described by an N^{th} order differential equation with one nonlinearity is considered. Conditions are stated under which oscillations composed of a sum of quasi-harmonical components are possible to exist in a system. A system of shortened equations is written. It is shown that under an approximation of nonlinearity by a cubic polynomial and in the absence of synchronous interaction only a single-frequency mode is possible in an autooscillations system.

УДК 621.391.193

О КЛАССИФИКАЦИИ ПО СИНДРОМУ

З. С. Баталова, И. М. Иванова, Ю. И. Неймарк

Вводится понятие диагностических сочетаний признаков и предлагаются способы отыскания весов, с которым они «голосуют» за тот или иной класс. Предлагается алгоритм распознавания объектов по «обобщенному синдрому». Приведены результаты решения конкретных задач.

В работе [1] предложена программа «Кора-3» распознавания нефтеносных и водоносных пластов. В основе этой программы лежит построение новых «признаков», представляющих собою сочетания из трех исходных двоичных признаков, которые встречаются либо только у нефтеносных пластов, либо только у водоносных. С помощью таких «признаков» удалось оценить информативность различных геофизических параметров, выяснить некоторые связи между ними, получить достаточно достоверные результаты распознавания. В клинической медицине сочетание признаков, характерное для определенного заболевания, называют синдромом. Понятие синдрома известно очень давно и используется врачами для дифференциальной диагностики и прогнозирования различных заболеваний. Знание синдрома позволяет обычно точно установить диагноз или прогноз заболевания. Однако лишь для немногих заболеваний известны синдромы (корь, скарлатина и др.). В большинстве случаев, представляющих затруднение для постановки врачебного диагноза, синдромы неизвестны. При поиске синдрома с помощью электронной цифровой машины на основе некоторого статистического материала в ряде конкретных задач медицинской диагностики и прогнозирования (непосредственный и отдаленный прогноз состояния больного после операции при желчно-каменной болезни, диагностика ревматизма и тиреотоксикоза и др.) оказалось, что синдром отсутствует. Это послужило причиной того, что в работах [2-4] авторов понятие синдрома было ослаблено и заменено близким к нему понятием диагностического сочетания, встречающегося в одном классе значительно чаще, чем в других. В дальнейшем аналогичное понятие диагностического сочетания было использовано в работе [5] для диагностики центрального рака легкого. Можно указать также на алгоритмы распознавания, в основе которых лежит поиск сочетаний признаков, близких по своему смыслу к диагностическим сочетаниям. Сюда относятся, например, алгоритм [6, 7], в котором находятся наиболее «информационные» подпространства признаков, алгоритм построения тупиковых тестов [8] и др. [9].

В работах [2, 3] принято, что между различными признаками и сочетаниями признаков отсутствует какая-либо зависимость. Это приводило к отбору большого числа (порядка нескольких тысяч) диагностических сочетаний и в связи с этим к невозможности объективной оценки каждого диагностического сочетания. В настоящей работе наряду с прежним вариантом рассмотрен случай, при котором появляется возможность учета зависимости между различными сочетаниями признаков. Независимых диагностических сочетаний оказывается значительное

меньше, причем каждое из них можно рассматривать как новый признак для описания объектов. Отметим, что новое пространство признаков можно считать полученным в результате существенно нелинейного преобразования исходного пространства. В работе предложен эвристический способ вычисления весов сочетаний и построения «обобщенного синдрома», позволяющего в ряде случаев получить достаточно достоверные результаты распознавания.

В качестве примера приведены данные решения конкретных задач медицинской диагностики.

1. Допустим, что описание объекта x характеризуется N действительными параметрами (признаками) x_1, x_2, \dots, x_m , каждый из которых принимает конечное число значений. Пусть обучающая последовательность состоит из двух классов X_1 и X_2 , содержащих J_1 и J_2 объектов соответственно. В простейшем случае, когда можно предположить отсутствие зависимостей между признаками и их сочетаниями, диагностическое сочетание определяется следующим образом. Допустим, что сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ встречается в классе X_1 m_1 раз, а в классе X_2 — m_2 раз. Определим частоту этого сочетания в каждом классе

$$\alpha_1 = \frac{m_1}{J_1}, \quad \alpha_2 = \frac{m_2}{J_2}.$$

Назовем сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ диагностическим в классе X_1 , если

$$\alpha_1 \geq \delta_k, \quad \alpha_1 \geq \beta_k \alpha_2, \quad (1)$$

где δ_k , β_k — так называемые пороги диагностичности. Порог β_k указывает, во сколько раз это сочетание встречается чаще в классе X_1 , чем в классе X_2 , порог δ_k ограничивает снизу частоту сочетания в классе X_1 . Например, выполнение условий (1) при значениях $\beta_k = 8$, $\delta_k = 0,2$ означает, что сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ в классе X_1 встречается не реже, чем в 20 случаях из 100, и в 8 раз чаще, чем в классе X_2 . При выполнении соотношений

$$\alpha_2 \geq \delta_k, \quad \alpha_2 \geq \beta_k \alpha_1$$

сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ будем называть диагностическим в классе X_2 .

Аналогично определяется диагностическое сочетание в случае, когда имеется S классов X_1, X_2, \dots, X_s , содержащих J_1, J_2, \dots, J_s объектов соответственно. При этом частота сочетания $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ в каждом классе равна

$$\alpha_s = \frac{m_s}{J_s} \quad (s = 1, 2, \dots, S), \quad (2)$$

Сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ называется диагностическим в классе X_s , если

$$\begin{aligned} \alpha_{s*} &= \max_s \{ \alpha_s \}, \\ \alpha_{s*} &\geq \delta_k, \quad \alpha_{s*} \geq \beta_k \alpha_s \quad (s \neq s*). \end{aligned} \quad (3)$$

Приведенное определение диагностического сочетания из k признаков не учитывает того, что некоторые составляющие его признаки также могут образовать диагностическое сочетание. В большинстве случаев это приводит к тому, что диагностичность указанного сочетания является прямым следствием диагностичности одного или нескольких входя-

ших в него сочетаний. Так, если частота сочетания $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ в классе X_1 во много раз превосходит частоты его в других классах, то большинство сочетаний, содержащих $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$, будет также диагностическим в этом классе, но эта диагностичность обусловливается в основном сочетанием $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$. При решении конкретных задач классификации объектов это обстоятельство приводит к отбору излишне большого числа диагностических сочетаний.

Для учета зависимости между различными диагностическими сочетаниями признаков и уменьшения тем самым числа диагностических сочетаний примем другой способ определения величин $\alpha_1, \alpha_2, \dots, \alpha_S$. В случае $k = 2$ он состоит в следующем. Пусть сочетание (x_{i_1}, x_{i_2}) встречается в классе X_s $m_{i_1 i_2}^s$ раз, а значения x_{i_1} и x_{i_2} встречаются в нем $m_{i_1}^s$ и $m_{i_2}^s$ раз соответственно. Тогда в качестве α_s берется величина

$$\alpha_s = \frac{m_{i_1 i_2}^s}{q_s}, \quad q_s = \min \{ m_{i_1}^s, m_{i_2}^s \}, \quad (4)$$

где q_s означает, сколько раз сочетание (x_{i_1}, x_{i_2}) могло встретиться в классе X_s . При значении q_s равном нулю, величина α_s также равна нулю.

В общем случае, когда $k = 3, 4, \dots$, величина α_s определяется аналогично. При этом сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ покрывается k сочетаниями из $(k - 1)$ символов $(x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}})$, $(x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}, x_{i_k})$, ..., $(x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}, x_{i_k})$. Пусть эти сочетания встречаются в классе X_s $m_{i_1 i_2 \dots i_{k-1} i_k}^s$ раз соответственно, а сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ встречается в нем $m_{i_1 i_2 \dots i_k}^s$ раз. Найдем

$$q_s = \min \{ m_{i_1 i_2 \dots i_{k-1}}^s, m_{i_1 i_2 \dots i_{k-2} i_k}^s, \dots, m_{i_2 i_3 \dots i_k}^s \}.$$

Тогда

$$\alpha_s = \frac{m_{i_1 i_2 \dots i_k}^s}{q_s} \quad (s = 1, 2, \dots, S).$$

Теперь сочетание $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ назовем диагностическим в классе X_{s^*} , если

$$\alpha_{s^*} = \max_s \{ \alpha_s \}, \quad (5)$$

$$\frac{q_{s^*}}{J_{s^*}} \geq \epsilon_k, \quad \alpha_{s^*} \geq \delta_k, \quad \alpha_{s^*} \geq \beta_k \alpha_s \quad (s \neq s^*),$$

где $\epsilon_k, \delta_k, \beta_k$ — пороги диагностичности. В отличие от (3), здесь введен еще один порог диагностичности ϵ_k , ограничивающий снизу наименьшую частоту диагностического сочетания из $(k - 1)$ символов.

Все сочетания находятся путем прямого перебора по одному, двум и т. д. градациям разных признаков. В результате отбора сочетаний, согласно условий (3) или (5), получается список диагностических сочетаний.

Примем каждое диагностическое сочетание в качестве нового двоичного признака z_n , n — порядковый номер диагностического сочетания. В зависимости от наличия или отсутствия у объекта этого сочетания будем кодировать переменное z_n значением 1 или 0. В результате пере-

кодировки объектов получим множества $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_s$ точек в пространстве признаков z_1, z_2, \dots, z_N , содержащие лишь необходимую информацию о разделимости этих множеств. Из условий (3) и (5) видно, что увеличение пороговых значений уменьшает размерность пространства признаков и вместе с тем увеличивает возможность полученного набора признаков оказаться недостаточным для разделимости множеств $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_s$. Поэтому пороговые значения выбираются так, чтобы сохранились свойства разделимости этих множеств без значительного увеличения размерности пространства признаков. Для распознавания объектов по построенным новым признакам (диагностическим сочетаниям) можно использовать любой алгоритм классификации.

В некоторых случаях достаточно достоверные результаты можно получить с помощью следующего алгоритма распознавания, основанного на вычислении весов диагностических сочетаний. Опишем его сначала для случая двух классов $X_1 \{ x^i(z_1^i, z_2^i, \dots, z_N^i) \} (i=1, 2, \dots, J_1)$ и $X_2 \{ x^j(z_1^j, z_2^j, \dots, z_N^j) \} (j=1, 2, \dots, J_2)$. В качестве веса диагностического сочетания z_n в классе X_1 берется величина

$$w_n^1 = \alpha_1 / (\alpha_2 + \delta), \quad (6)$$

где δ — малая положительная величина, отличная от нуля лишь тогда, когда $\alpha_2 = 0$. В классе X_2 вес этого сочетания $w_n^2 = 0$. Если сочетание z_n диагностическое в классе X_2 , то $w_n^1 = 0$, а величина

$$w_n^2 = \alpha_2 / (\alpha_1 + \delta). \quad (7)$$

Будем называть выражения

$$Z_1 = \sum_{n=1}^N w_n^1 z_n, \quad Z_2 = \sum_{n=1}^N w_n^2 z_n$$

«обобщенными синдромами» для классов X_1 и X_2 соответственно. Для установления, к какому из двух классов принадлежит объект $x(z_1, z_2, \dots, z_N)$, вычисляются величины Z_1 и Z_2 . Объект относится к классу X_1 , если имеет место неравенство

$$Z_1 - \gamma_1 Z_2 > d_1, \quad (8)$$

и к классу X_2 , если

$$Z_2 - \gamma_2 Z_1 > d_2. \quad (9)$$

Отказ в ответе дается в том случае, если удовлетворяются одновременно неравенства

$$Z_1 - \gamma_1 Z_2 \leq d_1, \quad Z_2 - \gamma_2 Z_1 \leq d_2. \quad (10)$$

Здесь величины $\gamma_s \geq 1$, $d_s \geq 0$ ($s = 1, 2$) являются параметрами алгоритма. Для их выбора построим на плоскости обобщенных синдромов Z_1, Z_2 точки с координатами

$$Z_1^i = \sum_{n=1}^N w_n^1 z_n^i, \quad Z_2^j = \sum_{n=1}^N w_n^2 z_n^j,$$

соответствующие множествам X_1 и X_2 (рис. 1). Прямые $Z_1 = \gamma_1 Z_2 + d_1$ и $Z_2 = \gamma_2 Z_1 + d_2$ проводятся на основании некоторого критерия качества распознавания, в качестве которого можно взять, например, величину $K = p + q$, возрастающую вместе с ростом частот ошибок и отказов p и q для точек обучающей последовательности. Величина σ характе-

ризует относительную нежелательность ошибок и отказов.

При наличии S классов X_1, X_2, \dots, X_s в качестве веса диагностического сочетания z_n в классе X_{s^*} берется величина

$$w_n^{s^*} = \frac{\alpha_{s^*}}{\max_{s \neq s^*} \{\alpha_s\} + \delta}.$$

Вес этого сочетания для остальных классов полагается равным нулю. Для распознавания нового объекта находятся обобщенные синдромы

$$Z_1 = \sum_{n=1}^N w_n^1 z_n,$$

$$Z_2 = \sum_{n=1}^N w_n^2 z_n, \dots, Z_s = \sum_{n=1}^N w_n^s z_n.$$

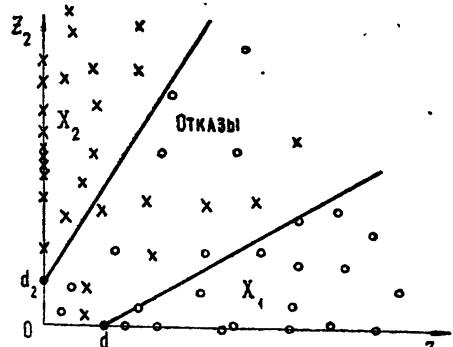


Рис. 1.

Объект относится к классу X_s , если выполняются $(s - 1)$ неравенств

$$Z_s - \gamma_{sr} Z_r > d_{sr} \quad (r \neq s).$$

Если выполняются хотя бы два неравенства

$$Z_s - \gamma_{sr} Z_r \leq d_{sr} \quad (s, r = 1, 2, \dots, S; \quad r \neq s),$$

то дается отказ в ответе.

Выбор параметров γ_{sr} и d_{sr} осуществляется аналогично случаю $S = 2$ при рассмотрении всевозможных плоскостей обобщенных синдромов Z_s, Z_r ($r \neq s$).

В некоторых случаях может оказаться, что множества $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_s$ в пространстве признаков z_1, z_2, \dots, z_N линейно разделимы. Тогда для их разделения можно использовать линейные алгоритмы [10, 11] классификации, с помощью которых в случае $S = 2$ можно найти линейную форму вида

$$\sum_{n=1}^N w_n z_n + w_{N+1} > 0, \quad (11)$$

такую, что если выполняется неравенство

$$\sum_{n=1}^N w_n z_n + w_{N+1} > \Delta, \quad (12)$$

то объект относится к классу X_1 , если же

$$\sum_{n=1}^N w_n z_n + w_{N+1} < -\Delta, \quad (13)$$

то он относится к классу X_2 . При

$$-\Delta \leq \sum_{n=1}^N w_n z_n + w_{N+1} \leq \Delta \quad (14)$$

дается отказ в ответе. Неравенства (8) — (10) приводятся к виду (12) — (14), если положить

$$Z_1 = \sum_{n=1}^N w_n z_n \quad (w_n > 0),$$

$$Z_2 = \sum_{n=1}^N |w_n| z_n \quad (w_n < 0),$$

$$\gamma_1 = \gamma_2 = 1, \quad d_1 = \Delta - w_{N+1}, \quad d_2 = w_{N+1} + \Delta.$$

2. Отбор диагностических сочетаний в соответствии с условиями* (2), (3) и (4), (5) был проведен в задаче дифференциального диагноза центрального рака легкого и хронической неспецифической пневмонии. Статистический материал для решения этой задачи составили историю болезни и рентгенограммы 207 больных раком легкого (класс X_1) и 156 больных хронической пневмонией (класс X_2). Схема обследования больного содержала 36 признаков, каждый из которых имел от двух до семи градаций.

Сочетания отыскивались при значениях $\beta_1 = 8$, $\beta_2 = 4$, $\varepsilon_k = \delta_k = 0,1$ ($k = 1, 2$) (эти параметры не являются оптимальными). В табл. 1 приведены результаты диагностики 363 больных по формулам (8)–(10) при $\gamma_1 = \gamma_2 = 3,6$, $d_1 = 17$, $d_2 = 53$. Как видно из таблицы, при условиях (4), (5), учитывающих зависимость диагностических сочетаний, получаются несколько лучшие результаты, чем при условиях (2), (3).

Таблица 1

Заболевание	Всего	Номер формулы	Число верных ответов	Число ошибок	Число отказов
Рак	207	2,3	179	9	19
		4,5	196	3	8
Хроническая пневмония	156	2,3	130	12	14
		4,5	140	7	9
Всего	363	2,3	309	21	33
		4,5	336	10	17

В табл. 2 представлены результаты распознавания по формулам (8)–(10) при различных значениях параметра β_2 , параметры ε_k , δ_k имеют прежние значения, параметры γ_s , d_s ($s = 1, 2$) выбирались по результатам экзамена.

С увеличением β_2 уменьшается число диагностических сочетаний из двух признаков. Вместе с тем уменьшается число ошибочных ответов, но возрастает число отказов. Дальнейшее увеличение β_2 нецелесообразно, так как уже при $\beta_2 = 8$ имеются случаи пересечения множеств \tilde{X}_1 и \tilde{X}_2 . Из таблицы следует, что наилучшее распознавание получается при $\beta_2 = 6$. При этом значении диагноз рака легкого устанавливается для 97,6% случаев с достоверностью 0,99, диагноз хронической пневмонии — для 93% с достоверностью 0,98. Заметим, что клинический диагноз рака легкого в этой группе больных был поставлен для 77,4% слу-

* Результаты, полученные при отборе признаков по формулам (2), (3), заимствованы из диплома О. А. Климовской.

чаев с достоверностью 0,8, клинический диагноз хронической пневмонии — для 75% случаев с достоверностью 0,78 [12].

Таблица 2

Заболевание	β_2	Число N диагностических сочетаний	Число верных ответов	Число ошибок	Число отказов
Рак (207)	6	320	200	2	5
	8	243	193	1	13
	10	226	189	1	17
Хроническая пневмония (156)	6	188	142	3	11
	8	176	139	2	15
	10	176	137	1	18

Найденные при значении $\beta_2 = 6\ 488$ диагностические сочетания были взяты в качестве новых признаков. Для разделения множеств \tilde{X}_1 и \tilde{X}_2 , полученных в результате перекодировки статистического материала, использовался минимаксный алгоритм [11]*.

В результате построения разделяющей гиперплоскости из 488 признаков-сочетаний было выделено 65 диагностических сочетаний. Данные экзамена 363 точек с помощью найденной линейной формы представлена в табл. 3. Здесь же приведены результаты, полученные с помощью формул (8)–(10) и алгоритма распознавания по общей близости признаков [2, 3]. Экзамен проводился поочередно для каждой точки обучающей последовательности на основе оставшихся 362 точек в предположении, что все исходные 36 признаков и все выделенные диагностические сочетания имеют одинаковое значение для постановки диагноза. Как видно из таблицы, этот алгоритм дает несколько худшие результаты распознавания. По-видимому, это связано с неверным допущением о равной значимости всех признаков и сочетаний при диагностике этих заболеваний**.

Таблица 3

Алгоритм	Число признаков (сочетаний)	Число верных ответов	Число ошибок	Число отказов
Распознавание по «обобщенному синдрому»	408	342	5	16
Минимаксный алгоритм	65	345	5	13
Распознавание по общей близости признаков	36	300	29	34
	488	311	25	27

* Линейная форма по минимаксному алгоритму была получена А. Н. Дурново.

Для ее получения были исключены 15 точек множеств \tilde{X}_1 и \tilde{X}_2 , наиболее близко расположенных друг к другу. Впоследствии некоторым из этих точек был дан ошибочный ответ или отказ в ответе.

** Ранее было выяснено [12], что различные признаки имеют разное влияние на диагностику центрального рака легкого и хронической пневмонии.

При сравнении весов диагностических сочетаний, найденных по формулам (6), (7), и весовых коэффициентов в формуле (11), полученных с помощью минимаксного алгоритма, оказалось, что выделяются одни и те же существенные диагностические сочетания, «голосующие» с относительно большим весом за свой класс. В отношении менее существенных диагностических сочетаний, имеющих незначительные весовые коэффициенты w_n , обнаружились небольшие расхождения: некоторые сочетания, «голосующие» за класс X_1 , были диагностическими в классе X_2 . В то же время большая группа диагностических сочетаний получила нулевые весовые коэффициенты.

Описанный в разделе 1 алгоритм применялся также для решения задачи диагностики бесприступно текущего ревматизма у детей. В качестве исходного материала были взяты истории болезни 506 больных детей ревматизмом (класс X_1) и 200 историй болезни больных, у которых были тонзилло-кардиальный синдром, туберкулезная интоксикация, синусит, холецистопатия и другие заболевания, имеющие сходный с ревматизмом клинический симптомокомплекс (класс X_2). Больные обследовались по определенной схеме, содержащей 32 признака, каждый из которых имел от двух до восьми градаций. Схема обследования приведена в [13]. Первоначально в обучающую последовательность были включены 143 больных из класса X_1 и 101 больной из класса X_2 , при этом из 32 исходных признаков были выбраны лишь 15 признаков. При значениях $\beta_1 = 8$, $\beta_2 = 6$, $\beta_3 = 4$, $\alpha_k = 0,1$ ($k = 1, 2, 3$) были отобраны 19 диагностических градаций признаков, 486 и 597 диагностических сочетаний из двух и из трех признаков соответственно. Число диагностических сочетаний и результаты распознавания обучающей последовательности по «обобщенному синдрому» с различными числами признаков k приведены в табл. 4. Для сопоставления в нижней строке таблицы приведены данные экзамена с помощью алгоритма по общей близости признаков.

Таблица 4

Число диагностических сочетаний (признаков)	k	Число верных ответов	Число ошибок	Число отказов
19	1	205	2	37
486	2	241	1	2
597	3	230	—	14
505	1, 2	242	—	2
1102	1, 2, 3	242	—	2
15	—	236	5	3

Как видно, наилучший результат получается при учете 505 диагностических сочетаний из одного и двух признаков. Эти сочетания были взяты в качестве новых признаков. После перекодировки статистического материала множества \tilde{X}_1 и \tilde{X}_2 оказались линейно разделимыми. При последующем экзамене 462 точек по найденной линейной формуле было получено 379 верных ответов, 31 ошибка, 52 отказа. Примерно такие же результаты были получены при экзамене по обобщенному синдрому, а также с помощью алгоритма распознавания по общей близости признаков. В дальнейшем была взята обучающая последовательность из 396 точек (229 из класса X_1 , 167—из класса X_2), причем учитывался

полный набор 32 исходных признаков. При $\beta_1 = 10$, $\beta_2 = 14$, $\epsilon_k = \delta_k = 0,1$ ($k=1, 2$) были отобраны 663 диагностических сочетания. При экзамене по обобщенному синдрому 706 больных для значений $\gamma_1 = 0,45$, $d_1 = 21$, $\gamma_2 = 9$, $d_2 = 100$ было получено 98% верных ответов, 2% ошибок, отказы в ответе даны в 4,5% случаев.

ЛИТЕРАТУРА

1. М. М. Бонгард, Н. Д. Вайнцайп, Ш. А. Губерман, М. П. Извекова, М. С. Смирнова, Проблемы расширения возможностей автоматов, изд. институт Автоматики и телемеханики, № 5, 1964.
2. Ю. И. Неймарк, З. С. Баталова, М. В. Введенская, Д. А. Рогова, Т. Н. Архангельская, И. В. Мамонова, А. Ф. Ходова, сб. Вопросы кардиологии, Волго-Вятское книжное изд-во, Г., 1966, стр. 243.
3. Ю. И. Неймарк, З. С. Баталова, Уч. зап. НИИ ПМК и ф-та ВМиК ГГУ, 1967, стр. 293.
4. Ю. И. Неймарк, М. Б. Шмерельсон, З. С. Баталова, Н. В. Кузнецова, Е. Н. Шкадаревич, А. С. Носкова, З. А. Горохова, Материалы к Всесоюезному симпозиуму, Волго-Вятское книжное изд-во, Г., 1965, стр. 146.
5. П. Е. Кунин, С. Я. Марморштейн, Е. Б. Вольфсон, Е. А. Мехтенштейн, В. А. Бояджан, Вестник рентгенологии и радиологии, № 5, 9 (1967).
6. В. Л. Браиловский, Автоматика и телемеханика, 23, № 12, 1643 (1962).
7. В. Л. Браиловский, Изв. АН СССР, Техническая кибернетика, № 2, 30 (1964).
8. А. Н. Дмитриев, Ю. И. Журавлев, Ф. П. Кренделеев, Дискретный анализ, вып. 7, 3 (1966).
9. И. А. Бергер, Кибернетика, № 3, 80 (1968).
10. Н. Нильсон, Обучающиеся машины, изд. Мир, М., 1967.
11. Ю. И. Неймарк, М. Д. Брейдо, А. Н. Дурново, Изв. АН СССР, Техническая кибернетика, № 2, 144 (1970).
12. Н. Е. Яхонтов, М. А. Кузнецова, Ю. И. Неймарк, З. С. Баталова, И. М. Иванова, сб. «Автоматизация. Организация. Диагностика», изд. Наука, М. (в печати).
13. Я. Г. Любавин, Ю. И. Неймарк, З. С. Баталова, Ю. Г. Васин, В. И. Кочкин, Н. Д. Образцова, сб. «Автоматизация. Организация. Диагностика», изд. Наука, М. (в печати).

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
14 мая 1970 г.

ON SYNDROME CLASSIFICATION

Z. S. Batanova, I. M. Ivanova, Yu. I. Neimark

The concept of diagnostic combinations of symptoms is introduced and ways to find weights with which the symptoms „vote for“ one or another class are put forward. An algorithm to recognize objects by „the generalized syndrome“ is proposed. The solutions of specific problems are given.

УДК 519.8

ВЫБОР ИСПЫТАНИЙ И УСЛОВИЕ ОСТАНОВКИ В ОДНОМЕРНОМ ГЛОБАЛЬНОМ ПОИСКЕ

Р. Г. Стронгин

Исследуется правило выбора испытаний и свойства условия остановки в одномерном алгоритме глобального поиска [3, 4]. Предлагается модифицированный вариант алгоритма, и иллюстрируется повышение скорости поиска, обусловленное модификацией

Настоящая работа является продолжением исследований [1–4], в которых развивается информационно-статистическая теория поиска экстремума функций.

СТАТИСТИЧЕСКОЕ ОПИСАНИЕ МИНИМИЗИРУЕМОЙ ФУНКЦИИ

Рассмотрим задачу поиска глобального минимума неизвестной одномерной действительной функции $\varphi(x)$, $x \in [a, b]$, при измерениях без помех. В качестве сценки для искомого глобального минимума примем точку x_{i^*} глобального минимума функции $\varphi(x_i)$ (в дальнейшем используется обозначение $\varphi(i) = \varphi(x_i)$), определенной на множестве равноотстоящих узлов ε -сети

$$x_0 = a, \quad x_1 = a + \varepsilon, \dots, \quad x_i = a + i\varepsilon, \dots, \quad x_n = a + n\varepsilon = b,$$

где ε определяет точность приближения x_{i^*} .

В процессе поиска точки $i^* \in I = \{0, 1, \dots, n\}$ производятся вычисления значений функции $\varphi(i)$ (будем называть их испытаниями) в некоторых точках множества I .

Определение 1. Множество $\omega = \omega_k = \{(i_s, z_s): 0 \leq s \leq k\}$ есть множество результатов $z_s = \varphi(i_s)$ испытаний, осуществленных за $k+1$ шагов поиска в точках i_s ($0 \leq s \leq k$), нумерация которых корректируется после каждого шага с тем, чтобы выполнялось условие $0 = i_0 < i_1 < \dots < i_k = n$ (предполагается, что первые два испытания всегда осуществляются в точках $i = 0$ и $i = n$).

В соответствии с [1–3] будем считать, что $\varphi(i)$, $i \in I$, является некоторой реализацией случайной функции, порождаемой разностным уравнением $\varphi(i) - \varphi(i-1) = \beta_i$ ($1 \leq i \leq n$), где β_i — независимые нормальные случайные величины со стандартами $\sigma_i = cm$, $c > 0$, $m > 0$ ($1 \leq i \leq n$) и математическими ожиданиями

$$m_i = m_i(\alpha) = \begin{cases} -m, & i \leq \alpha \\ m, & i > \alpha \end{cases} \quad (1 \leq i \leq n),$$

которые зависят от целочисленной случайной величины α ($\alpha \in I$) (мы будем называть ее состоянием природы), имеющей априорное распределение вероятностей $\xi(\alpha)$ (предполагается, что $\xi(\alpha) > 0$ при $\alpha \in I$).

Параметр m из (1), являющийся средним значением абсолютной величины первой разности функции $\varphi(i)$, $i \in I$, оценивается [2, 3] на основе ω_k по формуле

$$m = \begin{cases} rM & (M > 0) \\ 1 & (M = 0), \end{cases} \quad (2)$$

где

$$M = \max_{1 \leq s \leq k} |(z_s - z_{s-1})/(i_s - i_{s-1})| \quad (3)$$

и r — коэффициент (предполагается, что $r > 1$).

Из сделанных допущений вытекает [2, 3], что при достаточно малом $c > 0$ вероятность расположения глобального минимума функции $\varphi(i)$ в точке $\alpha \in I$ после k ($k \geq 1$) шагов поиска хорошо описывается априорной вероятностью $\xi(\alpha / \omega_k)$ для состояния природы α .

При этом, если $i_{s-1} \ll \alpha \ll i_s$ ($1 \leq s \leq k$), то

$$\xi(\alpha / \omega) = h(\omega, \alpha) \xi(\alpha) / \sum_{i \in I} h(\omega, i) \xi(i), \quad (4)$$

причем

$$h(\omega, \alpha) = \exp \left\{ -\frac{(\alpha - \alpha_s^*)^2}{2 \rho_s^2} \right\} \exp \left\{ \frac{R(s)}{2mc^2} \right\}, \quad (5)$$

$$\alpha_s^* = (i_s + i_{s-1})/2 - (z_s - z_{s-1})/2m, \quad (6)$$

$$R(s) = m(i_s - i_{s-1}) + \frac{(z_s - z_{s-1})^2}{m(i_s - i_{s-1})} - 2(z_s + z_{s-1}), \quad (7)$$

$$\rho_s^2 = (c/2)^2(i_s - i_{s-1}).$$

Определение 2. α^* ($\alpha^* \in I$) есть ближайшее целое число к α_t^* , где t ($1 \leq t \leq k$) удовлетворяет условию $R(t) = \max_{1 \leq s \leq k} R(s)$.

Лемма 1. При любом априорном распределении $\xi(\alpha)$ точка α^* является наивероятнейшей (при заданном ω) точкой расположения глобального минимума функции $\varphi(i)$, если c достаточно мало.

Доказательство немедленно следует из (4) и (5), если положить $c \rightarrow 0$.

Определение 3. $\mu_\alpha(l)$ есть математическое ожидание значения $\varphi(l)$ при условии, что природа находится в состоянии $\alpha \in I$.

При $i_{s-1} \ll l \ll i_s$ $\mu_\alpha(l)$ имеет вид [2, 3]

$$\mu_\alpha(l) = \mu^*(l) - v_\alpha(l), \quad (8)$$

где

$$v_\alpha(l) = \left(\frac{2m}{i_s - i_{s-1}} \right) \begin{cases} 0 & (\alpha \leq i_{s-1}, \alpha \geq i_s) \\ (i_s - l)(\alpha - i_{s-1}) & (i_{s-1} \leq \alpha \leq l), \\ (l - i_{s-1})(i_s - \alpha) & (l \leq \alpha \leq i_s) \end{cases} \quad (9)$$

$$\mu^*(l) = [l(z_s - z_{s-1}) + (z_{s-1}i_s - z_si_{s-1})]/(i_s - i_{s-1}). \quad (10)$$

Из (8) — (10) следует, что

$$\mu_\alpha(i_s) = \mu^*(i_s) = z_s \quad (0 \leq s \leq k, 0 \leq \alpha \leq n), \quad (11)$$

$$\mu^+(l) = \max_{\alpha \in I} \mu_\alpha(l) = \mu^*(l) \quad (0 \leq l \leq n), \quad (12)$$

$$\mu^-(l) = \min_{\alpha \in I} \mu_\alpha(l) = \mu_l(l) \quad (0 \leq l \leq n), \quad (13)$$

причем, если $i_{s-1} \leq l \leq i_s$ ($1 \leq s \leq k$), то

$$\mu_l(l) = \mu^*(l) - [2m/(i_s - i_{s-1})] (i_s - l)(l - i_{s-1}),$$

откуда, учитывая (10), получаем

$$\mu^-(l) = [2m/(i_s - i_{s-1})] (l - a_s^0)^2 - \mu_s^*, \quad (14)$$

где

$$a_s^0 = (i_s + i_{s-1})/2 - (z_s - z_{s-1})/4m, \quad (15)$$

$$\mu_s^* = \left(\frac{2m}{i_s - i_{s-1}} \right) \left[\left(\frac{i_s - i_{s-1}}{2} \right)^2 - \frac{(z_s + z_{s-1})(i_s - i_{s-1})}{4m} + \left(\frac{z_s - z_{s-1}}{4m} \right)^2 \right]. \quad (16)$$

Вид мажоранты $\mu^+(l)$ и миноранты $\mu^-(l)$ для функции регрессии $\mu_\alpha(l)$ изображен на рис. 1.

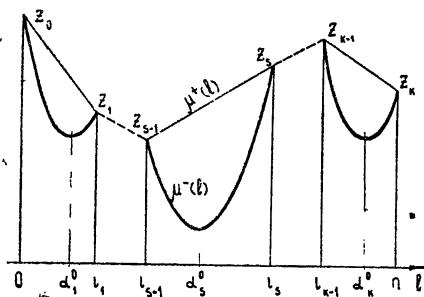


Рис. 1. Мажоранта $\mu^+(l)$ (верхняя, кусочно-линейная кривая) и миноранта $\mu^-(l)$ (нижняя, кусочно-квадратичная кривая) для функции $\mu_\alpha(l)$.

Лемма 2. Пусть $i_{s-1} < \alpha < i_s$. Тогда кривая $\mu_\alpha(l)$, рассматриваемая как функция непрерывного параметра l , $0 \leq l \leq n$, состоит из отрезков прямых, соединяющих точки

$$(0, z_0), (i_1, z_1), \dots, (i_{s-1}, z_{s-1}), (\alpha, \mu^-(\alpha)), (i_s, z_s), \dots, (n, z_k).$$

Доказательство вытекает из (8) — (11) и (13).

Лемма 3.

$$\mu_{\alpha^*}(\alpha^*) = \min_{l \in I} \mu_{\alpha^*}(l).$$

Доказательство. Полагая $\alpha^* \sim a_t^*$, из (6), (8) — (10) определяем

$$\mu_{\alpha^*}(l) = \begin{cases} z_{t-1} - m(l - i_{t-1}) & (i_{t-1} \leq l \leq \alpha^*), \\ z_t + m(l - i_t) & (\alpha^* \leq l \leq i_t) \end{cases}$$

откуда, учитывая лемму 2, получаем, что

$$\min_{l \in I} \mu_{\alpha^*}(l) = \min \{ z_0, z_1, \dots, z_{t-2}, \mu^-(\alpha^*), z_{t+1}, \dots, z_k \}.$$

Остается показать, что

$$\mu^-(\alpha^*) < \min \{ z_0, z_1, \dots, z_{t-2}, z_{t+1}, \dots, z_k \}.$$

Предположим противное, т. е. пусть существует j ($0 \leq j \leq k$) такое, что

$$\mu^-(\alpha^*) > z_j. \quad (17)$$

Определим $\mu^-(\alpha^*)$ из (6) и (13) — (16)

$$\mu^-(\alpha^*) = [z_t + z_{t-1} - m(i_t - i_{t-1})]/2, \quad (18)$$

и, используя неравенство

$$|z_s - z_{s-1}| < m(i_s - i_{s-1}) \quad (1 \leq s \leq k), \quad (19)$$

вытекающее из (2) и (3), получим оценку

$$R(t) < 2m(i_t - i_{t-1}) - 2(z_t + z_{t-1}),$$

которая, учитывая (17) и (18), приводит к соотношению

$$R(t) < -4z_j. \quad (20)$$

Далее, поскольку $dR(j)/dz_{j-1} < 0$, то, принимая во внимание (7) и (19), имеем*

$$R(j) > -4z_j. \quad (21)$$

Из (20) и (21) следует неравенство $R(t) < R(j)$, противоречащее определению 2.

Рассмотренное статистическое описание минимизируемой функции позволяет подойти к задаче выбора испытаний в процессе поиска экстремума с позиций теории статистических решений [6].

ВЫБОР ИСПЫТАНИЙ

Очередное испытание на $k+1$ -м шаге поиска условимся выбирать по правилу Байеса, предполагая, что задана функция потерь $Q(\alpha, l)$, описывающая потери, к которым приводит принятие решения l (т. е. проведение очередного испытания в точке $l \in I$), если природа находится в состоянии $\alpha \in I$ (экстремум находится в точке $\alpha \in I$).

В качестве конкретной функции потерь рассмотрим функцию

$$Q(\alpha, l) = \begin{cases} 0 & (l = \alpha) \\ D & (l \neq \alpha) \end{cases}, \quad (22)$$

где D есть некоторое положительное число. Тогда по принципу Байеса [5] следует принять решение l^* , удовлетворяющее условию

$$P(\omega_k, l^*) = \max_{\alpha \in I} P(\omega_k, \alpha), \quad (23)$$

причем $P(\omega_k, \alpha)$ есть плотность вероятности получения результата ω_k и одновременного пребывания природы в состоянии α .

Теорема 1. $l^* = \alpha^*$, если c достаточно мало.

Доказательство. Поскольку условие (23) эквивалентно условию $\xi(l^* | \omega) \geq \xi(\alpha^* | \omega)$ ($\alpha \in I$), то справедливость теоремы вытекает из леммы 1.

Потери в (22) зависят лишь от попадания испытания в точку экстремума. Другим интересным случаем является зависимость потерь от значения минимизируемой функции в точке испытания. Пусть, например,

$$Q(\alpha, l) = \mu_\alpha(l). \quad (24)$$

В этом случае Байесовское решение l^0 должно удовлетворять условию

$$M(l^0) = \min_{l \in I} M(l), \text{ где } M(l) = \sum_{\alpha \in I} \mu_\alpha(l) \xi(\alpha | \omega).$$

Теорема 2. $l^0 = \alpha^*$, если c достаточно мало.

Доказательство. При $c \rightarrow 0$ $M(l)$, как следует из (4), (5) и леммы 1, определяется членом $\mu_{\alpha^*}(l) \xi(\alpha^* | \omega)$, минимум которого, согласно лемме 3, достигается при $l = \alpha^*$.

* При $j = 0$ вместо $R(j)$ следует рассмотреть $R(j + 1)$.

Следствие 1. При достаточно малом r выбор решения α^* (используемый в алгоритмах [3, 4]) является оптимальным по Байесу для функций потерь из (22) и (24).

УСЛОВИЕ ОСТАНОВКИ

Теорема 3. При $r > 2$ α^* не совпадает с точками i_s ($0 \leq s \leq k$), в которых уже осуществлялись испытания, если $i_t - i_{t-1} > 1$, и $\alpha^* = i_t$ или $\alpha^* = i_{t-1}$, если

$$i_t - i_{t-1} \leq 1. \quad (25)$$

Доказательство следует из (2), (3), (6) и определения 2.

Определение 4. Условие (25) называется условием остановки.

Ниже мы рассмотрим некоторые свойства условия остановки в следующем варианте алгоритма глобального поиска (АГП) [4]*. Первые два испытания осуществляются на концах отрезка $[a, b]$. Каждое следующее испытание проводится в точке

$$x^* = (x_t + x_{t-1})/2 - (z_t - z_{t-1})/2m, \quad (26)$$

где t удовлетворяет условию $R(t) = \max_{1 \leq s \leq k} R(s)$, причем

$$R(s) = m(x_s - x_{s-1}) + \frac{(z_s - z_{s-1})^2}{m(x_s - x_{s-1})} - 2(z_s + z_{s-1}). \quad (27)$$

Поиск прекращается, если после выбора t выполняется неравенство

$$x_t - x_{t-1} \leq \epsilon. \quad (28)$$

Величина m оценивается из (2) при условии, что

$$M = \max_{1 \leq s \leq k} |(z_s - z_{s-1})/(x_s - x_{s-1})|. \quad (29)$$

Исходным пунктом для дальнейших рассуждений является неравенство $R(t) \geq R(s)$ ($1 \leq s \leq k$), определяющее выбор значения t в (26). Перепишем это условие в виде

$$A_s(r) = \pi_s(r) - \pi_t(r) \leq 0 \quad (1 \leq s \leq k), \quad (30)$$

где

$$\pi_s(r) = R(s) = r\Delta_s + \delta_s^2/r\Delta_s - 4\mu_s \quad (1 \leq s \leq k), \quad (31)$$

причем, $\Delta_s = M(x_s - x_{s-1})$, $\delta_s = z_s - z_{s-1}$ и $\mu_s = (z_s + z_{s-1})/2$.

Заметим, что в силу (2) и (29)

$$|\delta_s| < \Delta_s \quad (1 \leq s \leq k), \quad (32)$$

и условие остановки (28) переходит в (33)

$$\Delta_t \leq M\epsilon. \quad (33)$$

Определение 5. $J = \{s, 1 \leq s \leq k : \Delta_s > M\epsilon\}$.

Определение 6. $\alpha = \min_{s \in J} \Delta_s/\Delta_t$.

Теорема 4. Пусть условие (33) выполнено при некотором $r_0 > \sqrt{\alpha/(\alpha - 1)}$. Тогда оно будет выполнено для любого r из интервала

$$\sqrt{\alpha/(\alpha - 1)} \leq r \leq r_0$$

* Этот вариант получается в результате перехода от α^* к $x^* = a + \epsilon\alpha_t^*$ и соответствующей замены i_s ($0 \leq s \leq k$) на x_s в (3), (6), (7), (25).

(при этом $\alpha > 1$), и существует такое достаточно большое $r^0 > r_0$, что при любом $r > r^0$ условие остановки не будет выполняться.

Доказательство. Из (30) — (33) вытекает, что

$$dA_s(r)/dr \geq \Delta_t(\alpha - 1 - \alpha/r^2) \quad (s \in J),$$

и, следовательно, при $r \geq \sqrt{\alpha/(\alpha-1)}$ $dA_s(r)/dr \geq 0$.

Таким образом, условие $A_s(r) \leq 0$ ($s \in J$) справедливо при $\sqrt{\alpha/(\alpha-1)} \leq r \leq r_0$, поскольку $A_s(r_0) \leq 0$.

Далее, из (30) — (33) следует также оценка

$$r_0 \Delta_t(\alpha - 1 - 1/r_0^2) + 4(\mu_t - \mu_s) \leq 0 \quad (s \in J),$$

которая, очевидно, не выполняется при достаточно большом r .

Теорема 5. Если $\Delta_s \geq \beta \Delta_p$, где $\beta > 1$ и $\pi_s(r) \leq \pi_p(r)$, то из

$$r \geq \sqrt{\beta}/(\sqrt{\beta} - 1)$$

следует, что

$$\min(z_{p-1}, z_p) \leq \min(z_{s-1}, z_s).$$

Доказательство. Справедливость теоремы вытекает из оценки $\pi_s(r) - \pi_p(r) \geq r \Delta_p [\beta - 1 - \beta(2 - 1/r)/r] + 4 [\min(z_{p-1}, z_p) - \min(z_{s-1}, z_s)]$, которая выводится из (31), (32) и

$$\mu_s = \min(z_{s-1}, z_s) + |\delta_s|/2 \quad (1 \leq s \leq k).$$

Следствие 2. Если условие остановки выполнено и $r \geq \sqrt{\alpha}/(\sqrt{\alpha}-1)$, то $\min_{s \in J}(z_{t-1}, z_t) \leq \min(z_{s-1}, z_s)$.

Теорема 6. Пусть $K = \gamma M$ ($\gamma > 1$) есть константа Липшица для минимизируемой функции $\varphi(x)$. Тогда, если $\Delta_s \geq \beta \Delta_p$ ($\beta > 1$), $\pi_s(r) \leq \pi_p(r)$ и

$$r \geq \gamma(2\beta + 1)/(\beta - 1), \tag{34}$$

то $\varphi(x) \geq v_s = \mu_s - \gamma \Delta_s/2$, $x \in [x_{s-1}, x_s]$.

Доказательство. Из определения K вытекает, что

$$\varphi(x) \geq v_s = \mu_s - \gamma \Delta_s/2, \quad x \in [x_{s-1}, x_s] \quad (1 \leq s \leq k).$$

Тогда из (31), (32) и условий теоремы следует, что

$$\pi_s(r) - \pi_p(r) \geq 4(v_p - v_s) + \Delta_p [(r - 2\gamma)(\beta - 1) - 1/r],$$

и, учитывая неравенство

$$v_p \geq \min(z_{p-1}, z_p) - \gamma \Delta_p/2,$$

получаем

$$4 [\min(z_{p-1}, z_p) - v_s] + \Delta_p [r(\beta - 1) - 1/r - 2\gamma\beta] \leq 0,$$

откуда при $r^2(\beta - 1) - 2\gamma\beta r - 1 \geq 0$ следует утверждение теоремы. Последнее неравенство выполняется, если

$$r \geq \gamma \left(\frac{\beta}{\beta - 1} \right) \left[1 + \sqrt{1 + \frac{\beta - 1}{\gamma^2 \beta^2}} \right],$$

для чего достаточно выполнения (34).

Следствие 3. Пусть $\gamma = 1$ (т. е. константа Липшица совпадает с M) и условие остановки выполнено. Тогда из

$$r \geq (2\alpha + 1)/(\alpha - 1)$$

следует, что $\min(z_{t-1}, z_t) \leq \min_{x \in [x_{s-1}, x_s]} \varphi(x)$ ($s \in J$).

Графики оценок $\sqrt{\alpha/(\alpha-1)}$, $\sqrt{\alpha}/(\sqrt{\alpha}-1)$ и $(2\alpha+1)/(\alpha-1)$ изображены на рис. 2 (соответственно кривые 1, 2 и 3).

Полученные результаты позволяют судить о характере поиска и о его надежности в зависимости от значения r . Для иллюстрации рассмотрим любой интервал $[x_{s-1}, x_s]$, длина которого, например, в 4 раза превышает ε после выполнения условия остановки. Тогда, как следует из теорем этого параграфа и рис. 2, при $r > 2 \min[\varphi(x_{s-1}), \varphi(x_s)]$ больше минимального значения функции на концах интервала остановки, а при $r > 6$ (в предположении, что константа Липшица превышает M не более, чем вдвое) значение $\varphi(x)$ в любой точке $x \in [x_{s-1}, x_s]$ также пре-

восходит наименьшее из двух (уже измеренных) значений функции на концах интервала остановки. Наконец, если α (из определения 6) равно 4, то условие остановки будет выполняться при уменьшении r (от значения, использованного в ходе поиска) до 1,15.

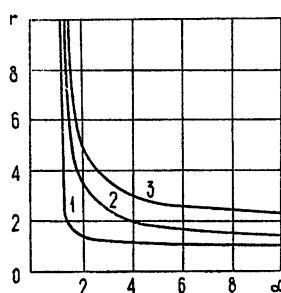


Рис. 2. Графики оценок: $\sqrt{\alpha/(\alpha-1)}$ (кривая 1), $\sqrt{\alpha}/(\sqrt{\alpha}-1)$ (кривая 2) и $(2\alpha+1)/(\alpha-1)$ (кривая 3).

Теорема 7. Максимальная длина интервала $\Delta_p/M = \max_{1 \leq s \leq k} (x_s - x_{s-1})$,

соответствующая окончанию поиска, убывает при увеличении r .

Доказательство. Пусть $\Delta_p = \beta M \varepsilon$. Тогда из (30) — (33) следует

$$r \Delta_t (\beta - 1) + [(\delta_p^2/\Delta_p - \delta_t^2/\Delta_t)/r + 4(\mu_t - \mu_p)] \leq 0,$$

где член в квадратных скобках может быть заменен нижней оценкой

$$- 4M \max [(x_t - x_{p-1}), (x_p - x_{t-1})] - \Delta_t/r.$$

Отсюда вытекает, что выполнение условия остановки невозможно при нарушении неравенства

$$\beta \leq \{ 4M \max [(x_t - x_{p-1}), (x_p - x_{t-1})] + \Delta_t(r + 1/r) \}/r \Delta_t.$$

Следствие 4. С ростом r увеличивается число испытаний предшествующих выполнению условия остановки.

Таким образом, возрастание надежности поиска при увеличении r , отмеченное ранее, сопровождается повышением затрат на поиск.

МОДИФИКАЦИЯ АЛГОРИТМА

В основе алгоритма (2), (26) — (29) лежит статистическая модель минимизируемой функции, содержащая условие (1). Согласно этому предположению, среднее значение модуля первой разности функции одинаково в различных точках области определения. Однако для гладких функций, глобальный минимум которых лежит внутри области, будут характерны малые (по сравнению с оценкой (29)) значения первой разности в окрестности экстремума. Неучет этого обстоятельства в принятой модели вызывает увеличение числа поисковых испытаний в окрест-

ности экстремума (фактически приводит к сканированию этой окрестности) при достаточно малом ϵ [4].

Возможное уточнение модели достигается тем, что алгоритм (2), (26)–(29) применяется не к самой функции $\varphi(x)$, а к ее преобразованию

$$\varphi^0(x) = f(\varphi(x) - \varphi^*), \quad (35)$$

где $\varphi^* = \min_{x \in [a, b]} \varphi(x)$, и f — выпуклая функция, причем $f(0) = 0$. Заметим, что точки минимума функций φ и φ^0 совпадают. Однако φ^* , входящее в (35), заранее обычно неизвестно. Поэтому вместо преобразования φ^0 можно использовать на каждом $k+1$ -м шаге поиска преобразование

$$\varphi_k^0(x) = f(\varphi(x) - \varphi_k^*), \quad (36)$$

где $\varphi_k^0 = \min_{0 \leq s \leq k} \varphi(x_s)$. Преобразование (36) значений $\varphi(x)$ целесообразно лишь в случае, когда экстремум находится не на границе области. Поэтому подстановка в (26)–(29) на $k+1$ -м шаге значений

$$z_s = f(\varphi(x_s) - \varphi_k^*) \quad (0 \leq s \leq k)$$

осуществляется лишь в случае, когда

$$\varphi_k^* \neq \varphi(x_0) = \varphi(a) \text{ и } \varphi_k^* \neq \varphi(x_k) = \varphi(b).$$

В противном случае, используются значения

$$z_s = \varphi(x_s) \quad (0 \leq s \leq k).$$

Для сравнения эффективности исходного алгоритма глобального поиска (АГП) и модифицированного (при $f(z) = \sqrt{z}$) варианта (МАГП) эти алгоритмы были применены для минимизации различных выборок по 100 функций, полученных случайным образом с помощью выражения

$$\varphi(x) = a_0 + \sum_{l=1}^N \left[a_l \sin\left(\frac{l\pi}{2}x\right) + b_l \cos\left(\frac{l\pi}{2}x\right) \right] \quad (x \in [0, 1]),$$

где

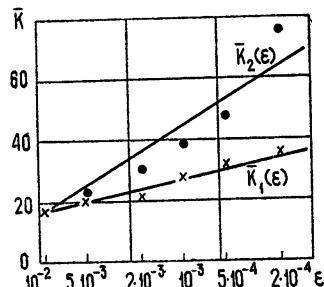
$$-1 \leq a_0, a_l, b_l \leq 1, \quad 1 \leq l \leq N, \quad 4 \leq N \leq 14,$$

причем коэффициенты и само число гармоник N — независимые, равномерно распределенные случайные величины (N — целочисленна).

Выборочное среднее число $\bar{K}_1(\epsilon)$ поисковых испытаний при минимизации с помощью МАГП изображено крестиками на рис. 3 (для различных точностей ϵ). Средняя эффективность $\bar{K}_2(\epsilon)$ исходного варианта АГП на тех же выборках также показана на рис. 3 (темные кружки). Средние значения $\bar{K}_1(\epsilon)$ и $\bar{K}_2(\epsilon)$ для различных точностей ϵ приведены на рис. 3.

Представленные на рис. 3 экспериментальные данные показывают, что с ростом точности модифицированный алгоритм (при $f(z) = \sqrt{z}$) становится заметно эффективнее исходного.

Рис. 3. Выборочное среднее число испытаний при минимизации функций с помощью МАГП ($\bar{K}_1(\epsilon)$) и АГП ($\bar{K}_2(\epsilon)$) как функция точности ϵ .



ЛИТЕРАТУРА

- 1 Ю. И. Неймарк, Р. Г. Стронгин, Изв. АН СССР, Техническая кибернетика, № 1, 17 (1966).
- 2 Р. Г. Стронгин, Изв. АН СССР, Техническая кибернетика, № 6, 118 (1969).
- 3 Р. Г. Стронгин, Автоматика и телемеханика, 31, № 7, 63 (1970).
- 4 Р. Г. Стронгин, Изв. высш. уч. зав. — Радиофизика, 13, № 4, 539 (1970).
- 5 Д. Блекуэлл, М. А. Гиршик, Теория игр и статистических решений, ИЛ, М, 1958.

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
8 июля 1970 г.

DECISION AND STOP RULES IN ONE-DIMENSIONAL GLOBAL SEARCH

R. G. Strongin

Decision and stop rules' properties in the one-dimensional absolute minimum seeking algorithm [3, 4] are investigated. Modified version of the original algorithm is suggested and the search speed up due to this modification is illustrated.

УДК 62—504 + 518.5

НЕКОТОРЫЕ КЛАССЫ ДИСКРЕТНЫХ УПРАВЛЯЕМЫХ СИСТЕМ И ЯЗЫКИ (АЛГОРИТМИЧЕСКИЕ ПРОБЛЕМЫ)

A. A. Бедина, Д. И. Коган

Вводятся некоторые классы управляемых векторных систем. В рамках этих понятий может быть сформулирован ряд задач дискретного управления. Исследуются вопросы разрешимости проблемы перевода системы из стартового состояния в предписанное для различных классов систем. Вводится и исследуется класс Δ -языков; каждый такой язык — множество стратегий некоторой системы, переводящих стартовый вектор в финальный вектор или финальное подмножество векторов. Понятие Δ -языка оказывается обобщением понятия регулярного множества.

Будем рассматривать ряд конкретизаций введенного в [1] понятия управляемой векторной системы (ΛR -системы, Δ -системы, Δ -det-системы). В рамках таких конкретизаций можно сформулировать ряд проблем дискретного управления, в частности, теории расписаний. Последовательности управления в системе (стратегии системы) — слова некоторого алфавита (алфавита управляющих воздействий). Множество стратегий, переводящих стартовый вектор S в финальный вектор F или в один из векторов финального подмножества Φ (задаваемого формулой арифметики Пресбургера [2], раздел 1), есть язык (в смысле Хомского), определяемый самой системой, вектором S и вектором F (подмножеством Φ). Разрешимость проблемы непустоты языков такого типа эквивалентна разрешимости проблемы осуществимости перевода системы из стартового состояния в предписанное. Результаты проведенного исследования сформулированы теоремами 1 и 2 (разделы 1, 2). Введенный и частично исследованный в разделе 1 класс Δ -языков подробно изучается в разделе 3. В разделе 4 рассматриваются некоторые подклассы класса Δ -языков и связанные с этими подклассами алгоритмические проблемы. Далее вводятся языки, определяемые автоматами со счетчиками с конечным числом занулений (раздел 5). Эти языки являются подклассом Δ -языков. Вопросы соотношения между Δ - и бесконтекстными языками рассмотрены в разделе 6.

1. Пусть D — множество всех действительных, I — целых, N — натуральных чисел, $n \in N$; M — выпуклое многогранное множество в пространстве D^n , координаты всех его точек неотрицательны; K — конечное множество векторов из I^n , S — вектор, принадлежащий $M \cap N^n$. Последовательности $S = S_0, S_1, \dots, S_r$ векторов из $M \cap N^n$, такие, что для $i = 1, \dots, r$ разность $S_{i+1} - S_i$ либо принадлежит K , либо равна нулю, назовем KS — траекториями в M . Множество последних векторов всех KS -траекторий в M обозначим $V(M, K, S)$.

Введем граф Γ_K с вершинами p_1, p_2, \dots, p_n . Граф Γ_K содержит дугу $p_i p_j$ тогда и только тогда, когда в K имеется вектор с отрицательной i -й и положительной j -й координатами; K назовем z -множеством ($K \in Z$), если Γ_K не содержит циклов,

Лемма 1. Для любого множества $V(M, K, S)$, $K \in Z$, существует формула арифметики Пресбургера Φ_V , такая, что $\bar{Y} = (y_1, y_2, \dots, y_n)$ принадлежит $V(M, K, S)$ тогда и только тогда, когда $\Phi_V(y_1, y_2, \dots, y_n)$ имеет место.

ΔR -системой ($\Delta^z R$ -системой, если $K \in Z$) назовем совокупность $\Sigma = \langle n, K, K_1, \dots, K_m, A, R \rangle$, где $n \in N$ (размерность системы); K —конечное множество ненулевых векторов из I^n (векторы управлений); K_1, \dots, K_m — множества векторов такие, что $\bigcup_{i=1}^m K_i = K$; A — алфавит (множество управляемых воздействий); R — регулярное множество слов из алфавита A (стратегии системы), $A = \{a_1, a_2, \dots, a_m\}$.

Пусть $\alpha = a_{i_1} a_{i_2} \dots a_{i_p}$ — стратегия и $S = S_0, S_1, \dots, S_p$ — KS -траектория в первом ортанте пространства D^n . Назовем эту траекторию S -реализацией стратегии α в системе Σ , если $S_d = S_{d-1} \in K_d$ при $d = 1, 2, \dots, p$. Последний вектор реализации будем называть результатом реализации. Если для стратегии α векторов S и T существует S -реализация α с результатом T , то будем говорить, что стратегия α переводит вектор S в вектор T .

Множество всех слов алфавита A обозначим A^* ; ΔR -систему назовем Δ -системой, если $R = A^*$; Δ -систему назовем Δ - U -det-системой ($U \subseteq N^n$), если для каждой стратегии при $S \in U$ может существовать лишь одна ее S -реализация; Δ - N^n -det-системы будем называть Δ -det-системами. Как легко видеть, необходимым и достаточным условием для того, чтобы Δ -система была Δ -det-системой, является однозначность множеств K_1, K_2, \dots, K_m .

Пусть Φ — формула арифметики Пресбургера от n переменных (далее — n -ф. а. П.) и вектор $S \in N^n$. Обозначим $P(\Sigma, S, \Phi)$ множество стратегий системы Σ , переводящих вектор S в векторы, координаты которых удовлетворяют формуле Φ . Множество слов $X \subseteq A^*$ будем называть Δ -языком (Δ^z -языком), если существуют Δ -система Σ (Δ^z -система Σ) вектор S и ф. а. П. Φ , такие, что $X = P(\Sigma, S, \Phi)$.

Теорема 1. Разрешима проблема определения по тройке Σ, S, Φ , является ли пустым Δ^z -язык, к которому определяемый. Эта теорема — непосредственное следствие леммы 1 и разрешимости арифметики Пресбургера.

Лемма 2. Пусть $X_1 = P(\Sigma_1, S_1, \Phi_1)$ и $X_2 = P(\Sigma_2, S_2, \Phi_2)$ — Δ -языки. Тогда множества $X_1 \cup X_2, X_1 \cap X_2$ также являются Δ -языками.

По $\Sigma_1 = \langle p, K', K'_1, \dots, K'_v, A, A^* \rangle$ и $\Sigma_2 = \langle q, K^2, K^2_1, \dots, K^2_w, B, B^* \rangle$ следующим образом построим систему Σ . Размерность системы полагается равной $p + q + 2$. Предполагаем, что алфавиты $A = \{a_1, a_2, \dots, a_v\}$ и $B = \{b_1, b_2, \dots, b_w\}$ упорядочены так, что только первые t символов их совпадают, $A \cap B$ есть множество $\{a_1, a_2, \dots, a_t\}$. Определим K_j ($j = 1, 2, \dots, t$) как множества векторов, имеющих следующую структуру: вектор $(x_1, x_2, \dots, x_{p+q+2})$ принадлежит K_j тогда и только тогда, когда а) $(x_1, x_2, \dots, x_p) \in K^1_j$; б) $(x_{p+1}, x_{p+2}, \dots, x_{p+q}) \in K^2_j$; в) $x_{p+q+1} = 0, x_{p+q+2} = 0$. Далее, K_j ($j = t + 1, \dots, v$) — множества векторов следующей структуры: вектор $(x_1, x_2, \dots, x_{p+q+2})$, принадлежит K_j тогда и только тогда, когда а) $(x_1, x_2, \dots, x_p) \in K^1_j$; б) $x_{p+1} = x_{p+2} = \dots = x_{p+q} = 0$; в) $x_{p+q+1} = 1, x_{p+q+2} = 0$. Наконец, K_j ($j = v + 1, \dots, v + w - t$) — множества векторов такой структуры: вектор $(x_1, x_2, \dots, x_{p+q+2})$ принадлежит K_j тогда и только тогда, когда а) $x_1 = x_2 = \dots = x_{p+q} = 0$; б) $x_{p+1}, \dots, x_{p+q} \in K^2_j$; в) $x_{p+q+1} = 0, x_{p+q+2} = 1$; $K = \bigcup_{j=1}^t K'_j$. Алфавит $C(v + w - t)$ — буквенный, $c_1 = a_1 = b_1, c_2 = a_2 = b_2, \dots, c_t = a_t = b_t, c_{t+1} = a_{t+1}, \dots, c_v = a_v, c_{v+1} = b_{t+1}, \dots, c_{v+w-t} = b_w$.

Система Σ есть совокупность $\langle p+q, K, K_1, \dots, K_{v+w-t}, C, C^* \rangle$. Пусть $S = (p+q+2)$ -мерный вектор, первые p координат которого совпадают (соответственно) с координатами вектора S_1 , а следующие q также соответственно совпадают с координатами вектора S_2 . Последние две координаты вектора S полагаются нулевыми. Положим

$$\Psi_1(x_1, \dots, x_{p+q+2}) = [\Phi_1(x_1, \dots, x_p) \vee \Phi_2(x_{p+1}, \dots, x_{p+q})]$$

$$\wedge [(x_{p+q+1} = 0) \vee (x_{p+q+2} = 0)];$$

$$\Psi_2(x_1, \dots, x_{p+q+2}) = [\Phi_1(x_1, \dots, x_p) \wedge \Phi_2(x_{p+1}, \dots, x_{p+q})]$$

$$\wedge [(x_{p+q+1} = 0) \wedge (x_{p+q+2} = 0)].$$

Как легко видеть, $X_1 \cup X_2 = P(\Sigma, S, \Psi_1)$ и $X_1 \cap X_2 = P(\Sigma, S, \Psi_2)$.

Лемма 3. Всякое регулярное множество слов есть Δ -язык.

Пусть регулярное множество R выделяется конечным автоматом $U = \langle A, Q, f, q_1, F \rangle$, где $A = \{a_1, a_2, \dots, a_m\}$ — входной алфавит автомата ($R \subseteq A^*$); $Q = \{q_1, q_2, \dots, q_n\}$ — множество внутренних состояний; f — отображение из $A \times Q$ в Q (функция переходов автомата); $q_1 \in Q$ — начальное состояние автомата, $F \subseteq Q$ (выделенное подмножество состояний). Не теряя общности рассуждений, полагаем, что граф автомата U не содержит петель; K_j ($j = 1, 2, \dots, m$) — множества n -мерных векторов, определяемые следующим образом: вектор (x_1, x_2, \dots, x_n) принадлежит K_j тогда и только тогда, когда а) только две координаты вектора ненулевые, одна из них равна плюс, а другая — минус единице, б) $[(x_s = 1) \wedge (x_t = -1)] \leftrightarrow q_s = f(a_j, q_t)$ ($s, t = 1, 2, \dots, n$).

Пусть $K = \bigcup_{j=1}^m K_j$, рассматриваем систему $\Sigma = \langle n, K, K_1, \dots, K_m, A, A^* \rangle$. Положим $S = (1, 0, 0, \dots, 0)$ и $\Phi(x_1, \dots, x_n) = [(x_{i_1} = 1) \vee \dots \vee (x_{i_k} = 1)]$, где i_1, i_2, \dots, i_k — полный перечень номеров состояний из F . Тогда слово $\alpha \in A^*$ переводит автомат U из начального состояния в состояние, скажем, q_z тогда и только тогда, когда S -реализация стратегии α в системе Σ имеет своим результатом вектор, у которого только z -ая координата отлична от нуля, она равна единице. Поэтому $R = P(\Sigma, S, \Phi)$.

Множество слов $X \subseteq A^*$ будем называть ΔR -языком, если существует ΔR -система Σ , вектор S и формула а.П. Φ , такие, что $X = P(\Sigma, S, \Phi)$.

Лемма 4. Всякий ΔR -язык является Δ -языком.

Пусть ΔR -язык X определяется системой $\Sigma = \langle n, K, K_1, K_2, \dots, K_m, A, R \rangle$, вектором S и ф. а. П. Φ . Положим $\Sigma^1 = \langle n, K, K_1, \dots, K_m, A, A^* \rangle$. Очевидно, что пересечение множеств $P(\Sigma^1, S, \Phi)$ и R есть X . Но как $P(\Sigma^1, S, \Phi)$, так и R являются Δ -языками (см. лемму 3). Из леммы 2 следует, что X является Δ -языком.

Проблема непустоты для Δ -языков остается открытой. Как можно показать, она эквивалентна известной проблеме непустоты для бесконечных языков с регулярным ограничением на последовательность применений правил вывода.

2. Пусть $\Sigma = \langle n, K, K_1, \dots, K_m, A, R \rangle$ — произвольная ΔR -система; $\alpha = a_{i_1} a_{i_2}, \dots, a_{i_k}$ — одна из стратегий этой системы; KS -траекторию $S = S_0, S_1, S_2, \dots, S_k$ в первом ортантне D^n назовем S -квази-реализацией стратегии α , если а) разность $S_d - S_{d-1}$ ($d = 1, 2, \dots, k$) равна нулю тогда и только тогда, когда сумма вектора S_{d-1} с любым вектором из множества K_{i_d} есть вектор, не принадлежащий N^n ; б) из

того, что разность $S_d - S_{d-1}$ отлична от нуля, следует ее принадлежность множеству K_{ℓ_d} . Последний вектор квазиреализации назовем результатом квазиреализации. Пусть Φ — ф. а. П. от n переменных и вектор $S \in N^n$. Обозначим $Q(\Sigma, S, \Phi)$ множество стратегий, для которых существуют такие их S -квазиреализации, что координаты векторов-результатов удовлетворяют формуле Φ . Пусть для вектора $D \in N^n$ множество $Q(\Sigma, S, D)$ есть совокупность таких стратегий, что для каждой существует S -квазиреализация с результатом D .

Теорема 2. Существуют ΔR -система Σ и вектор $S \in N^n$, такие, что проблема существования стратегий Σ , квазиреализации которых переводят вектор S в вектор $D \in N^n$ (т. е. проблема непустоты множеств $Q(\Sigma, S, D)$) является неразрешимой.

Пусть Π — конечный автомат с двумя счетчиками (см. [1] или [3]), работающий в однобуквенном входном алфавите [1]. Выделено одно состояние этого автомата, $U_\Pi(r)$ — множество входных слов, переводящих автомат из начального в выделенное состояние, причем перед началом работы над словом первый счетчик автомата содержит r единиц, второй счетчик пуст, а после окончания работы оба счетчика пусты. Будем считать, что проблема непустоты множеств $U_\Pi(r)$ неразрешима (существование автомата с таким свойством следует из [1] или теоремы 2, [3], § 15.2). Мы покажем, что по автомату Π строятся ΔR -система Σ^0 , вектор D^0 и эффективным образом вычислимая функция $S = S(r)$, такие, что $U_\Pi(r)$ пусто тогда и только тогда, когда пусто $Q(\Sigma^0, S(r), D^0)$. Это и будет доказательством теоремы, так как проблема непустоты $Q(\Sigma^0, S(r), D^0)$ при фиксированном S и проблема непустоты $Q(\Sigma^0, S(r), D^0)$ при фиксированном D^0 одновременно разрешимы или неразрешимы.

Пусть Π имеет k состояний, q_1 — начальное, а q_k — выделенное состояние. Тогда система Σ^0 будет иметь размерность $k+6$, вектор $S(r)$ имеет своей первой координатой единицу, 2-я, 3-я, ..., k -я координаты равны нулю, $(k+1)$ -я — число, $(k+2)$ -я — нуль, $(k+3)$ -я и $(k+4)$ -я — единица, $(k+5)$ -я и $(k+6)$ -я — нуль. У вектора D^0 первые $k-1$ координаты — нули, k -я координата равна единице, $(k+1)$ -я и $(k+2)$ -я — нули, $(k+3)$ -я и $(k+4)$ -я — единицы, $(k+5)$ -я и $(k+6)$ -я — нули. Как будет видно дальше, первые k координат вектора из квазиреализации характеризуют состояние автомата, $(k+1)$ -я и $(k+2)$ -я показывают содержание его счетчиков, $(k+3)$ -я — $(k+6)$ -я служат для определения пустоты или непустоты счетчиков. Алфавит A системы Σ^0 четырехбуквенный, $A = \{a, b, 1, c\}$. Множество R имеет следующую структуру: $R = (ab1c)^*$. Здесь символ $*$ означает итерацию ([1], гл. 2, § 3). Букве a соответствует один вектор a , его $(k+1)$ -я и $(k+3)$ -я координаты равны минус единице, а $(k+5)$ -я — плюс единице. Букве b соответствует также единственный вектор b , у которого $(k+2)$ -я и $(k+4)$ -я координаты равны минус единице, а $(k+6)$ -я — плюс единице. Все остальные координаты векторов a и b нулевые. Пусть у вектора S ($S \in N^{k+6}$) $(k+3)$ -я и $(k+4)$ -я координаты равны единице, а $(k+5)$ -я и $(k+6)$ -я — нулю. Тогда по последним четырем координатам вектора, являющегося результатом S -квазиреализации стратегии ab , можно ответить на вопрос о том, какими были $(k+1)$ -я и $(k+2)$ -я координаты S . Именно, если они были нулевыми, то у вектора-результата $(k+3)$ -я и $(k+4)$ -я координаты равны единице, а $(k+5)$ -я и $(k+6)$ -я — нулю. Если s_{k+1} отлично от нуля, а $s_{k+2} = 0$, то у вектора-результата $(k+3)$ -я и $(k+6)$ -я координаты равны нулю, а $(k+4)$ -я и $(k+5)$ -я равны плюс единице. Если $s_{k+1} = 0$, $s_{k+2} \neq 0$, то $(k+3)$ -я и $(k+6)$ -я координаты вектора-результата равны единице, а $(k+4)$ -я и $(k+5)$ -я равны нулю. И, наконец, если $s_{k+1} \neq 0$, $s_{k+2} \neq 0$, то у вектора-результата последние две координаты равны единице, а предпоследняя пара координат — нули,

Букве I соответствуют $4k$ векторов, l_1, l_2, \dots, l_{4k} . У первого, пятого, ..., $(4k-3)$ -го векторов $(k+3), (k+4), (k+5), (k+6)$ -я координаты равны соответственно $-1, -1, 0, 0$. У второго, шестого, ..., $(4k-2)$ -го эти же координаты равны соответственно $-1, 0, 0, -1$. У третьего, седьмого, ..., $(4k-1)$ -го $-0, -1, -1, 0$. У четвертого, восьмого, ..., $4k$ -го $-0, 0, -1, -1$. Будем говорить, что счетчики автомата обладают первым качеством, если оба они пусты; вторым качеством, если пуст только первый; третьим качеством, если пуст только второй; четвертым качеством, если оба счетчика непусты. Если автомат Π , счетчики которого обладают p -м качеством ($p = 1, 2, 3, 4$), под действием входного символа переходит из i -го состояния в j -е, то i -я координата вектора l_{4i+p} равна минус единице, а j -я — плюс единице. Остальные из первых k координат нулевые ($i = 1, 2, \dots, k$). Если при этом автомат меняет содержание первого счетчика на δ_1 , а второго на δ_2 (δ_1 и δ_2 равны $0, -1$ или $+1$), то в векторе l_{4i+p} $(k+1)$ -я и $(k+2)$ -я координаты равны соответственно $\delta_1 + \varepsilon_p^1$ и $\delta_2 + \varepsilon_p^2$, где $\varepsilon_p^1 = 1$ при $p = 3, 4$; $\varepsilon_p^2 = 1$ при $p = 2, 4$; $\varepsilon_1^1 = \varepsilon_2^1 = \varepsilon_1^2 = \varepsilon_2^2 = 0$.

Букве c соответствует вектор с равными плюс единице $(k+3)$ -й и $(k+4)$ -й координатами. Без затруднений проверяется, что построенные ΔR -система Σ^0 и функция $S(r)$ таковы, что выполняются следующие условия: а) каждой стратегии отвечает единственная $S(r)$ -квазиреализация, б) если автомат Π , начавший работу, имеет r единиц в первом счетчике и пустой второй счетчик, то оказывается, что по z тактов работы автомат Π в состоянии q , имеет содержимое счетчиков s_1 и s_2 тогда и только тогда, когда результат $S(r)$ — квазиреализации стратегии $(ab1c)^z$ есть вектор, имеющий своей i -й координатой единицу, $(k+1)$ -й — s_1 , $(k+2)$ -й — s_2 , $(k+3)$ -й и $(k+4)$ -й единицы, остальные координаты нулевые. Таким образом, $U_\Pi(r)$ пусто тогда и только тогда, когда пусто $Q(\Sigma^0, S(r), D^0)$, где D^0 — вектор с равными единице k -й, $(k+3)$ -й, $(k+4)$ -й координатами, остальные координаты равны нулю. Доказательство закончено.

3. Пусть $A = \{a_1, a_2, \dots, a_m\}$ — алфавит. Введем функции $s_1(a), s_2(a), \dots, s_m(a)$, определенные на словах этого алфавита; $s_i(a)$ — количество вхождений буквы a_i в слово a . Если $L \subseteq A^*$, то $S(L)$ — множество m -мерных векторов, такое, что $(i_1, i_2, \dots, i_m) \in S(L)$ тогда и только тогда, когда существует $a \in L$ и $s_1(a)=i_1, s_2(a)=i_2, \dots, s_m(a)=i_m$. Множество векторов называется полулинейным, если существует ф. а. Π , его определяющая (подробно см. [2]).

Теорема 3. Пусть L — Δ^z -язык. Тогда множество $S(L)$ полулинейно. Положим, что $L = P(\Sigma, S, \Phi)$, где $\Sigma = \langle n, K, K_1, \dots, K_m, A, A^* \rangle$. Введем в рассмотрение $(n+m)$ -мерную систему $\Sigma^1 = \langle n+m, K^1, K_1^1, \dots, K_m^1, A, A^* \rangle$. Множества K_i^1 ($i = 1, 2, \dots, m$) определяются так: вектор $\{b_1, \dots, b_{m+n}\} \in K_i^1$ тогда и только тогда, когда вектор $(b_1, b_2, \dots, b_n) \in K_i$, b_{n+t} ($t = 1, \dots, m$) отлична от нуля и равно единице лишь при $t = i$ ($i = 1, 2, \dots, m$); $K^1 = \bigcup_{i=1}^m K_i^1$. Пусть S^1 — вектор из N^{n+m} , первые n координат которого совпадают соответственно с координатами вектора S , а последние m равны нулю. Пусть M^1 — первый ортант пространства D^{n+m} ; $\Phi_V(x_1, x_2, \dots, x_{n+m})$ — ф. а. Π , такая, что $X^0 = \{x_1^0, x_2^0, \dots, x_{n+m}^0\} \in V(M^1, K^1, S)$ тогда и только тогда, когда имеет место $\Phi_V(x_1^0, x_2^0, \dots, x_{n+m}^0) \in E$ (см. лемму 1). Рассмотрим ф. а. Π . $D(x_{n+1}, \dots, x_{n+m}) \leftrightarrow \exists x_1 \exists x_2, \dots, \exists x_n [\Phi_V(x_1, \dots, x_{n+m}) \wedge \Phi(x_1, \dots, x_n)]$. Покажем, что она задает $S(L)$.

Пусть для $Y = (y_1, y_2, \dots, y_m)$ имеет место $D(y_1, y_2, \dots, y_m)$, т. е. существуют x_1, x_2, \dots, x_n , такие, что $\Phi_V(x_1, \dots, x_n, y_1, \dots, y_m) \wedge \Phi(x_1, \dots, x_n)$.

Рассмотрим K^1S^1 -траекторию в M^1 , результатом которой является вектор $(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m)$. Пусть это траектория $T: S^1 = Z_0, Z_1, Z_2, \dots, Z_p = (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m)$. Положим $A_i = Z_i - Z_{i-1}$ ($i = 1, 2, \dots, p$).

Рассмотрим слово $\alpha = a_{i_1}a_{i_2} \dots a_{i_p}$, связанное с векторами A_i следующим образом: $A_i \in K^1_{ii}$. По траектории T это слово определяется однозначно. Оно принадлежит языку L , и выполняются равенства $S_1(\alpha) = y_1, S_2(\alpha) = y_2, \dots, S_m(\alpha) = y_m$. Пусть, обратно, $\alpha \in L$. Тогда существует S^1 -реализация стратегии α в системе Σ^1 с результатом, координаты которого удовлетворяют формулам $\Phi(x_1, x_2, \dots, x_n)$ и $\Phi_V(x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_{n+m})$, а следовательно, и формуле $D(x_{n+1}, x_{n+2}, \dots, x_{n+m})$. Здесь, как видно из структуры системы Σ^1 , x_{n+i} — количество вхождений символа a_i в стратегию α ($i = 1, 2, \dots, m$). Теорема доказана.

Множество U слов алфавита A назовем универсальным, если $U = A^*$.

Теорема 4. Существуют Δ -система Σ и ф. а. П. Φ , такие, что проблема универсальности множеств $P(\Sigma, S, \Phi)$ является неразрешимой.

Пусть Π — автомат с двумя счетчиками, рассматривавшийся при доказательстве теоремы 2. Слово в алфавите $B = \{a, b, l_{aa}, l_{ab}, l_{ba}, l_{bb}\}$ назовем r -описывающим автомат Π , а) если оно принадлежит регулярному множеству $\{aal_{aa}, abl_{ab}, bal_{ba}, bbl_{bb}\}^*$, б) если автомат Π , начав работать в состоянии q_1 , имея число r содержимым первого счетчика и пустой второй счетчик, перед началом k -го такта работы имеет пустой первый счетчик (нет пустой первый счетчик), то $(3k-2)$ буква слова α есть a (есть b). Таким же образом определяется связь между содержимым второго счетчика и $(3k-1)$ -й буквой. Обозначим через $E(r)$ множество слов α , таких, что а) α — r -описывающее для Π , б) если $l(\alpha)$ — количество букв в слове α , то Π , начав работу в состоянии q_1 , имея число r в первом и нуль во втором счетчике, после $l(\alpha)/3$ тактов работы приходит в выделенное состояние. Может быть построена Δ -система Σ , эффективно вычислимая функция $S = S(r)$ и ф. а. П. Φ , такие, что $\alpha \in B^*$ принадлежит $P(\Sigma, S(r), \Phi) \iff \alpha \in E(r)$. Итак, множество $P(\Sigma, S(r), \Phi)$ совпадает с B^* тогда и только тогда, когда $E(r)$ пусто. Разрешимость проблемы универсальности для Δ -языков означала бы разрешимость проблемы пустоты для множеств $U(\Pi, r)$.

Пусть W_1 и W_2 — множество слов алфавита A . Конкатенацией этих множеств назовем множество

$$W = \{\alpha | (\alpha = \beta\gamma) \wedge (B \in W_1) \wedge (\gamma \in W_2)\}.$$

Лемма 5. Класс Δ -языков замкнут относительно операции конкатенации.

Теорема 5. Пусть P — свойство, определенное на множестве Δ -языков и такое, что а) P должно хотя бы для одного Δ -языка, б) P истинно для всех регулярных множеств, в) P сохраняется при инверсии gsm — отображения (отображения, осуществляющего обобщенную последовательной машиной, определение см. [5]), объединении с пустым множеством и пересечении с пустым множеством и пересечении с любым регулярным множеством. Тогда неразрешима проблема определения по тройке Σ, S, Φ , обладает ли определяемый ею Δ -язык свойством P .

Данная теорема является следствием теоремы 1 из работы [6] вместе с леммами 2, 3, 5 и теоремой 4 настоящей работы.

В качестве иллюстрации применения теоремы приведем следующее следствие.

Следствие. Неразрешима проблема определения по Δ -языку, является ли он регулярным множеством.

Назовем Δ -язык X n — Δ -языком, если $X = P(\Sigma, S, \Phi)$, где Σ — n -мерная Δ -система.

Теорема 6. Для любого натурального n существует Δ -язык L^{n+1} , являющийся $(n+1)$ - Δ -языком и не являющийся n - Δ -языком.

Пусть A_n — n -буквенный алфавит; $\alpha \in A_n^*$; $s_1(\alpha), s_2(\alpha), \dots, s_n(\alpha)$ —количество вхождений букв a_1, a_2, \dots, a_n (соответственно) в слово α . Пусть $X^n = \{\alpha \mid (\alpha \in A_n^*) \wedge s_1(\alpha) = s_2(\alpha) = \dots = s_n(\alpha)\}$. Методом индукции несложно показать, что языки X^n являются языками, существование которых утверждается теоремой.

4. Если Δ -язык L есть $P(\Sigma, S, \Phi)$, где Σ — Δ - U -det-система и $S \in U'$ то будем называть L Δ - d -языком.

Если Δ -язык L есть $P(\Sigma, S, \Phi)$, где Σ — Δ -det-система, мы будем называть L Δ -det-языком.

Теорема 7. Проблема универсальности для Δ - d -языков является разрешимой.

Доказательство теоремы основывается на возможности записи условия универсальности языков такого рода в рамках арифметики Пресбургера.

Из доказательства леммы 3 легко усмотреть, что всякое регулярное множество является Δ - d -языком. Вместе с тем имеет место следующая теорема.

Теорема 8. Существуют регулярные множества, не являющиеся Δ -det-языками.

Пусть $A = \{a, b\}$ —алфавит. Рассмотрим регулярное множество слов $X = \{\alpha \mid (\alpha \in A^*) \wedge (\alpha = a_0 a)\}$. Предположим, что система $\Sigma = \langle n, K, K_1, K_2, A, A^* \rangle$, вектор S и ф. а. $\Pi \Phi$ таковы, что $X = P(\Sigma, S, \Phi)$. Совокупности K_1, K_2 полагаются одноэлементными. Для любой стратегии α в системе существует ее S -реализация, так как по предположению она существует для стратегии αa . Поэтому для двух стратегий $\alpha_1 = ab$ и $\alpha_2 = ba$ существуют S -реализации, при этом векторы-результаты этих реализаций совпадают, их координаты одновременно удовлетворяют или не удовлетворяют формуле Φ ; стратегии α_1 и α_2 одновременно принадлежат или не принадлежат X . Но ведь по условию $\alpha_1 \in X, \alpha_2 \in X$. Полученное противоречие доказывает теорему.

5. Конечным автоматом с n ($n \in N$) счетчиками и конечным числом занулений назовем совокупность $W \{A, Q, q_1, f, F, z\}$, где $A = \{a_1, a_2, \dots, a_k\}$ (входной алфавит); $Q = \{q_1, q_2, \dots, q_m\}$ (множество состояний автомата); $q_1 \in Q$ (начальное состояние); f —отображение из $A \times Q \times \times \{0, 1\}$ в $Q \times \{-1, 0, +1\}^n$ (f —функция переходов автомата); $F \subseteq Q$ (выделенное подмножество состояний); $z \in N$ (допустимое число занулений по каждому счетчику). Предполагается, что из $f(a_i, q_j, \pi_1, \pi_2, \dots, \pi_n) = (q_l, \sigma_1, \sigma_2, \dots, \sigma_n)$ и $\pi_r = 0$ следует, что σ_r отлично от единицы ($r = 1, 2, \dots, n$). Позициями W назовем элементы множества $Q \times N^n$. Слову $\alpha \in A^*$ ($\alpha = a_{p_1} a_{p_2} \dots a_{p_t}$) поставим в соответствие последовательность $\Pi_0(\alpha), \Pi_1(\alpha), \dots, \Pi_t(\alpha)$ позиций автомата W . Здесь $\Pi_0(\alpha) = (q_1, 0, 0, \dots, 0)$; если $\Pi_i(\alpha) = (q_j, s_1^i, s_2^i, \dots, s_n^i)$, то $\Pi_{i+1}(\alpha) = (q_m, s_1^i + \sigma^1, s_2^i + \sigma^2, \dots, s_n^i + \sigma^n)$, где $f(a_{p_i}, q_j, \text{sign } s_1^i, \text{sign } s_2^i, \dots, \text{sign } s_n^i) = (q_m, \sigma^1, \sigma^2, \dots, \sigma^n)$ ($i = 0, 1, \dots, t-1$) ($\text{sign } x$ —функция, заданная на N и отличная от единицы лишь в точке нуль, $\text{sign } 0 = 0$). Вторая, третья, ..., $(n+1)$ -я компоненты позиции $\Pi_i(\alpha)$ —содержатые счетчики, первого, второго, ..., n -го соответственно. Если в $(i-1)$ -й позиции содержимое некоторого счетчика отлично от нуля, a в i -й оно нулевое, то будем говорить, что на i -м такте работы

автомата произошло зануление этого счетчика. Слово α назовем приемлемым, если а) $\Pi_{\alpha} \in F \times N^n$, б) количество занулений каждого из счетчиков в процессе работы над α не превосходит z . Множество всех приемлемых слов обозначим $Z(W)$.

Теорема 9. Пусть W — конечный автомат с счетчиками и конечным числом занулений. Тогда множество слов $Z(W)$ является Δ -языком.

Доказательство теоремы, которое авторы опускают, состоит в построении алгоритма, конструирующего по заданному W тройку Σ, S, Φ .

Содержимое k -го счетчика в момент времени t обозначим C_t^k . Пусть $C_t^k < C_{t+1}^k = C_{t+2}^k = \dots = C_{t+d-1}^k > C_{t+d}^k$. Тогда будем говорить, что k -й счетчик в момент $t + d$ производит возврат. Положим, что W^0 — конечный автомат с n счетчиками, как это определялось в начале данного раздела, но константа z понимается не как ограничение на число занулений, а как ограничение на число возвратов. Автоматы с одним магазином и ограничением такого рода вместе с выделяемыми ими языками изучались в [7]. Очевидно, что ограничение числа возвратов влечет ограничение на зануления. Поэтому язык, определяемый автоматом со счетчиками с конечным числом возвратов, одновременно является языком, определяемым автоматом со счетчиками с конечным числом занулений.

Теорема 10. Если разрешима проблема пустоты для Δ -языков, то разрешима проблема пустоты языков, определяемых конечными автоматами с n счетчиками с конечным числом занулений (возвратов). Неразрешима проблема пустоты языков, определяемых двухмагазинными автоматами с одним возвратом.

Первая часть этой теоремы — следствие теоремы 9. Вторая часть доказывается сведением к формулируемой проблеме комбинаторной проблемы Поста.

6. Очевидно, что существуют Δ -языки, не являющиеся бесконтекстными. Например, $a^n b^n c^n$ ($n \in N$). Вместе с тем формулируемая ниже теорема 11 показывает, что существуют бесконтекстные языки весьма простой структуры, не являющиеся Δ -языками.

Теорема 11. Существуют языки, определяемые конечными автоматами с одним счетчиком (S_1 -множества в терминах работы [1]), не являющиеся Δ -языками.

Примером такого языка является множество слов вида

$$1^{n_1} 0 1^{n_2} 0 1^{n_3} \dots 0 1^{n_k} 0 1^{n_{k+1}} \quad (k > 0).$$

Легко доказывается следующая теорема.

Теорема 12. Всякий ограниченный бесконтекстный язык [8] является Δ -языком.

Авторы благодарны Ю. В. Глебскому за интерес к работе.

ЛИТЕРАТУРА

1. Л. П. Гаврилова, Д. И. Коган, Изв. высш. уч. зав. — Радиофизика, № 4, 13, 584 (1970).
2. S. Ginsburg, E. Spanier, Pacif. J. Math., 16, 285 (1966).
3. А. И. Мальцев, Алгоритмы и рекурсивные функции, изд. Наука, М., 1965.
4. В. М. Глушков, Синтез цифровых автоматов, изд. Наука, М., 1962.
5. A. Aho, J. Ullman, Math. Syst. Theory, 2, № 2, 97 (1968), (русский перевод, Кiberneticheskiy sbornik, vyp. 6 (naya seriya), 145).
6. S. Greibach, Math. Syst. Theory, 2, № 1, 1 (1968).
7. J. Ginsburg, E. Spanier, J. SIAM Control, 4, 429 (1966).
8. С. Гинзбург, Математическая теория контекстно-свободных языков, изд. Мир, М., 1970.

**SOME CLASSES OF DIGITAL CONTROLLED SYSTEMS AND
LANGUAGES (ALGORITHHMICAL PROBLEMS)***A. A. Bedina., D. I. Kogan*

Some classes of controlled vector systems are introduced. Within these concepts a number of problems in digital control can be stated. Solvability problems of the system's conversion from start state into the one prescribed for various classes of systems are investigated. Δ -language class is introduced and investigated, each of the languages being a set of a certain system's strategies transforming a start vector into a final subset of vectors. The Δ -language concept appears to be a generalization of a regular set notion.

УДК 62—50 + 519.2

ОБ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

B. M. Перельмутер

Рассматриваются оценки характеристик динамических систем методом максимального правдоподобия для случаев, когда корреляционная функция помехи априорно неизвестна и должна оцениваться при идентификации. Приводятся результаты экспериментов.

В [1] рассмотрены некоторые вопросы определения динамических характеристик управляемых объектов методом максимального правдоподобия (ММП), при применении которого по сравнению с методом наименьших квадратов (МНК) достигается существенное уменьшение ошибок идентификации для сильно коррелированной помехи при малом времени наблюдения. Предполагалось, что корреляционная функция помехи известна. Представляет интерес распространение оценок ММП на случаи, когда корреляционная функция помехи априорно неизвестна и должна также оцениваться в процессе идентификации [2, 3].

В настоящей статье проводится анализ некоторых способов определения динамических характеристик при неизвестной корреляционной функции помехи.

Пусть объект моделируется m параллельно соединенными звенями с заданными динамическими характеристиками и переменными коэффициентами усиления u_i . Схема идентификации приведена на рис. 1

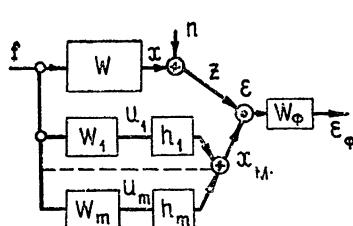


Рис. 1

Здесь W — идентифицируемый объект, W_j ($j = 1, \dots, m$) — заданные динамические звенья, $f(t)$ — входной сигнал — случайный или детерминированный процесс, для которого выполнены условия идентифицируемости (сигналы $u_j(t)$ — линейно независимы). Помеха $n(t)$ предполагается нормально распределенной с нулевым средним, дисперсией D_n и корреляционной функцией $K_n(\tau)^*$. Звездочка (*) обозначает истинное значение величины, а волннистая линия над обозначением — ее оценку.

1. Рассмотрим дискретные системы и сигналы. Выход модели объекта в момент времени k $x_m(k) = \sum_{j=1}^m h_j u_j(k)$. Пусть проведено N измерений. Образуем векторы $u(k) = |u_1(k), \dots, u_m(k)|^t$, $z = |z(1), \dots, z(N)|^t$ и матрицу F размером $N \times m$ $F = |u(1), \dots, u(N)|^t$.

Тогда ММП сводится к решению уравнения

$$(F^t \Lambda F) h = F^t \Lambda z, \quad (1.1)$$

где $\Lambda = K_n^{-1}$, K_n — корреляционная матрица помехи.

Обозначим $\Delta h = \tilde{h} - h^*$ ошибку идентификации. Тогда ковариационная матрица ошибки

$$D_h = M[\Delta h \Delta h^t] = M_f[(F^t \Lambda F)^{-1}]. \quad (1.2)$$

Если $m = 1$, то можно определить границы для D_h :

$$\frac{D_n \mu_{\min}}{D_u N} \leq D_h \leq \frac{D_n \mu_{\max}}{D_u N}, \quad (1.3)$$

где D_u — средний квадрат $u(t)$, μ_{\max} и μ_{\min} — соответственно наибольшее и наименьшее собственные числа матрицы $D_n^{-1} K_n$ (нормированной корреляционной матрицы). Из [4] следует, что μ_{\max} и μ_{\min} асимптотически равны максимальному и минимальному значениям функции спектральной плотности. Например, для $K_n(i) = \exp[-\beta|i|\Delta]$, $\rho = \exp[-\beta\pi]$,

$$\mu_{\max} \approx (1 + \rho)/(1 - \rho), \quad \mu_{\min} \approx (1 - \rho)/(1 + \rho). \quad (1.4)$$

Экспериментальные данные по эффективности ММП по сравнению с МНК приведены в [1].

Предположим, что корреляционная функция помехи известна с точностью до l -мерного вектора параметров p . Дифференцируя логарифм функции правдоподобия по p , получаем систему уравнений для определения p :

$$\frac{\partial \ln L}{\partial p_i} = \frac{1}{2} |\Lambda|^{-1} \frac{\partial |\Lambda|}{\partial p_i} - \frac{1}{2} \tilde{n}^t \frac{\partial \Lambda}{\partial p_i} \tilde{n} = 0 \quad (i=1, \dots, l), \quad (1.5)$$

где $n = z - F\tilde{h}$, так как реализация $n(t)$ непосредственно не наблюдается. Таким образом, в процессе идентификации необходимо производить совместное решение систем уравнений (1.1) и (1.5).

Систему (1.1), (1.5) возможно решать различными итерационными способами. Наиболее целесообразно последовательное решение систем (1.1) и (1.5). На основе некоторой априорной информации выбираем $\Lambda_{(1)}$ и по имеющимся реализациям $u_j(t)$ и $z(t)$ по (1.1) вычисляем оценку $\tilde{h}_{(1)}$; подставляя ее в (1.5), производим оценку p , уточняем матрицу $\Lambda_{(2)}$ и используем ее в (1.1) и т. д. Если априорная информация отсутствует, то на первом шаге Λ заменяется единичной матрицей

Пусть $K_n(i) = D_n \rho^{1/2}$. Дифференцируя $\ln L$ по D_n , находим

$$\tilde{D}_n = \frac{(1+\rho^2) \sum_{i=1}^N \tilde{n}^2(i) - 2\rho \sum_{i=1}^{N-1} \tilde{n}(i) \tilde{n}(i+1) - \rho^2(\tilde{n}^2(1) + \tilde{n}^2(N))}{(1-\rho^2)N}, \quad (1.6)$$

где

$$\tilde{n}(i) = z(i) - u^t(i) \tilde{h}. \quad (1.7)$$

Дифференцируя $\ln L$ по ρ и подставляя (1.6), получаем уравнение для оценки ρ :

$$(N-1) \sum_{i=2}^{N-1} \tilde{n}^2(i) \rho^3 - (N-2) \sum_{i=1}^{N-1} \tilde{n}(i) \tilde{n}(i+1) \rho^2 - ((N+1) \sum_{i=2}^{N-1} \tilde{n}^2(i) + \tilde{n}^2(1) + \tilde{n}^2(N)) \rho + N \sum_{i=1}^{N-1} \tilde{n}(i) \tilde{n}(i+1) = 0. \quad (1.8)$$

Если $\tilde{h} = h^*$, то при $N \rightarrow \infty$ $\sum_{i=2}^{N-1} \tilde{n}^2(i) \rightarrow D_n(N-2)$, $\sum_{i=1}^{N-1} \tilde{n}(i) \tilde{n}(i+1) \rightarrow D_n(N-1)$, ρ^* и ρ^* являются решением получающегося при этом уравнения

ния, т. е. оценка состоятельна. Остальные два корня получающиеся уравнения по модулю больше единицы.

Дисперсия оценки ρ может быть найдена путем линеаризации (1.8) в окрестности ρ^* и в окрестностях математических ожиданий величин коэффициентов. Получаем

$$M[\Delta\rho^2] = \Delta\bar{\rho}^2 = \frac{[N(1-\rho^{*2}) + 3\rho^{*2}-1](1-\rho^{*2})^2 N^2}{(N-1)^2[N(1-\rho^{*2})+2\rho^{*2}]^2}. \quad (1.9)$$

Для возможности применения этого метода средний квадрат ошибки аппроксимации $\eta(k) = x(k) - u^t(k) h^*$ выхода объекта при отсутствии помех должен быть значительно меньше дисперсии помехи. Даже при нулевой ошибке аппроксимации существует взаимное влияние уравнений оценки параметров h_i и уравнений для определения параметров корреляционной функции, что приводит к увеличению ошибки по сравнению с величинами, даваемыми формулами (1.2), (1.9), и к смещению оценки ρ . Действительно, дисперсии оценок коэффициентов h_i зависят от математического ожидания оценки вектора p : $D_h = f_1(M[p])$; с другой стороны, наличие дисперсии D_h оказывается на точности определения p : $M[\tilde{p}] = f_2(D_h)$. Совместное решение этих двух уравнений дает возможность найти D_h и $M[\tilde{p}]$, которые определяются в результате вычислений, а также проанализировать сходимость. Подробнее этот вопрос рассмотрен ниже при анализе непрерывных алгоритмов.

Эксперимент 1. Был проведен ряд опытов по определению коэффициента усиления звена первого порядка $x(k+1) - 0,37x(k) = h^t(k)$ $K_n(i) = 0,37^{11}$, $N = 5$, которые показали работоспособность метода. Ниже приведены результаты одного из опытов: 1 шаг — $\rho = 0$, $\Delta h^2 h^{*-2} = 0,112$; 2 шаг — $\rho = 0,18$; $\Delta h^2 h^{*-2} = 0,072$; 3 шаг — $\rho = 0,24$; $\Delta h^2 h^{*-2} = 0,067$.

Эксперимент 2. Передаточная функция объекта $W(\zeta) = h_1(\zeta - 0,37)^{-1} + h_2(\zeta - 0,72)^{-1}$; $h_1^* = 0,93$; $h_2^* = 0,72$ (ζ — оператор сдвига), остальные условия как в первом примере. Результаты одного из опытов выглядят так: 1 шаг — $\rho = 0$; $\Delta h_1^2 + \Delta h_2^2 = 0,038$; 2 шаг — оценка $\rho = 0,12$; $\Delta h_1^2 + \Delta h_2^2 = 0,023$; 3 шаг — $\rho = 0,22$, $\Delta h_1^2 + \Delta h_2^2 = 0,015$.

2. Изложенный метод требует обращения матрицы, и его целесообразно применять в том случае, если известна аналитическая форма обратной матрицы и ее определителя. Для определения обратной матрицы в том случае, когда помеха представляет собой процесс авторегрессии, возможно использовать метод, приведенный в [5]. В ряде случаев возможно применить более простые методы, имеющие меньшую эффективность при конечном времени наблюдения, однако при $T \rightarrow \infty$ их эффективность стремится к эффективности ММП. В [2, 3] предлагается, например, минимизировать средний квадрат ошибки $\epsilon_{\Phi}(t)$ после фильтра W_{Φ} , обратного формирующему (рис. 1). Рассмотрим этот вопрос подробнее.

Пусть спектральная плотность помехи — дробно-рациональная функция. Представим корреляционную матрицу K_n в виде произведения нижней и транспонированной к ней верхней треугольных матриц. Тогда $\Lambda = \Lambda'_1 \Lambda_1$ и логарифм правдоподобия

$$\ln L = \ln c + \frac{1}{2} \ln |\Lambda_1|^2 - \frac{1}{2} m^t m, \quad (2.1)$$

где $m = \Lambda_1 n$, $M[m m^t] = E$. Таким образом, матрица Λ_1 осуществляет преобразование последовательности $n(k)$ в последовательность независимых величин $m(k)$. Пусть $K_n(i)$ такова, что уравнение дискретного формирующего фильтра $n(k) + \dots + a_l n(k-l) = \xi(k)$, где $\xi(k)$ — дискретный белый шум. Так как это уравнение можно рассматривать как обратное, формирующее последовательность независимых величин из процесса с корреляционной функцией $K_n(i)$, то отсюда следует, что начиная с $l+1$ строки, соответствующие элементы Λ_1 должны совпадать с коэффициентами этого уравнения. Далее, поскольку, начиная с $l+1$ строки, диагональные элементы Λ_1 равны σ_ξ^{-1} , то $|\Lambda_1| = \sigma_\xi^{-N} \varphi(a)$, где $\varphi(a)$ — некоторая функция коэффициентов a_i , не меняющаяся при увеличении N . Следовательно,

$$\ln \Delta = \ln c + \frac{1}{2} \ln \varphi^2(a) - \frac{N}{2} \ln \sigma_\xi^2 - \frac{1}{2} \sum_{i=1}^N m_i^2, \quad (2.2)$$

и при $N \rightarrow \infty$ максимизация функции правдоподобия сводится к нахождению таких коэффициентов a_i и b_i , при которых минимален средний квадрат величины после фильтра, обратного формирующему.

Если уравнение формирующего фильтра $n(k) + \dots + a_l n(k-l) = \xi(k) + \dots + b_q \xi(k-q)$ ($q < l$), то $m(k) = \sum_{j=0}^{\infty} \psi(j) n(k-j)$, $\psi(j)$ — весовая функция фильтра, обратного формирующему. Можно принять $m(k) = \sum_{j=0}^s \psi(j) n(k-j)$. Тогда, начиная с $s+1$ строки в Λ_1 , элементы этих строк будут представлять собой величины соответствующих дискрет $\psi(j)$. Таким образом, минимизация среднего квадрата ошибки на выходе фильтра, обратного формирующему, путем настройки параметров этого фильтра эквивалентна ММП только при $N \gg s$, что обычно является более тяжелым условием, чем $N \gg l$ в первом случае.

Для $b_i = 0$ коэффициенты a_i находятся из линейной системы уравнений

$$\sum_{j=0}^l a_j \left(\sum_{i=l+1}^N \tilde{n}(i-j) \tilde{n}(i-k) \right) = 0 \quad (a_0 = 1; k = 1, \dots, l). \quad (2.3)$$

Система (2.3) решается совместно с системой (1.1), как это описано выше.

В общем случае $b_i \neq 0$ применяются различные итерационные методы [2, 3, 6].

Если $K_n(i) = D_n \rho^{|i|}$, то, применяя разложение в ряд, можно найти (при $\mathbf{h} = \mathbf{h}^*$)

$$\overline{\Delta^2 p} \simeq \frac{1 - \rho^2}{N - 1}, \quad (2.4)$$

так что с учетом (1.9) эффективность оценки ρ

$$\frac{\overline{\Delta^2 p}_{\text{ММП}}}{\overline{\Delta^2 p}_{\text{МНК}}} = \frac{N^2 (1 - \rho^{*2}) [(1 - \rho^{*2}) N + 3\rho^{*2} - 1]}{(N - 1) [N(1 - \rho^{*2}) + 2\rho^{*2}]^2}. \quad (2.5)$$

Эксперимент 3. Для определения эффективности ММП была проведена серия опытов по определению коэффициента ρ корреляционной функции $K_n(i) = \rho^{|i|}$, $N = 10$. Для ММП коэффициент ρ рассчитывался по (1.8); для упрощенного метода — по (2.3). Для малых ρ ($\rho < 0,3$) раз-

ницы практически не наблюдалось. Для $\rho = 0,7$ получены следующие величины дисперсий оценок: для ММП $\bar{\Delta}^2 \rho = 0,04$, для упрощенного метода — 0,055, что дает хорошее совпадение с (1.9), (2.4).

3. Переходим к рассмотрению непрерывных систем и сигналов. Формулы ММП для таких систем приведены в [1]. Ковариационная матрица дается формулой (3.13) [1]. Пусть производится оценка одного параметра h при корреляционной функции помехи $K_n = D_n \exp[-\alpha |\tau|]$. Для синусоидального входного сигнала $f(t) = A \sin(\omega t)$ эффективность оценки при $\omega T = k\pi$ ($k = 1, 2, \dots$)*.

$$\frac{D_{h\text{ММП}}}{D_{h\text{МНК}}} = \frac{\chi (\chi^2 + 9,87 k^2)}{\chi (\chi^2 + 9,87 k^2) + 19,74 k^2 (1 - \exp(-\chi)) (-1)^k}, \quad (3.1)$$

где $\chi = \alpha T$.

Если $f(t)$ — случайная функция, то дисперсии оценок ММП и МНК находятся путем разложения в ряде выражений

$$D_{h\text{МНК}} = M_u \left\{ \int_0^T \int_0^T K_n(t_1 - t_2) u(t_1) u(t_2) dt_1 dt_2 \left[\int_0^T u^2(t) dt \right]^{-2} \right\}, \quad (3.2)$$

$$D_{h\text{ММП}} = 2\alpha D_n M_u \left\{ \left[\int_0^T ((p + \alpha) u(t))^2 dt \right]^{-1} \right\}. \quad (3.3)$$

Пусть $K_u(\tau) = D_u \exp[-\beta |\tau|] (1 + \beta |\tau|)$, тогда

$$D_{h\text{ММП}} \approx \frac{2\alpha D_n}{TD_u(\alpha^2 + \beta^2)} \left[1 + \frac{5\alpha^2 + 2\alpha^2\beta^2 + \beta^4}{T\beta(\alpha^2 + \beta^2)^2} \right]. \quad (3.4)$$

Можно показать, что для этого случая при больших T

$$\frac{D_{h\text{МНК}}}{D_{h\text{ММП}}} \approx 1 + \frac{2\beta^3}{\alpha(\alpha + \beta)^2}, \quad (3.5)$$

так что оценки МНК асимптотически не эффективны.

Для непрерывных сигналов предел плотности вероятностей наблюдаемых координат процесса не существует, а предел отношения правдоподобия существует не во всех случаях [7]. Для рациональных спектральных плотностей этот предел существует, если предел отношения спектральных плотностей сравниваемых процессов при $\omega \rightarrow \infty$ равен единице. Таким образом, дисперсия процесса должна быть связана с параметрами корреляционной функции. Например, для возможности сравнения экспоненциальных корреляционных функций $K_n = D_n \times \exp[-\beta_i |\tau|]$ должно выполняться условие $D_{n_i} \beta_i = \text{const}$. Если предел отношения вероятностей существует, то для спектральной плотности $S_n(p) = a^2 [\alpha_0 + \alpha_1 p + \dots + \alpha_{q-1} p^{q-1} + p^q]^{-2}$ достаточными статистиками являются $\int_0^T n^{(i)}(t) dt$ ($i = 0, 1, \dots, q-1$); a^2 можно определить по одной

реализации теоретически с вероятностью 1, однако при этом используются свойства выборочных функций, связанные с высокочастотной частью спектра, так что на практике этот способ неприменим из-за неизбежного действия высокочастотных составляющих помехи, имеющих малую амплитуду [8]. Кроме того, получение производных высокого порядка также часто затруднено.

* Эта формула уточняет формулу (3.32) [1], которая приближенно справедлива только при $\omega T \gg 1$.

Поэтому для определения параметров корреляционной функции необходимо вводить дополнительные уравнения. Эти уравнения можно получить путем обработки сигнала $n(t)$, пропущенного через различные динамические звенья. При этом порядок производной должен быть не выше $q - 1$. Эти звенья выбираются такими, чтобы все сигналы, используемые для определения параметров корреляционных функций, были бы линейно независимыми и получающиеся уравнения легко разрешимыми. Такой же метод целесообразно применять в том случае, когда числитель $S(p)$ не равен постоянной; в этом случае достаточных статистик не существует.

Пусть $K_n(\tau) = D_n \exp[-\beta|\tau|]$. Используем статистики $c_1 = \frac{1}{T} \int_0^T \tilde{n}^2(t) dt$ и $c_2 = \frac{1}{T} \int_0^T \tilde{n}_1^2(t) dt$, где $\tilde{n}_1(p) = \frac{T_1 p + 1}{T_2 p + 1} \tilde{n}(p)$.

Тогда

$$\tilde{\beta} = \frac{T_2(c_2/c_1 - 1)}{T_1^2 - T_2^2 c_2/c_1}. \quad (3.6)$$

Оценка состоятельная с дисперсией (при $\tilde{n} = n^*$)

$$\frac{D_\beta}{\beta^{*2}} = \frac{2b\delta\nu(1+\delta)^2(1+b) + (\delta-1)[b^2(3\delta+1) - \delta - 3] - 8b\delta}{b^2\delta^2\nu^2(1+\delta)^2}, \quad (3.7)$$

$$b = T_2 \tilde{\beta}, \quad \delta = T_1/T_2, \quad \nu = T/T_1.$$

Если $K_n = D_n \exp[-\alpha|\tau|] (\cos \beta\tau + \alpha\beta^{-1} \sin \beta|\tau|)$, то используем статистики c_1 и c_2 , как в предыдущем случае, а также $c_3 = \frac{1}{T} \int_0^T \tilde{n}^{12} dt$. Тогда $\gamma^2 = \alpha^2 + \beta^2 = c_3 c_1^{-1}$,

$$\alpha = \frac{(c_2/c_1)(1 + T_2^2 \gamma^2) - 1 - T_1^2 \gamma^2}{2T_2(1 - c_2/c_1)}. \quad (3.8)$$

Выше говорилось о том, что для определения математического ожидания оценки \tilde{p} и ковариационной матрицы ошибок D_h с учетом их взаимного влияния необходимо решить два уравнения:

$$D_h = f_1(M[\tilde{p}]), \quad M[\tilde{p}] = f_2(D_h).$$

Рассмотрим этот вопрос подробнее для $K_n = D_n \exp[-\beta|\tau|]$.

При достаточно большом T ММП для оценки \tilde{p} при фиксированном β заключается в минимизации среднего квадрата величины ε_Φ , причем $W_\Phi = p + \beta$. Сигналы, пропущенные через звено $p + \beta$, будем обозначать с индексом 0 (нуль):

$$D_h = A^{-1} D_L A^{-1} = f_1(\tilde{\beta}),$$

$$D_L = \frac{1}{T^2} \int_0^T \int_0^T K_{n_0}(t_1 - t_2) M[u_0(t_1) u_0^t(t_2)] dt_1 dt_2, \quad (3.9)$$

где A — матрица с элементами $a_{ij} = \frac{1}{T} \int_0^T u_{0i}(t) u_{0j}(t) dt$.

Далее,

$$c_1 = \frac{1}{T} \int_0^T (n(t) + h^{*t} u(t) - \tilde{h}^t u(t))^2 dt \simeq \\ (3.10)$$

$$\simeq D_n + \frac{1}{T} \int_0^T u^t(t) \Delta h \Delta h^t u(t) dt,$$

$$c_2 = \frac{1}{T} \int_0^T \left[\frac{T_1 p + 1}{T_2 p + 1} (n(t) + h^{*t} u(t) - \tilde{h}^t u(t)) \right]^2 dt \simeq \\ (3.11)$$

$$\simeq \frac{1}{T} \int_0^T u_2^t(t) \Delta h \Delta h^t u_2(t) dt + D_{n_2},$$

где u_2 и n_2 — сигналы $u(t)$ и $n(t)$, пропущенные через звено $(T_1 p + 1) \times (T_2 p + 1)^{-1}$. Подставляя (3.10) и (3.11) в (3.6), проводя усреднение, приближенно полагая, что математическое ожидание функции равно функции от математического ожидания, заменяя интегралы математическими ожиданиями подынтегральных выражений, находим уравнение $M[\beta] = f_2(D_h)$.

Рассмотрим случай $m = 1$. Для определенности примем $S_u(p) = 4D_u \beta^{*3} (p + \beta^*)^{-2} (-p + \beta^*)^{-2}$. Выполняя описанные выше преобразования, находим

$$D_h = \frac{D_n}{D_u \beta^* T} \frac{3\beta^{*4} + 3\tilde{\beta}_0^4 + 2\beta^{*2}\tilde{\beta}_0^2}{(\beta^{*2} + \tilde{\beta}_0^2)^2 + (2/\beta^* T)(\beta^{*4} + 2\beta^{*2}\tilde{\beta}_0^2 + 5\tilde{\beta}_0^4)}, \quad (3.12)$$

$$\tilde{\beta}_0 = \beta^* \frac{D_n/D_u + D_h T_2 \beta^* (1 + T_2 \beta^*)^{-1}}{D_n D_u^{-1} + D_h (1 + 2\beta^* T_2) (1 + \beta^* T_2)^{-1}},$$

где для упрощения записи обозначено $\beta_0 = M[\beta]$. Пусть $D_n/D_u = 1$, $T_1 = 1$ сек, $T_2 = 0,1$ сек, $\beta^* = 1$, $T = 5$ сек. Решая (3.12) и (3.13), находим $M[\beta] = 0,82$, $D_h = 0,248$. Отметим, что если β было бы известно точно, $\tilde{\beta} = \beta^* = 1$, то $D_h = 0,22$. При применении МНК $D_h = 0,30$.

Уравнения (3.12), (3.13) можно использовать для анализа сходимости выбранного итерационного процесса, используя известные критерии сходимости.

Эксперимент 4. Оценка коэффициента усиления $h^* = 1$, $u = \sin(2,8t)$, $K_n = 0,25 e^{-0,3|t|}$, $T = 4,2$ сек. По полученной реализации по МНК находим $h = 0,875$, затем производим оценку помехи по формуле $\tilde{n} = z - \tilde{h}u$ и по (3.6) для $T_1 = 1$ сек, $T_2 = 0,1$ сек вычисляем $\tilde{\beta} = 0,25$.

По ММП уточняем оценку h , получаем $\tilde{h} = 0,93$.

Другой способ оценки параметров корреляционной функции помехи основан на дискретизации непрерывных сигналов. Пусть известна корреляционная функция помехи с точностью до параметров. Выбирая интервал дискретности θ , по известным методам находим выражение

для корреляционной функции дискретного процесса с точностью до этих же параметров. Фиксируя ординаты процесса $\tilde{n}(t)$ через θ и применяя методы первого раздела, находим параметры корреляционной функции. Этот метод особенно удобен при использовании ЭЦВМ.

Найдем дисперсию ошибки при $K_n = D_n \exp[-\beta|\tau|]$. Величина $\tilde{\beta}$ находится по формуле $\tilde{\beta} = \ln \rho^{\theta-1}$. Тогда

$$\Delta \beta = \frac{1}{\theta} \frac{1}{\rho^*} \Delta \rho \quad (3.13)$$

и

$$\overline{\Delta \beta^2} = \frac{1}{\theta^2 \rho^{*2}} \overline{\Delta \rho^2}. \quad (3.14)$$

Подставляя в (3.14), (2.4) и учитывая $N-1 = T/\theta$, получаем

$$D_\beta = (e^{2\beta\theta} - 1)/T\theta. \quad (3.15)$$

Это выражение имеет минимум при $\theta = 0$, равный $2\beta/T$. Аналогично при использовании (1.9)

$$D_{\beta_{\min}} = D_{\beta|\theta=0} = \frac{2\beta}{T + \beta^{-1}}, \quad (3.16)$$

так что эффективность

$$\frac{D_{\beta_{\min \text{ ММП}}}}{D_{\beta_{\min \text{ МНК}}}} = \frac{\beta T}{1 + \beta T}. \quad (3.17)$$

Для определения коэффициента корреляции ρ может также использоваться формула

$$\tilde{\rho}_\theta = e^{-\tilde{\beta}|\theta|} = \frac{\int_0^{\theta} n(t) n(t+\theta) dt}{\int_0^{\theta} n^2(t) dt}. \quad (3.18)$$

Дисперсия этой оценки также минимальна при $\theta \rightarrow 0$ и равна

$$D_{\beta_{\min}} = D_{\beta|\theta=0} = \frac{2\beta}{T}. \quad (3.19)$$

На практике выбору малого шага θ препятствует ограниченные быстродействие и память вычислительного устройства.

Кроме того, приведенные формулы являются приближенными, так как получены линеаризацией нелинейных зависимостей. Далее, при определении ρ могут иметь место ошибки, не связанные с величиной θ , например, из-за погрешности вычислительных устройств, округления и т. д. Для такого рода погрешностей ошибка $\Delta \beta$ минимальна при $\beta\theta = 1$ и при уменьшении $\beta\theta$ до 0,4 и менее резко возрастает. Указанное обстоятельство следует иметь в виду при выборе интервала θ .

Проведенный в статье анализ и экспериментальные исследования показывают возможность применения ММП для идентификации объектов автоматического управления в том случае, когда корреляционная функция помехи априорно неизвестна и должна оцениваться при идентификации. При этом возможно использование приведенных выше методов и формул. Хотя изложение в статье ведется применительно

к идентификации объекта, передаточная функция которого представляется в виде суммы заданных передаточных функций с переменными коэффициентами усиления, рассмотренные методы могут быть легко распространены на другие способы описания объекта, например, с помощью дифференциальных или интегральных уравнений.

ЛИТЕРАТУРА

1. В. М. Перельмутер, Изв. высш. уч. зав. — Радиофизика, 13, № 4, 506 (1970).
2. К. Ж. Астром, Т. Болин, сб. Теория самонастраивающихся систем управления, изд. Наука, М., 1969.
3. A. E. Rogers, K. Steiglitz, IEEE International Convention Record, part 3, 74 (1967).
4. У. Гренандер, Г. Сеге, Теплицевые формы и их приложения, ИЛ, М., 1961.
5. S. A. Tretter, IEEE Trans., AC—13, 3, 302 (1968).
6. J. H. Eaton, IEEE International Convention Record, part 3, 38 (1967).
7. Б. Р. Левин, Теоретические основы статистической радиотехники, кн. 2, изд. Сов. радио, М., 1968.
8. В. Ф. Писаренко, Радиотехника и электроника, 6, № 4, 515 (1961).

Научно-исследовательский
электротехнический институт

Поступила в редакцию
7 апреля 1970 г.

ON PLANTS IDENTIFICATION BY MAXIMUM LIKELIHOOD METHOD

V. M. Perelmutter

Dynamic system characteristics estimation by the maximum plausibility method is considered when noise correlation function is a priori unknown and is to be estimated in the process of identification. Experimental results are given.

УДК 621.373.93

К ТЕОРИИ РЕЗИСТИВНОГО ПАРАМЕТРИЧЕСКОГО ГЕНЕРАТОРА

T. B. Мошкина, A. C. Таранович

Методом Ван-дер-Поля рассмотрена динамика симметричного генератора на туннельных диодах. Найдено уравнение семейства амплитудно-частотных характеристик. Рассмотрены режимы параметрического возбуждения и синхронизации генератора при воздействии внешней ЭДС, составляющей с собственной частотой системы произвольное рациональное отношение. Исследовано влияние постоянного смещения туннельных диодов на режимы работы схемы. Найдены фазово-частотные характеристики. Отдельно рассмотрено влияние активной нагрузки.

В радиотехнике большое распространение получили активные делители и умножители частоты как автономные, так и неавтономные [1]. Симметричная схема резистивного параметрона [2, 3] по классификации, предложенной в работе [1], является неавтономным двухтактным делителем частоты на два. Кроме режима неавтономного деления в схеме возможен режим автономного деления. В этом случае внешний сигнал вызывает синхронизацию возбужденного генератора на субгетронике. В работе [4] осуществлен общий подход к рассмотрению возможных в такой схеме режимов. Однако принятые в этой работе представление характеристики туннельного диода (ТД) полиномом третьей степени не позволило описать некоторые наблюдающиеся на практике режимы [2–4], в том числе режимы деления и умножения частоты с произвольной кратностью. Повышение степени аппроксимирующего полинома приводит к резкому усложнению анализа, к утрате непосредственной связи с физическими параметрами системы [3], поэтому уточнение аналитических результатов на этом пути встречает значительные трудности.

В предлагаемой работе для вольтамперной характеристики использована иная аппроксимация, содержащая две экспоненты. Такая аппроксимация обеспечивает хорошую точность на всем рабочем интервале характеристики [5, 6], позволяет качественно рассмотреть возможные в системе режимы деления и умножения частоты с любой кратностью и найти уравнения семейств амплитудно-частотных и фазово-частотных характеристик.

1. УКОРОЧЕННЫЕ УРАВНЕНИЯ СИСТЕМЫ

Рассмотрим генератор на ТД, принципиальная схема которого приведена на рис. 1. Как и в [1], предположим, что индуктивности L_1 и L_2 и коэффициент взаимоиндукции M равны между собой и постоянны; емкости контура C_1 и C_2 , включающие емкости ТД, также равны и постоянны; $\mathcal{E} = E_0 + E \cos(\omega t)$ — суммарная ЭДС смещения и синхронизации; i_1 и i_2 — токи в индуктивностях. Пунктиром на рис. 1 обозначено сопротивление внешней активной нагрузки, влияние которой рассмотрено отдельно. Вольтамперные характеристики ТД считаются одинаковы-

ми, т. ч. симметрия схемы сохраняется и по нелинейным элементам. Для их аппроксимации используется уравнение

$$J_j = a u_j \exp(-b u_j) + J^* [\exp(c u_j) - 1], \quad (1)$$

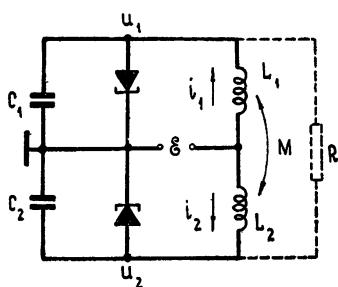


Рис. 1.

в котором коэффициенты a , b , J^* , c определяются по реальной характеристике диода [5, 6].

При сделанных предположениях система уравнений, описывающая поведение генератора без учета нагрузки, имеет следующий вид:

$$2LC\ddot{u} + 2L\frac{dJ}{dt} + u = 0, \quad (2)$$

$$u_0 = E_0 + E \cos(\bar{\omega} t),$$

где

$$u = \frac{u_1 - u_2}{2}, \quad J = \frac{J_1 - J_2}{2}, \quad u_0 = \frac{u_1 + u_2}{2},$$

$$J = a u \exp(-b u_0) \operatorname{ch}(b u) - a u_0 \exp(-b u_0) \operatorname{sh}(b u) +$$

$$+ J^* \exp(c u_0) \operatorname{sh}(c u).$$

Вводя новые переменные $\tau = \bar{\omega} t$; $x = b u$ и параметры

$$2\varepsilon = bE, \quad \beta = c/b, \quad \alpha = \exp(-bE_0)/bJ^*, \quad \gamma_1 = (E_0/J^*) \exp(-bE_0),$$

$$\gamma_2 = \exp(cE_0)/a, \quad \omega = \bar{\omega} \sqrt{2LC}, \quad \delta = abJ^* \sqrt{2L/C},$$

получим безразмерное уравнение

$$\omega^2 \ddot{x} + \delta \omega \frac{df}{d\tau} + x = 0, \quad (3)$$

где

$$f = \exp(-2\varepsilon \cos \tau) \{ \alpha x \operatorname{ch} x - \gamma_1 \operatorname{sh} x - 2\alpha\varepsilon \cos \tau \operatorname{sh} x +$$

$$+ \gamma_2 \exp[2(\beta + 1)\varepsilon \cos \tau] \operatorname{sh}(\beta x) \}.$$

Будем искать решение уравнения (3) в виде

$$x = z_1 \exp\left(-jk \frac{\tau}{\nu}\right) + z_2 \exp\left(jk \frac{\tau}{\nu}\right), \quad (4)$$

$$\dot{x} = -j \frac{k}{\nu} \left[z_1 \exp\left(-jk \frac{\tau}{\nu}\right) - z_2 \exp\left(jk \frac{\tau}{\nu}\right) \right],$$

где k/ν — рациональное отношение, а z_1 и z_2 будем считать медленно меняющимися комплексно-сопряженными функциями времени. Подставляя искомое решение (4) в уравнение (3), вводя новый масштаб времени $\vartheta = \tau/\nu$ и производя усреднение правой части уравнения [7], получим систему укороченных уравнений (комплексно-сопряженное уравнение может быть найдено аналогичным путем) в виде

$$2j \frac{\omega^2}{\nu^2} k \dot{z}_1 = \left(1 - \frac{k^2 \omega^2}{\nu^2}\right) z_1 - j \delta \frac{k}{\nu} \omega x, \quad (5)$$

где

$$\times = \frac{1}{T} \int_0^T f e^{ik\vartheta} d\vartheta,$$

а точкой обозначено дифференцирование по ϑ . Из уравнения (5) следует, что усреднение правомерно, если $\omega k/v \approx 1$ и $\delta \ll 1$. Подставляя \times в уравнение (5) и переходя к действительным переменным ρ и φ (амплитуда и фаза колебаний) с помощью подстановки $z_1 = \rho e^{i\varphi}$, $z_2 = \rho e^{-i\varphi}$, получим следующую систему укороченных уравнений:

$$\begin{aligned} \dot{\rho} &= -\frac{\delta v \rho}{2\omega} \Phi_0(\rho, \varepsilon) - \frac{\delta v \rho}{2\omega} \sum_{l>0} \Phi_{1l}(\rho, \varepsilon) \cos(vl\varphi), \\ \dot{\varphi} &= \frac{[(k\omega/v)^2 - 1] v^2}{2k\omega^2} + \frac{\delta v}{2\omega} \sum_{l>0} \Phi_{2l}(\rho, \varepsilon) \sin(vl\varphi). \end{aligned} \quad (6)$$

В этой системе уравнений введены следующие обозначения:

$$\begin{aligned} \Phi_0(\rho, \varepsilon) &= \sum_{p, q=0}^{\infty} \frac{\varepsilon^{2q}}{(q!)^2} \frac{\rho^{2p}}{p!(p+1)!} M_s, \\ \Phi_{1l}(\rho, \varepsilon) &= \sum_{p, q=0}^{\infty} \frac{\varepsilon^{2q+kl}}{q!(q+kl)!} \frac{\rho^{2p+v l-2}}{p!(p+v l)!} (2p + v l) M_{s+r}, \\ \Phi_{2l}(\rho, \varepsilon) &= \sum_{p, q=0}^{\infty} \frac{\varepsilon^{2q+kl}}{q!(q+kl)!} \frac{\rho^{2p+v l-2}}{p!(p+v l)!} v l M_{s+r}, \end{aligned} \quad (7)$$

где

$$\begin{aligned} M_s &= \alpha s - \gamma_1 + \gamma_2 \beta^s, \\ M_{s+r} &= \alpha (-1)^{kl}(s+r) - (-1)^{kl}\gamma_1 + \gamma_2 \beta^{s+r}, \\ s &= 2p + 2q + 1, \\ r &= kl + v l - 2, \end{aligned} \quad (8)$$

а l — целочисленный параметр, изменяющийся от единицы до бесконечности и принимающий все четные значения при нечетном v , любые значения — при четном v .

Коэффициенты M_s и M_{s+r} имеют вполне определенную связь с вольтамперной характеристикой тунNELьного диода (1). Действительно, находя n -ю производную от выражения (1) в рабочей точке и переходя к параметрам α , β , γ_1 , γ_2 , получим

$$\begin{aligned} \left. \frac{d^n J}{d u^n} \right|_{E_0} &= J^* a b^n [(-1)^{n+1} \alpha n + \\ &+ (-1)^n \gamma_1 + \gamma_2 \beta^n]. \end{aligned} \quad (9)$$

Сравнивая (9) с (8), легко заметить, что M_s с точностью до постоянных множителей $J^* a b^n$ определяет ряд нечетных производных, а M_{s+r} — ряд нечетных производных, если kl четное, и ряд четных производных, если kl нечетное. По коэффи-

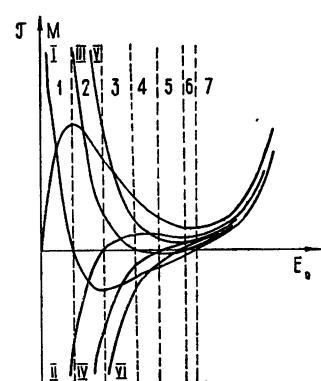


Рис. 2.

циентам a, b, c, J^* , которые определяются из реальной характеристики ТД одним из методов, предложенных в [5, 6], могут быть построены зависимости M_s и M_{s+r} от смещения E_0 . На рис. 2 приведена экспериментальная характеристика ТД и построены для нее коэффициенты $M_1 - M_6$ (соответствующие кривые отмечены римскими цифрами). При построении для простоты считалось, что $\beta = 1$. Для реальных характеристик германиевых ТД β колеблется в пределах 0,9 — 1.

2. СТАЦИОНАРНЫЕ РЕШЕНИЯ

Стационарные решения системы (6) найдем, приравнивая нулю ее правые части. В дальнейшем для простоты ограничимся в разложениях по l этой системы только первыми членами. При этом в случае четных ν должны быть взяты значения Φ_{1l} и Φ_{2l} , соответствующие $l=1$, а в случае нечетных ν — значения, соответствующие $l=2$.

Кроме тривиального решения $\rho = 0$, стационарные решения системы (6) определяются уравнениями

$$\begin{aligned} \Phi_{1l}(\rho, \varepsilon) \cos(\nu l \varphi) &= -\Phi_0(\rho, \varepsilon), \\ \delta\Phi_{2l}(\rho, \varepsilon) \sin(\nu l \varphi) &= -\frac{[(k \omega/\nu)^2 - 1] \nu}{k \omega} = -\zeta. \end{aligned} \quad (10)$$

Исключая φ из уравнений (10), получим уравнение, связывающее «расстройку» ζ , амплитуду установившихся колебаний ρ и амплитуду начаеки ε — амплитудно-частотную характеристику

$$\zeta = \pm \frac{\delta\Phi_{2l}(\rho, \varepsilon)}{\Phi_{1l}(\rho, \varepsilon)} [\Phi_{1l}^2(\rho, \varepsilon) - \Phi_0^2(\rho, \varepsilon)]^{1/2}. \quad (11)$$

Отсюда ясно, что стационарное решение вида (4) существует, если выполняется условие

$$\Phi_{1l}^2(\rho, \varepsilon) > \Phi_0^2(\rho, \varepsilon). \quad (11)$$

На рис. 3 представлены графики возможных вариантов поведения

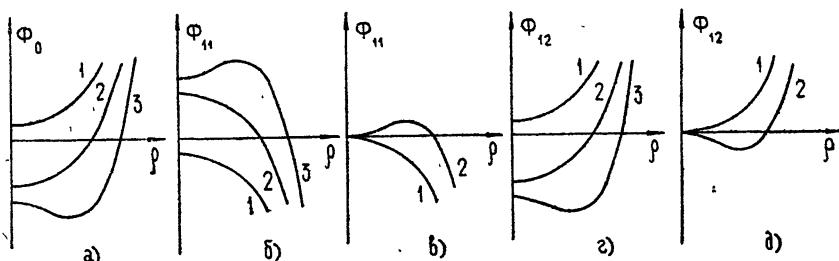


Рис. 3.

функций $\Phi_0(\rho, \varepsilon)$ и $\Phi_{1l}(\rho, \varepsilon)$ для различных значений ν и k и фиксированного значения $\varepsilon \neq 0$, полученные на основании общих методов исследования функций. Из (7) и (8) видно, что $\Phi_0(\rho, \varepsilon)$ не зависит от ν и k , а в зависимости от соотношения знаков M_1 и M_3 может представиться три варианта ее поведения, приведенных на рис. 3а. Поведение $\Phi_{1l}(\rho, \varepsilon)$ зависит от ν и k . Поэтому целесообразно рассмотреть отдельно следующие случаи: 1) $k = 1$, ν -четное — деление частоты на четных субгармониках, 2) $k = 1$, ν -нечетное — деление на нечетных субгармониках, 3) $k > 1$, $\nu = 1$ — умножение частоты.

Первому случаю соответствует функция $\Phi_{11}(\rho, \varepsilon)$, возможные варианты поведения которой в зависимости от соотношения знаков M_2 и

M_4 приведены на рис. 3б для случая $\nu = 2$, а на рис. 3в — для случая $\nu \geq 4$ в зависимости от знака M_ν .

Второму случаю соответствует функция $\Phi_{12}(\rho, \varepsilon)$. На рис. 3г представлено ее поведение в зависимости от соотношения знаков M_3 и M_5 для случая $\nu = 1$, а на рис. 3д — для случая $\nu \geq 3$ в зависимости от знака $M_{2\nu+1}$.

Третьему случаю также соответствует функция $\Phi_{12}(\rho, \varepsilon)$, однако ее поведение зависит от соотношения знаков M_{2k+1} и M_{2k+3} , причем для приведенной на рис. 2 характеристики ТД возможен только один вариант поведения $\Phi_{12}(\rho, \varepsilon)$ — кривая 1, рис. 3г.

Рис. 2 дает представление об изменении первых шести номеров коэффициентов M_s или M_{s+r} при изменении смещения E_0 . Весь интервал изменения E_0 разбит на семь областей (обозначены арабскими цифрами), в которых эти коэффициенты сохраняют одно и то же соотношение знаков. Оказывается, что каждой такой области соответствует одна из возможных комбинаций поведения функций $\Phi_0(\rho, \varepsilon)$ и $\Phi_{11}(\rho, \varepsilon)$, представленных на рис. 3.

Используя рис. 2, рис. 3 и уравнение (11), можно качественно построить амплитудно-частотные характеристики системы для различных значений ν , k , E_0 , проследить как меняется их вид при изменении смещения E_0 и амплитуды внешней ЭДС, установить характерные для каждой из семи областей смещения режимы работы системы.

Следует отметить, что уравнения (7), (8), если известна характеристика ТД, позволяют, в случае необходимости, сделать и количественные построения.

На рис. 4 изображены возможные виды амплитудно-частотных характеристик системы, построенные на основе качественного анализа и исследования устойчивости в малом стационарных решений, в предложении того, что в системе используются ТД, характеристика которых изображена на рис. 2. Кружками обозначены устойчивые узлы или фокусы, крестиками — седла, кружками с крестиками — неустойчивые узлы или фокусы.

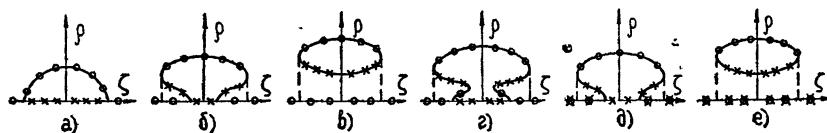


Рис. 4.

В зависимости от смещения в невозбужденной системе (ось $\rho = 0$ устойчива при $\varepsilon = 0$) в определенной полосе расстроек могут наблюдаться режимы мягкого (рис. 4а) и жесткого (рис. 4б) параметрического возбуждения, причем в последнем случае на краях полосы возбуждения появляются характерные зоны «затягивания». Особый случай жесткого возбуждения изображен на рис. 4в — колебания возникают только при задании определенных начальных условий, зоны затягивания отсутствуют. Как отмечено в [2], такой режим наблюдается экспериментально, однако в рамках аппроксимации характеристики ТД кубическим полиномом не поддается описанию, как и обычный жесткий режим. Анализ показывает, что режим рис. 4а наблюдается в области I (рис. 2) при $\nu = 2$ и в области 2 при $\nu = 1$. Режим 4б — в области 7 при $\nu = 2$, режим 4в — в области 7 при четных ν . При нечетных ν параметрическое возбуждение в областях 1 и 7 невозможно, при четных $\nu \geq 4$ — в области I.

В самовозбужденной системе (ось $\rho = 0$ неустойчива при $\varepsilon = 0$)

в определенной полосе расстроек наблюдаются режимы параметрической синхронизации колебаний, т. е. колебания системы на частоте, кратной ЭДС накачки и близкой к собственной частоте генератора (4д, е). Анализ показывает, что режим 4д наблюдается в области 2 при $\nu = 1$ и в областях 2 и 6 при $\nu = 2$. Режим 4е — в областях 2—6 для четных и нечетных ν .

В случае умножения частоты ($k > 1$, $\nu = 1$) вид резонансных кривых и возможные режимы работы подобны случаю $k=1$, $\nu = 1$. Однако, как это видно из (7) и (11), полоса синхронизации или возбуждения с ростом k ввиду малости ε резко сужается.

3. ФАЗОВО-ЧАСТОТНЫЕ ХАРАКТЕРИСТИКИ

Фазово-частотная характеристика системы, полученная из уравнений (10) и (11), имеет следующий вид:

$$\operatorname{tg}(\nu l \varphi) = \pm \frac{[\Phi_{1l}^2(\rho, \varepsilon) - \Phi_0^2(\rho, \varepsilon)]^{1/2}}{\Phi_0(\rho, \varepsilon)}. \quad (13)$$

Отсюда

$$\varphi = \pm \varphi_0 / \nu l \pm n\pi / \nu l, \quad (14)$$

где

$$\varphi_0 = \operatorname{arctg} \frac{[\Phi_{1l}^2(\rho, \varepsilon) - \Phi_0^2(\rho, \varepsilon)]^{1/2}}{\Phi_0(\rho, \varepsilon)} \quad (n = 0, 1, 2, \dots).$$

Знак плюс перед φ_0 соответствует положительным расстройкам, а минус — отрицательным. Выражением (14) определяются возможные значения фазы, удовлетворяющие уравнению (13). Однако допустимыми для исследуемой системы являются фазы, одновременно удовлетворяющие и системе (10), которую можно представить в несколько ином виде:

$$\begin{aligned} \cos(\nu l \varphi) &= -\Phi_0(\rho, \varepsilon) / \Phi_{1l}(\rho, \varepsilon), \\ \sin(\nu l \varphi) &= -(\pm A) / \Phi_{1l}(\rho, \varepsilon), \end{aligned} \quad (15)$$

где $A = [\Phi_{1l}^2(\rho, \varepsilon) - \Phi_0^2(\rho, \varepsilon)]^{1/2}$. Выбор знака перед A делается аналогично выбору знака перед φ_0 .

Выражение (14) показывает, что количество устойчивых фаз в рассматриваемой симметричной системе всегда четно и не зависит от номера k .

На рис. 5 представлены два варианта фазово-частотной характеристики, которые могут наблюдаться в системе, и отмечены пределы полосы синхронизации или возбуждения. Качественный вид характеристики построен на основе анализа выражений (13) — (15). Какой из них реализуется в конкретном режиме работы системы, зависит от значения ν и относительного расположения нулей функций $\Phi_0(\rho, \varepsilon)$ и $\Phi_{1l}(\rho, \varepsilon)$,

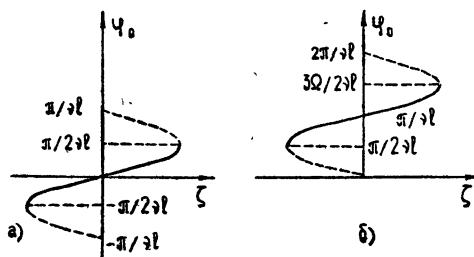


Рис. 5

определеняющих амплитудно-частотную характеристику рассматриваемого режима. Следует отметить, что вид изображенных на рис. 5 характеристик соответствует «замкнутым» амплитудно-частотным характеристикам (рис. 4в, е), причем пунктиром обозначены фазы неустойчивых ветвей амплитудно-частотных характеристик. В случае «незамкнутых» амплитудно-частотных характеристик (рис. 4а, б, г, д) ветви фазово-частотных характеристик, соответствующие неустойчивым ветвям, реализуются неполностью или вовсе отсутствуют.

4. УЧЕТ ВЛИЯНИЯ НАГРУЗКИ

Рассмотрим влияние активной нагрузки R , подключение которой обозначено на рис. 1 пунктиром. С учетом нагрузки и в прежних обозначениях безразмерное уравнение (3) примет вид

$$\omega^2 \ddot{x} + \delta\omega \frac{df}{d\tau} + \sigma\dot{x} + x = 0, \quad (16)$$

где $\sigma = (2/R) \sqrt{2L/C}$.

Приближенная система укороченных уравнений записывается в следующем виде:

$$\begin{aligned} \dot{\rho} &= -\frac{\delta\omega\rho}{2\omega} \left[\Phi_0(\rho, \varepsilon) + \frac{\sigma}{\delta} \right] - \frac{\delta\omega\rho}{2\omega} \Phi_{1l}(\rho, \varepsilon) \cos(\nu l \varphi), \\ \dot{\varphi} &= \xi \frac{\nu}{2\omega} + \frac{\delta\omega}{2\omega} \Phi_{2l}(\rho, \varepsilon) \sin(\nu l \varphi), \end{aligned} \quad (17)$$

откуда получается уравнение амплитудно-частотной характеристики стационарного решения

$$\zeta = \pm \frac{\delta\Phi_{2l}(\rho, \varepsilon)}{\Phi_{1l}(\rho, \varepsilon)} \{ \Phi_{1l}^2(\rho, \varepsilon) - [\Phi_0(\rho, \varepsilon) + \sigma/\delta]^2 \}^{1/2}. \quad (18)$$

Условие существования стационарного решения теперь определяется неравенством

$$\Phi_{1l}^2(\rho, \varepsilon) > [\Phi_0(\rho, \varepsilon) + \sigma/\delta]^2, \quad (19)$$

которое выполняется труднее, чем неравенство (12). Отсюда видно, что при увеличении нагрузки (уменьшение $|R|$) полоса синхронизации или возбуждения сужается, а амплитуда установившихся колебаний становится меньше. С помощью рис. 3 можно детально проследить это влияние.

Кроме того, подключение нагрузки вызывает изменение условий самовозбуждения системы, в той или иной степени подавляя генерацию. Оказывается, что выбором нагрузки при $\nu = 2$ в областях 2, 3, а при $\nu = 1$ в областях 2, 3, 6 можно установить режим 4а; при $\nu = 2$ в областях 4, 5, 6, а при $\nu = 1$ в областях 4, 5 — жесткий режим 4б или в. При $\nu = 1$ и $\nu = 2$ в областях 4, 5 при определенных значениях нагрузки может наблюдаться режим 4г, отличающийся в определенной полосе расположением двух устойчивых значений амплитуды колебаний. При $\nu > 2$ выбором нагрузки может быть установлен режим 4б только в областях 4, 5 для нечетных ν и только в областях 4, 5, 6 для четных ν . В областях 2, 3 и под нагрузкой сохраняются режимы 4е, причем с непрерывным ростом нагрузки полоса синхронизации и амплитуда синхронизованных колебаний непрерывно уменьшаются до исчезновения области синхронизации.

Найдя новое выражение для фазово-частотной характеристики

и сравнивая его с выражением (13), нетрудно установить, что пределы изменения фазы в полосе расстроек при подключении нагрузки не меняются.

Подводя итоги, можно отметить следующее.

Внешняя ЭДС, поданная в схему выбранным способом, периодически меняет дифференциальную проводимость туннельного диода, т. е. меняет условия самовозбуждения генератора. Это приводит к возникновению колебаний в невозбужденной системе или к синхронизации колебаний в возбужденной системе. Общность причины возникновения явлений и характер воздействия на систему позволяют в рассматриваемом случае возбуждение и синхронизацию называть параметрическими.

Принятая аппроксимация вольт-амперной характеристики туннельного диода (2) позволила получить уравнение семейства амплитудно-частотных характеристик, справедливое при малых расстройках и амплитудах накачки для любого кратного отношения частот, а также объяснить появление большинства наблюдаемых на опыте режимов работы системы.

В невозбужденной системе мягкий режим возникновения колебаний возможен только в случаях $\nu = 1$ и $\nu = 2$. В случае $\nu > 2$ режим возникновения колебаний жесткий.

В возбужденной системе, при любых значениях ν , возможен режим синхронизаций, когда частота колебаний в системе близка к собственной и кратна внешней ЭДС, причем кратность может быть дробной. Ширина полосы синхронизации пропорциональна δ и сужается при увеличении значения ν . Выражения (7) и (11) позволяют сделать вывод, что полоса синхронизации на четных субгармониках шире, чем на соседних нечетных.

Увеличение значения k , определяющего кратность умножения частоты, ведет к резкому сужению полосы синхронизации или возбуждения. Этот вывод базируется на выражениях (7) и (11), из которых видно, что ширина полосы зависит от значения $e^{k\nu}$.

В рассматриваемой симметричной системе количество устойчивых фаз возбужденных или синхронизованных колебаний, как это видно из (14), не зависит от k и всегда четно, причем равно ν или 2ν в случае четных и нечетных значений ν соответственно.

Как видно из рис. 5, фаза устойчивых ветвей амплитудно-частотной характеристики, при изменении расстройки в пределах полосы синхронизации от отрицательных значений к положительным, меняется от $-\pi/2\nu l$ до $\pi/2\nu l$ относительно допустимых значений фаз, соответствующих нулевой расстройке ($l = 1$, если ν — четное, и $l = 2$, если ν — нечетное). В случае умножения частоты, как это следует из (14), может быть только две устойчивых фазы, независимо от значения k .

Подключение нагрузки приводит к сужению полосы синхронизации или возбуждения, уменьшению амплитуды установленныхся колебаний и появлению режимов параметрического возбуждения, мягких или жестких, в области падающего участка характеристики ТД для любых значений ν .

Пределы изменения фазы внутри полосы синхронизации или возбуждения при подключении нагрузки не меняются.

ЛИТЕРАТУРА

- 1 И. Х. Ризкин, Умножители и делители частоты, изд. Связь, М., 1966.
- 2 Ф. П. Жарков, Магнитные элементы, изд. Наукова думка, Киев, 1964.
- 3 В. А. Редько, Резистивные параметрические генераторы, изд. Зинатне, Рига, 1968.
- 4 А. С. Тарапотович, Изв. высш. уч. зав. — Радиофизика, 11, № 7, 1046 (1968).
- 5 A. Ferendeci, W. H. Ko, Proc. IRE, 50, № 8 (1962).

6. Ю. Л. Симонов, Радиотехника, 20, № 4, 62 (1965).
 7. А. А. Андронов, А. А. Витт, С. Э. Хайкин, Теория колебаний, ГИФМЛ, М., 1959.

Научно-исследовательский физико-технический
институт при Горьковском университете

Поступила в редакцию
25 мая 1970 г.

TO THE THEORY OF A RESISTIVE PARAMETRIC GENERATOR

T. V., Moshkina, A. S. Tarantovich

Dynamics of a tunnel diode symmetric generator is considered by the Van der Pol method. The equation for amplitude-frequency characteristics family is found. Modes of parametric excitation and synchronization of the generator are considered under the influence of external E. M. F., constituting an arbitrary rational ratio with the system's natural frequency. The effect of fixed bias voltage in diodes on the circuit's operation modes is investigated. Phase-frequency characteristics are found. A separate consideration of resistive load influence is carried out.

Примечание при корректуре. На рис. 5 б значение $\varphi_0 = 3\Omega/2\nu l$ следует читать $\varphi_0 = 3\pi/2\nu l$.

УДК 62 — 50

АНАЛИЗ ЭЛЕКТРОПНЕВМАТИЧЕСКОГО СЕРВОМЕХАНИЗМА С СУХИМ ТРЕНИЕМ СПОСОБОМ ГАРМОНИЧЕСКОЙ ЛИНЕАРИЗАЦИИ НЕЛИНЕЙНОСТЕЙ

A. Л. Глик

Исследована устойчивость линеаризованного методом гармонического баланса электропневматического сервомеханизма с большим коэффициентом усиления с сухим трением.

Функционирование электропневматического сервомеханизма с сухим трением можно описать уравнениями

$$\begin{aligned} \ddot{x} + \alpha \dot{x} + x_0 \operatorname{sgn} \dot{x} &= f; \\ T_2 \dot{f} + f &= K_2 \alpha' \\ \alpha' &= y - K_3 \dot{x}; \\ T_1 \dot{y} + y &= K_1 I; \\ I &= (U - K_{oc} x) K_y. \end{aligned} \tag{1}$$

В уравнениях x — перемещение поршня; $x_0 = \frac{R}{m}$, m , R , α характеризует соответственно массу подвижных частей, сумму сил сухого трения и вязкое трение двигателя; $f = \frac{F}{m}$, F — развиваемое усилие; y и I — положение управляющего органа и разность токов в обмотках управления электромеханического преобразователя; U — управляющее напряжение; K_{oc} — коэффициент обратной связи; K_1 , K_2 , K_3 , K_y , T_1 , T_2 — постоянные коэффициенты.

Метод гармонического баланса позволяет приближенно заменить нелинейную функцию $x_0 \operatorname{sgn} \dot{x}$ линейной, так что

$$x_0 \operatorname{sgn} \dot{x} \approx \frac{q(A)}{\Omega} \dot{x}, \tag{2}$$

где коэффициент $q(A)$ для однозначной характеристики сухого трения равен

$$q(A) = \frac{4x_0}{\pi A}, \tag{3}$$

A и Ω — амплитуда и частота периодического решения [1].

Для линеаризованной системы получим следующее характеристическое уравнение

$$T_1 T_2 k' p^4 + \left\{ T_1 T_2 \left[\alpha + \frac{q(A)}{\Omega} \right] + T_1 + T_2 \right\} k' p^3 + \left\{ T_1 \left[\alpha + \frac{q(A)}{\Omega} \right] + \right.$$

$$+ T_2 \left[\alpha + \frac{q(A)}{\Omega} \right] + 1 + T_1 K_2 K_3 \left\{ k' p^2 + \left[1 + \frac{k' q(A)}{\Omega} \right] p + k_a = 0, \quad (4)$$

где $k_a = K_1 K_y K_{oc} K_2 k' = K_1 K_y K_{oc} K_2 \left(\alpha \left[1 + k_2 k_3 \frac{1}{\alpha} \right] \right)^{-1}$ — коэффициент усиления линейной части сервомеханизма.

При $p = j \Omega$ имеем

$$k_a + T_1 T_2 k' \Omega^4 - (T_1 \alpha + T_2 \alpha + 1 + T_1 K_2 K_3) k' \Omega^2 - (T_1 + T_2) q(A) k' \Omega = 0, \quad (5)$$

$$\Omega + k' q(A) - (T_1 T_2 \alpha + T_1 + T_2) k' \Omega^3 - T_1 T_2 q(A) k' \Omega^2 = 0.$$

Разрешая уравнения (5) относительно A и k_a , имеем

$$A = \frac{4x_0 (1 - T_1 T_2 \Omega^2) k'}{\pi [(T_1 T_2 \alpha + T_1 + T_2) k' \Omega^2 - 1] \Omega};$$

$$k_a = - T_1 T_2 k' \Omega^4 + (T_1 \alpha + T_2 \alpha + 1 + T_1 K_2 K_3) k' \Omega^2 + \quad (6)$$

$$+ (T_1 + T_2) \frac{4x_0}{\pi A} k' \Omega.$$

Зная A и Ω , можно использовать второе уравнение (6) для определения параметра нелинейной системы k_a как функции текущего значения амплитуды колебаний и наоборот.

При $A \rightarrow \infty$ (*)

$$k_a = k_{1kp} = \frac{(T_1 T_2 \alpha + T_1 + T_2) k' (T_1 \alpha + T_2 \alpha + 1 + T_1 K_2 K_3) - T_1 T_2}{(T_1 T_2 \alpha + T_1 + T_2)^2 k'}$$

Кривая $A = f(k_a)$ стремится к оси k_a и к асимптоте k_{1kp} при $k_a \rightarrow \infty$. Эта кривая в зависимости от значений параметров может заходить в область $k_a < k_{1kp}$ или располагаться полностью правее асимптоты $k_a = k_{1kp}$ (рис. 1, 2).

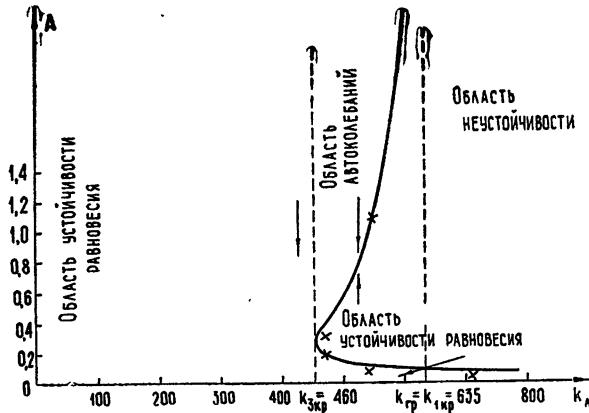


Рис. 1.

Так как $A \geq 0$, то должно быть $1 - T_1 T_2 \Omega^2 \geq 0$ и $(T_1 T_2 \alpha + T_1 + T_2) \times k' \Omega^2 - 1 \geq 0$ (или $1 - T_1 T_2 \Omega^2 \leq 0$ и $(T_1 T_2 \alpha + T_1 + T_2) k' \Omega^2 - 1 \leq 0$). Это определяет значения Ω .

* k_{1kp} совпадает со значением коэффициента усиления порождающей линейной системы, находящейся на границе устойчивости.

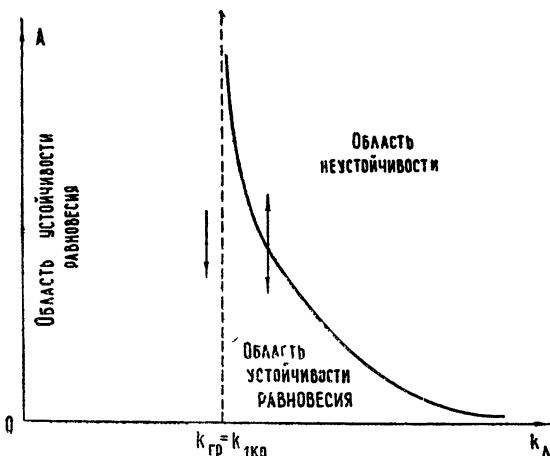


Рис. 2.

$$\frac{1}{\sqrt{T_1 T_2}} \geq \Omega \geq \frac{1}{\sqrt{(T_1 T_2 \alpha + T_1 + T_2) k'}} \quad (7)$$

$$\left(\frac{1}{\sqrt{T_1 T_2}} \leq \Omega \leq \frac{1}{\sqrt{(T_1 T_2 \alpha + T_1 + T_2) k'}} \right).$$

Полученные точки соответствуют устойчивому решению при

$$\left(\frac{\partial X}{\partial a} \right)^* \left(\frac{\partial Y}{\partial \omega} \right)^* - \left(\frac{\partial X}{\partial \omega} \right)^* \left(\frac{\partial Y}{\partial a} \right)^* > 0$$

(X — вещественная и Y — мнимая части выражения кривой Михайлова), что приводит к неравенству

$$\left[(T_1 \alpha + T_2 \alpha + 1 + T_1 K_2 K_3) \Omega + (T_1 + T_2) \frac{4x_0}{\pi A} - 2 \dot{T}_1 T_2 \Omega^3 \right] \times \\ \times (T_1 T_2 \Omega^2 - 1) > \left[(T_1 T_2 \alpha + T_1 + T_2) \Omega + T_1 T_2 \frac{4x_0}{\pi A} \right] (T_1 + T_2) \Omega^2. \quad (8)$$

Проведенные расчеты приводят к следующей картине. При значениях параметров $T_1 = 0,005$ сек, $T_2 = 0,02$ сек, $\alpha = 642$ сек $^{-1}$, $K_2 = 1670$ м сек $^{-2}$ град $^{-1}$, $K_3 = 6,2$ град сек м $^{-1}$, $x_0 = 100$ м сек $^{-2}$ верхней ветви кривой на рис. 1 соответствуют устойчивые периодические режимы (стрелки показывают направление изменения амплитуд колебаний). При стремлении амплитуды колебаний A к нулю функция $q(A) = \frac{4x_0}{\pi A}$ стремится к бесконечности. Поэтому, условие (8) не выполняется и предельные циклы, соответствующие малым A , неустойчивы. Отсутствуют автоколебания и для всех точек нижней ветви рис. 1, (а также кривой рис. 2).

В соответствии с изложенным неустойчивая линейная система под влиянием трения не будет иметь расходящихся колебаний при всех отклонениях A , меньших тех, которые соответствуют точкам нижней ветви кривой рис. 1 и кривой рис. 2, т. е. сухое трение делает неустойчивую линейную систему устойчивой в малом. Наряду с этим сухое трение в случае, представленном на рис. 1, для значений k_A , соответст-

вующих устойчивой порождающей линейной системе, может вызвать развитие устойчивых автоколебаний при больших возмущениях.

На рис. 1 показаны точки, полученные электромоделированием. Как видно, количественные отклонения, даваемые методом гармонического баланса, могут быть весьма незначительными. Изложенные результаты качественно подтверждены экспериментально.

ЛИТЕРАТУРА

1. Е. И. Попов, И. П. Пальтов, Приближенные методы исследования нелинейных автоматических систем, Физматгиз, М., 1960.

Тульский политехнический институт

Поступила в редакцию
7 июля 1969 г.

ANALYSIS OF ELECTROPNEUMATIC SERVOMECHANISM WITH DRY FRICTION BY APPLYING DESCRIBING FUNCTION METHOD TO NONLINEARITIES

A. L. Glik

Stability of a linearized electropneumatic servomechanism with a high gain by dry friction is analyzed by the describing function method.

УДК 519.8

ТЕСТИРОВАНИЕ МЕТОДА КУСОЧНО-ЛИНЕЙНОЙ АППРОКСИМАЦИИ НА ОДНОМ КЛАССЕ МНОГОЭКСТРЕМАЛЬНЫХ ФУНКЦИЙ

Д. И. Батищев, А. В. Литвер

На классе функций с ограниченным сверху числом локальных экстремумов исследуется эффективность метода кусочно-линейной аппроксимации. Приводятся экспериментальные зависимости эффективности от точности и параметров, характеризующих класс тестовых функций и алгоритм поиска.

1. Задача поиска с заданной точностью координат локальных минимумов непрерывной функции $f(x)$, определенной на отрезке $[a, b]$, часто встречается в вычислительной практике. При этом допускается, что вид функции неизвестен, но ее значение может быть вычислено без ошибки в любой точке $x \in [a, b]$. В качестве примера можно привести задачи, связанные с вычислением корней одномерного нелинейного уравнения, определением полинома наилучшего приближения при равноволновой аппроксимации.

В [1] для решения сформулированной задачи предложен метод кусочно-линейной аппроксимации, процесс поиска по которому осуществляется в два этапа.

На первом этапе, состоящем из k шагов, в равноотстоящих точках интервала $[a, b]$, число которых определяется по итерационной формуле

$$N_r = 2N_{r-1} - 1 \quad (r = 2, 3, \dots, k), \quad (1)$$

$$N_1 = 5,$$

вычисляются значения функции $f(x)$. Так, на первом шаге ($r = 1$) имеется пять точек, на втором шаге ($r = 2$) — девять точек, но в число последних входят те пять, в которых на предыдущем шаге уже были вычислены значения $f(x)$, и т. д.

По полученным значениям $(x_j, f(x_j))$, $j = 1, 2, \dots, N_r$, на каждом шаге строится кусочно-линейная модель $F(N_r)$, для которой по знаку разностей $(f(x_j) - f(x_{j-1}))$, $j = 1, 2, \dots, N_r$, определяется число J , подынтервалов «подозрительных» на существование в них локального минимума функции $f(x)$. Процесс последовательного построения кусочно-линейных функций $F(N_r)$ ($r = 1, 2, \dots, k$) заканчивается, как только достигается соответствие между моделью, полученной на k -м шаге, и исследуемой функцией. В качестве критерия такого соответствия принимается значение параметра M , показывающего, сколько раз подряд число «подозрительных» подынтервалов кусочно-линейной модели повторялось при ее уточнении. Таким образом, при повторении структуры кусочно-линейных функций $F(N_r)M$ раз подряд считается, что функция $f(x)$ имеет такое же число локальных минимумов, как и кусочно-линейная модель $F(N_k)$, построенная на k -м шаге.

Второй этап поиска сводится к локальной минимизации при помо-

щи метода Кифера—Джонсона [2, 3] функции $f(x)$ в каждом из «подозрительных» подынтервалов, полученных на k -й модели.

Общее число вычислений функции $f(x)$, необходимое для определения координат локальных минимумов с точностью ε , равно

$$N = N_k(M) + J_k R(N_k, \varepsilon), \quad (2)$$

где $N_k(M)$ — число вычислений функции $f(x)$, необходимое для построения кусочно-линейных моделей, структура которых совпадает M раз подряд; J_k — число «подозрительных» подынтервалов k -й модели; $R(N_k, \varepsilon)$ — число вычислений функции $f(x)$, необходимое для поиска локального минимума с точностью ε в каждом из «подозрительных» подынтервалов $2(b-a)/(N_k - 1)$.

2. Рассмотрим эффективность метода кусочно-линейной аппроксимации на класс функций $\Phi(J)$, обладающих следующими свойствами.

1) Функции $\varphi(x) \in \Phi(J)$ определены и непрерывны на интервале $[a, b]$.

2) Максимальное число локальных минимумов любой из функций $\varphi(x) \in \Phi(J)$ не больше некоторого значения J

$$\max_{\varphi \in \Phi} J_\varphi \leq J. \quad (3)$$

Этот класс функций типичен для задач аппроксимации полиномами и дробно-рациональными функциями, наименее уклоняющимися от нуля [4].

Нетрудно показать, что класс тестовых функций, адекватный классу функций $\Phi(J)$, может быть получен в виде тригонометрических полиномов степени n

$$\varphi(x) = a_1 + \sum_{k=1}^n \left(a_k \cos \frac{(k-1)\pi}{b-a} x + a_{n+k} \sin \frac{(k-1)\pi}{b-a} x \right). \quad (4)$$

При этом очередная тестовая функция получается путем «генерирования» [5] коэффициентов a_j , $j = 1, 2, \dots, 2n-1$, в соответствии с определенным законом $p(a_j)$ их распределения на интервале $[A, B]$. Тогда эффективность исследуемого алгоритма поиска может быть вычислена по формуле

$$\bar{N}(J, \varepsilon, M) = \int_A^B N(J, \varepsilon, M, \varphi(x, \alpha)) \times p(\alpha) d\alpha. \quad (5)$$

Верхняя оценка J на число локальных минимумов тригонометрического полинома зависит от числа слагаемых гармоник, т. е. определяется степенью полинома n . Для получения этой зависимости проведем следующее вспомогательное построение.

Для полинома (4) справедливы следующие оценки:

$$|\varphi(x)| \leq K, \quad |\varphi'(x)| \leq L \text{ и } \overline{\varphi}(x) \leq S,$$

где

$$K = m(\sqrt{2}(n-1) + 1);$$

$$L = \frac{m\pi}{(b-a)\sqrt{2}} n(n-1);$$

$$S = m \left(1 + \frac{2}{\pi} (\ln(n-1) + 0,577) \right);$$

$$m = \max(|A|, |B|).$$

На интервале $[a, b]$ строим кусочно-линейную функцию

$$F(x) = \begin{cases} -Lx, & (x \in [0, Q/L], [3Q/L, 5Q/L]...) \\ Lx, & (x \in [Q/L, 2Q/L], [5Q/L, 7Q/L]...) \end{cases}, \quad (7)$$

где $Q = K - |S|$. Координата j -го локального минимума функции $F(x)$ вычисляется по формуле

$$x_j = \frac{Q}{L}(4j - 3) \quad (j = 1, 2, \dots, J). \quad (8)$$

Для J -го локального минимума из построения $F(x)$ имеем

$$E\left(\frac{b-a}{x_j}\right) = 1, \quad (9)$$

где $E(z)$ — целая часть от z . Тогда, учитывая (8) и подставляя из (6) значения K, L и S , согласно (9) получаем

$$J = E\left(\frac{\pi n(n-1)}{8\left(n-1-\frac{\sqrt{2}}{\pi}(\ln(n-1)+0,577)\right)} + \frac{3}{4}\right) + 1. \quad (10)$$

Исследование формулы (10) показало, что каждые пять новых членов суммы тригонометрического полинома увеличивают максимально возможное число локальных минимумов J на два. Например, для степеней полинома $n = 5, 10, 15$ и 20 параметр J , соответственно, имеет значения $4, 6, 8$, и 10 . Экспериментальная проверка формулы (10) для каждого из n по 1000 реализациям полинома (5) с коэффициентами $\alpha_i \in [-100, 100], i = 1, 2, \dots, 2n-1$, показала ее справедливость для $n = 5, 10, 15, \dots, 40$. На рис. 1 в качестве примера приведены гистограммы числа локальных минимумов J_φ для тригонометрических полиномов 10 и 15 степени.

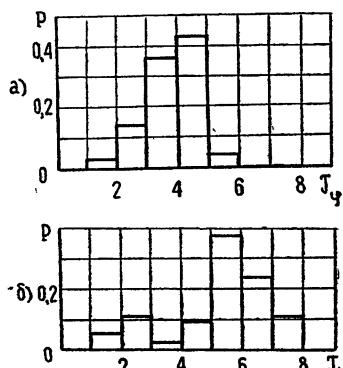


Рис. 1. Гистограммы числа локальных минимумов J_φ для полиномов степени n .

Рис. 2. Зависимость эффективности метода \bar{N} от точности поиска ε

Рис. 2.

Рис. 2. Зависимость эффективности метода \bar{N} от точности поиска ε

Таким образом, тестовый класс $\Phi(J)$ определяется числом гармоник тригонометрического полинома (5), а очередная тестовая функция $\varphi(x) \in \Phi(J)$ получается путем разыгрывания компонент вектора α в соответствии с определенным законом распределения вероятностей.

Для исследования зависимости эффективности метода кусочно-линейной аппроксимации $\bar{N}(J, \epsilon, M)$ для каждой из выборок (J, ϵ, M) , где $J = 4, 6, 8, 10$; $\epsilon = 10^{-1}, 10^{-2}, 10^{-6}$ и $M = 2, 3, 4, 5$, было решено по 100 тестовых задач. При этом функции $\varphi(x)$ были определены на интервале $[0, 100]$, а коэффициенты $a_j, j = 1, 2, \dots, 2n - 1$, с равновероятным законом распределения генерировались в диапазоне $[-100, +100]$. Результаты экспериментального тестирования метода кусочно-линейной аппроксимации приведены в таблице 1 и на рис. 2.

Таблица 1

M	$\epsilon = 10^{-1}$				$\epsilon = 10^{-2}$				$\epsilon = 10^{-6}$			
	$J=4$	$J=6$	$J=8$	$J=10$	$J=4$	$J=6$	$J=8$	$J=10$	$J=4$	$J=6$	$J=8$	$J=10$
2	53	90	121	158	64	110	149	196	106	189	258	374
3	56	116	189	258	68	161	217	291	110	230	357	466
4	92	180	271	317	98	190	295	347	123	305	424	518
5	101	263	311	328	112	280	330	353	154	396	460	524

Из таблицы 1 и рис. 2 видно, что метод кусочно-линейной аппроксимации обладает высокой эффективностью поиска на классе функций, имеющих ограниченное сверху число локальных минимумов. При этом увеличение значения параметра M , характеризующего алгоритм поиска, слабо влияет на эффективность метода.

ЛИТЕРАТУРА

1. Д. И. Батищев, Сб. Тезисы докладов I Всесоюзного межвузовского симпозиума по прикладной математике и кибернетике, Горький, 1967.
2. J. Kiefer, Proc. of Amer. Math. Society, v. 4, № 3, 502 (1953).
3. Р. Беллман, Динамическое программирование, ИЛ, 1960.
4. Н. И. Ахиезер, Лекции по теории аппроксимации, изд. Наука, М., 1965.
5. Л. А. Растигин, Автоматика и вычислительная техника, № 2, 1 (1968).

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
28 июня 1970 г.

TESTING OF PIECEWISE-LINEAR APPROXIMATION METHOD OVER ONE CLASS OF POLYEXTREMAL FUNCTIONS

D. I. Batishchev, A. V. Litver

The efficiency of the piecewise-linear approximation method is investigated over a class of functions with a bounded from above number of local extrema. Experimental dependences of the efficiency upon accuracy and parameters characterizing the class of testing functions and the search algorithm are given.

**КРАТКИЕ СООБЩЕНИЯ
И ПИСЬМА В РЕДАКЦИЮ**

УДК 621.396 6 078 6

**О ВЛИЯНИИ ФОРМЫ ХАРАКТЕРИСТИКИ ФД НА РАБОТУ СИСТЕМЫ ФАП
ПРИ ДЕЙСТВИИ ПОМЕХ**

Б. Э. Рахимов

В современных радиотехнических устройствах для стабилизации частоты широкое применение находит фазовая автоподстройка частоты (ФАП). Системы ФАП часто работают в условиях постоянно действующей помехи. Изучение работы системы ФАП при наличии помех с синусоидальной характеристикой фазового детектора (ФД) без фильтра и с фильтром первого порядка проводилось в [1-5]. В [6] рассмотрено воздействие помех на бесфильтровую систему ФАП с прямоугольной характеристикой ФД.

Форма характеристики ФД влияет, как известно, на ряд динамических показателей системы ФАП, таких как полосу захвата, время переходного процесса и т. д. В настоящей работе рассмотрено влияние формы характеристики ФД на некоторые простейшие статистические показатели системы ФАП: функцию распределения разности фаз и относительную среднюю остаточную расстройку по частоте. Фильтр в цепи обратной связи полагается единичным.

Рассматривается кусочно-нелинейная аппроксимация характеристики ФД следующего вида:

$$F(\varphi) = \begin{cases} [2/\pi(\varphi + \pi/2)]^n & (-\pi \leq \varphi < 0) \\ [-2/\pi(\varphi - \pi/2)]^n & (0 \leq \varphi \leq \pi) \end{cases}, \quad (1)$$

где $n = \dots, 1/5, 1/3, 1, 3, 5, \dots$.

Характеристика $F(\varphi)$ является периодической функцией φ с периодом 2π . Форма характеристики $F(\varphi)$ определяется значением параметра n , так при $n = 1$ характеристика является треугольной, при $n \rightarrow 0$ стремится к прямоугольной*. На рис. 1 в каче-

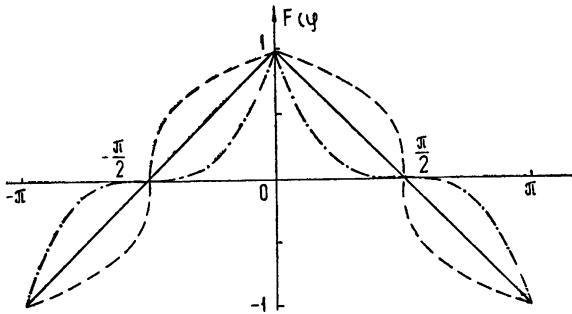


Рис. 1

стве приведен вид характеристики $F(\varphi)$ при $n = 1/3, 1$ и 3 (соответственно этим значениям n пунктирными, сплошными и штрих-пунктирными линиями обозначены кривые также на рис. 2 и 3).

Стochasticическое дифференциальное уравнение системы ФАП при флуктуациях фаз сигналов, поступающих на ФД, может быть записано в виде (см., например, [5]):

$$p\varphi + \Delta K(p) F(\varphi) = \Delta_0 + \xi(t), \quad (2)$$

* Этот случай сводится к рассмотренному в [6] заменой $\tilde{\varphi} = \varphi + \pi/2$.

где φ — разность фаз колебаний подстраиваемого и эталонного генераторов (ПГ и ЭГ); $p \equiv \frac{d}{dt}$ — оператор дифференцирования; $K(p)$ — коэффициент передачи фильтра низких частот, Δ — полоса синхронизма, Δ_0 — начальная расстройка ПГ и ЭГ, $F(\varphi)$ — нормированная характеристика ФД; $\xi(t) = \frac{d}{dt} (\Theta_{\text{ПГ}}(t) - \Theta_{\text{ЭГ}}(t))$, $\Theta_{\text{ПГ}}(t)$ и $\Theta_{\text{ЭГ}}(t)$ — случайные фазы сигналов, поступающих на ФД.

При $K(p) = 1$ уравнение (2) имеет вид

$$\dot{\varphi} + \Delta F(\varphi) = \Delta_0 + \xi(t), \quad (3)$$

здесь и в (2) $F(\varphi)$ определена по формуле (1). Случайный процесс $\xi(t)$ считается дельта-коррелированным с корреляционной функцией $K_\xi(\tau) = K\delta(\tau)$, тогда $\varphi(t)$ есть марковский процесс [7]. В силу периодичности $F(\varphi)$ $\varphi(t)$ рассматривается на окружности длиной 2π . Функция плотности вероятности разности фаз $W(\varphi, t)$ в этом случае является решением уравнений Колмогорова — Фоккера — Планка:

$$\begin{aligned} \frac{\partial W(\varphi, t)}{\partial t} &= - \frac{\partial}{\partial \varphi} [K_{11}(\varphi) W(\varphi, t)] + \frac{1}{2} \frac{\partial^2}{\partial \varphi^2} [K_{21}(\varphi) W(\varphi, t)] \\ &\quad (-\pi < \varphi < 0), \\ \frac{\partial W(\varphi, t)}{\partial t} &= - \frac{\partial}{\partial \varphi} [K_{12}(\varphi) W(\varphi, t)] + \frac{1}{2} \frac{\partial^2}{\partial \varphi^2} [K_{22}(\varphi) W(\varphi, t)] \end{aligned} \quad (4)$$

где

$$(0 < \varphi < \pi)$$

$$K_{11}(\varphi) = \Delta_0 - \Delta [2/\pi (\varphi + \pi/2)]^n,$$

$$K_{12}(\varphi) = \Delta_0 - \Delta [-2/\pi (\varphi - \pi/2)]^n, \quad (5)$$

$$K_{21}(\varphi) = K_{22}(\varphi) = \int_{-\infty}^{\infty} K_\xi(\tau) d\tau = K.$$

При $\varphi = 0, \pm \pi$ $W(\varphi, t)$ удовлетворяет условиям сопряжения [8, 9]:

$$\begin{aligned} W(\varphi, t) \Big|_{-\pi+0} &= W(\varphi, t) \Big|_{\pi-0}, \\ \left[\frac{2K_{11}(\varphi)}{K} W(\varphi, t) - \frac{d}{d\varphi} W(\varphi, t) \right]_{-\pi+0} &= \left[\frac{2K_{12}(\varphi)}{K} W(\varphi, t) - \frac{d}{d\varphi} W(\varphi, t) \right]_{\pi-0}. \end{aligned} \quad (6)$$

Используя условия сопряжения (6) и условие нормировки $\int_{-\pi}^{\pi} W(\varphi, t) d\varphi = 1$, решение уравнений (4) в стационарном случае $\left(\frac{\partial W}{\partial t} = 0 \right)$ находится в виде:

$$W(\varphi) = L \begin{cases} [1 + \beta I_1(\varphi)] \Psi_1(\varphi) & (-\pi < \varphi < 0) \\ [1 + \beta I_2(\varphi)] \Psi_2(\varphi) & (0 < \varphi < \pi) \end{cases}, \quad (7)$$

где

$$\Psi_i(\varphi) = \exp \int_0^\varphi \frac{2K_{1i}(x)}{K} dx, \quad I_i(\varphi) = \int_0^\varphi \frac{dx}{\Psi_i(x)}, \quad (i = 1, 2)$$

$$\beta = \frac{\Psi_2(\pi) - \Psi_1(-\pi)}{I_1(-\pi) \Psi_1(-\pi) - I_2(\pi) \Psi_2(\pi)},$$

$$L^{-1} = \int_{-\pi}^0 [1 + \beta I_1(\varphi)] \Psi_1(\varphi) d\varphi + \int_0^\pi [1 + \beta I_2(\varphi)] \Psi_2(\varphi) d\varphi.$$

В рассматриваемом случае

$$\Psi_1(\varphi) = \exp \left\{ D_0 \varphi - \frac{D}{n+1} \left[\left(\frac{2}{\pi} \right)^n \left(\varphi + \frac{\pi}{2} \right)^{n+1} - \frac{\pi}{2} \right] \right\},$$

$$\Psi_2(\varphi) = \exp \left\{ D_0 \varphi - \frac{D}{n+1} \left[\left(-\frac{2}{\pi} \right)^n \left(\varphi - \frac{\pi}{2} \right)^{n+1} + \frac{\pi}{2} \right] \right\},$$

где $D = 2\Delta/K$, $D_0 = 2\Delta_0/K$.

На рис. 2 приведены графики функции плотности вероятности $W(\varphi)$ при $D = 5$,

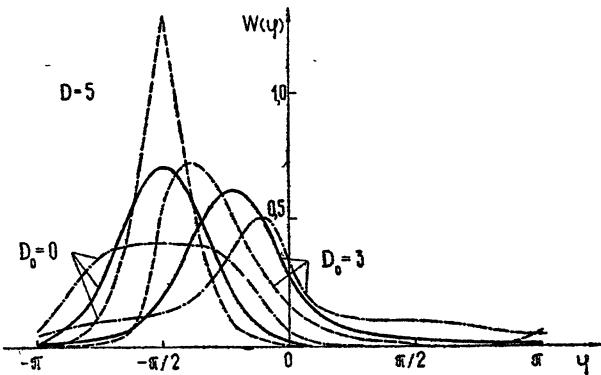


Рис. 2

$D_0 = 0$ и 3 , построенные* по формуле (7) для значений $n = 1/3, 1$ и 3 . Сравнение графиков показывает, что максимум и минимум распределения $W(\varphi)$ при $n = 1/3$ имеют, соответственно, большее и меньшее значения по сравнению с $n = 1$ и $n = 3$ при одинаковых и тех же D и D_0 .

Относительное среднее значение разности частот ПГ и ЭГ находится из (3) усреднением сначала по быстро меняющимся величинам, а затем по медленным [1]:

$$\frac{\langle \dot{\varphi} \rangle}{\Delta} = \frac{\Delta_0}{\Delta} - \int_{-\pi}^{\pi} F(\varphi) W(\varphi) d\varphi. \quad (8)$$

На рис. 3 даны графики $\frac{\langle \dot{\varphi} \rangle}{\Delta} = f\left(\frac{\Delta_0}{\Delta}\right)$ для $D = 1$ и 5 при $n = 1/3, 1$ и 3 . Из сравнения кривых при одном и том же D следует, что система ФАП при $n = 1/3$ имеет меньшую остаточную расстройку по частоте.

Проведенное рассмотрение позволяет сделать вывод о том, что с приближением формы характеристики ФД к прямоугольной (см. [6]) уменьшается «ширина» функции распределения $W(\varphi)$ (определенная шириной эквивалентного прямоугольника с высотой $W_{\max}(\varphi)$ и площадью единица) и уменьшается относительная средняя остаточная расстройка по частоте $\frac{\langle \dot{\varphi} \rangle}{\Delta}$. Если интерпретировать уравнение (3) как уравнение движения точки в поле регулярных внешних сил при случайных возмущениях [1, 7], то полученные результаты получают следующее простое объяснение. При переходе к нелинейности прямоугольного вида потенциальные «ямы» видоиз-

меняются так, что уменьшаются «разброс по дну ямы» и частота «выбивания» точки из «ямы». В этом смысле бесфильтровую систему ФАП с прямоугольной характеристикой ФД можно назвать более помехоустойчивой.

Автор благодарит М. Л. Тая за обсуждение и полезные советы.

ЛИТЕРАТУРА

1. В. И. Тихонов, Автоматика и телемеханика, 20, № 9, 1188 (1959).
2. В. И. Тихонов, Автоматика и телемеханика, 21, № 3, 301 (1960).

* Численный счет проведен на ЭВМ БЭСМ — ЗМ.

3. В. И. Тихонов, Б. И. Шахтарин, Автоматика и телемеханика, 26, № 9, 1563 (1965).
4. Витерби, ТИИЭР, 51, № 12, 1704 (1963).
5. В. Шахгильдян, А. А. Яховкин, Фазовая автоподстройка частоты, Связь, М., 1966.
6. Б. Э. Рахимов, Ученые записки, вып. 105, сер. радиофизическая, Горький, 1970.
7. Р. Л. Стратонович, Избранные вопросы теории флюктуации в радиотехнике, изд. Сов. радио, М., 1961.
8. Э. М. Хазен, Теория вероятностей и ее применения, 6, вып. 1, 130 (1961).
9. М. Л. Тай, Изв. высш. уч. зав. — Радиофизика, 8, № 2, 341 (1965).

Научно-исследовательский институт
прикладной математики и кибернетики
при Горьковском университете

Поступила в редакцию
8 апреля 1970 г.

УДК 531.391.314 + 621.396.6.078.6

КРИТЕРИЙ СИНХРОНИЗАЦИИ ДЛЯ СИСТЕМЫ ФАЗОВОЙ АВТОПОДСТРОЙКИ С НЕЛИНЕЙНЫМ ФИЛЬТРОМ

Ю. Н. Бакаев

Рассмотрим систему фазовой автоподстройки с нелинейным интегрирующим $\mathcal{R}C$ фильтром [1]. Согласно [2], ее дифференциальные уравнения можно записать в следующем виде:

$$\begin{aligned} \lambda \frac{d\varphi}{d\tau} &= u - \gamma, \\ \frac{G(u)}{\lambda} \frac{du}{d\tau} &= -F(\varphi) - u. \end{aligned} \quad (1)$$

В сравнении с [2] в (1) φ заменено на $(-\varphi)$ и F на $(-F)$, где $F(\varphi)$ — периодическая с периодом 2π , непрерывная и непрерывно дифференцируемая функция, $G(u)$ — непрерывная, всюду положительная функция; γ, λ — постоянные параметры, причем $\lambda > 0$.

Физический смысл переменных и параметров следующий (более подробно см. [2]): φ — разность фаз двух колебаний, поступающих на фазовый детектор (чувствительный элемент системы); u — напряжение, развиваемое фазовым детектором, нормированное по отношению к его максимально возможному значению; λ — коэффициент затухания системы; γ — начальная расстройка, отнесенная к максимально возможной расстройке, создаваемой управляющим элементом; τ — безразмерное время.

При этом $F(\varphi)$ есть уравнение характеристики фазового детектора, нормированной по максимуму; $G(u)$ — уравнение нормированной характеристики нелинейной емкости C : $G(u) = C(u)/C(0)$. Примерный вид функций $F(\varphi)$ и $G(u)$ показан на рис. 1, 2.

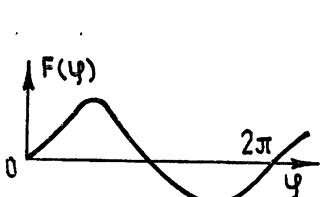


Рис. 1

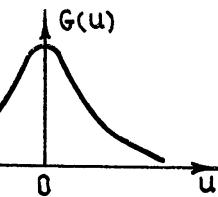


Рис. 2

С весьма общих позиций дифференциальные уравнения с нелинейностями периодического вида изучались в [3].

В [1, 2] было выполнено качественное исследование при кусочно-линейных аппроксимациях функций $F(\varphi)$ и $G(u)$. В [2], кроме того, рассматривается случай $F(\varphi) = \sin \varphi$, $G(u) = (1 + u^2 H^{-2})^{-1}$.

Цель заметки состоит в том, чтобы выделить более широкий класс нелинейностей, в пределах которого может быть сформулирован достаточный критерий устойчивости решений (1) при любых начальных возмущениях.

Будем считать, что уравнение $F(\varphi) - \gamma = 0$ имеет на полуинтервале $[0, 2\pi]$ два (и только два) действительных корня. Тот из них, для которого $F'(0) > 0$, обозначим

через φ_0 , второй — φ_1 . Выполним в (1) замену $\varphi - \varphi_0 = \psi$, $u - \gamma = y$. Обозначив $G(y + \gamma) = G_1(y)$, $F(\psi + \varphi_0) - \gamma = \Phi(\psi)$, получим вместо (1)

$$\begin{aligned} \lambda \frac{d\psi}{d\tau} &= y, \\ \frac{G_1(y)}{\lambda} - \frac{dy}{d\tau} &= -\Phi(\psi) - y. \end{aligned} \quad (2)$$

Корнями уравнения $\Phi(\psi) = 0$ будут $\psi = 0$ и $\psi_1 = \varphi_1 - \varphi_0$; без ограничения общности можно считать, что $\int_0^{2\pi} \Phi(\psi) d\psi = \mu < 0$. Противоположный случай сводится к предыдущему заменой ψ на $(-\psi)$ и y на $(-y)$.

Введем в рассмотрение функцию V :

$$V = \lambda \int_0^{\psi} \Phi(\psi) d\psi + \frac{1}{\lambda} \int_0^y y G_1(y) dy, \quad (3)$$

такую, чтобы она была неограниченной по координате y , т. е.

$$\int_0^y y G_1(y) dy \rightarrow \infty \quad (y \rightarrow \infty). \quad (4)$$

На интервале $(\psi_1, 2\pi - \psi_1)$ функция V является положительно определенной в общепринятом смысле, непрерывной; ее производная по времени в силу уравнений (2)

$$\frac{dV}{d\tau} = -y^2 \quad (5)$$

знакоотрицательна, причем уравнение $y^2 = 0$ не определяет целых траекторий, кроме состояний равновесия $\psi = 0$ и $\psi = \psi_1$.

Продолжим функцию V периодически в другие части фазовой плоскости. Построенная таким образом функция V будет при $\mu = 0$ непрерывна всюду. При этом, согласно (4), (5), гарантируется асимптотическая устойчивость решений (2) в целом. Однако для приложений случай $\mu = 0$, не представляет сколько-нибудь существенного интереса. Если же $\mu < 0$, то функция V будет на прямых $\psi = \psi_1 \pm 2k\pi$ ($k = 0, 1, 2, \dots$), испытывать разрыв непрерывности первого рода, и убывание функции V вдоль траекторий системы (2) уже не гарантирует устойчивости ее решений в целом, если не наложить дополнительных ограничений.

Скачок функции V будет отрицательным в нижней полуплоскости ($y < 0$) и положительным в верхней. Поэтому далее достаточно рассмотреть только случай, когда $y > 0$. Докажем следующее утверждение: если на прямой $\psi = \psi_1 - 2\pi$ найдется такая точка M с положительной ординатой ($y > 0$), определяемая ей траектория проходит через точку $\psi = \psi_1$, $y = 0$, то решения (2) асимптотически устойчивы в целом.

Обозначим в целях доказательства траекторию, определяемую точкой M , через $y_*(\psi)$, а через $y_1(\psi)$ — траекторию, определяемую другой точкой, такой, что $y_1(\psi_1 - 2\pi) > y_*(\psi_1 - 2\pi)$. Для соответствующих функций V , согласно (5), будем иметь

$$\frac{d(V_1 - V_*)}{d\psi} = -(y_1 - y_*). \quad (6)$$

Так как траектории на фазовой плоскости не пересекаются, и, следовательно, $y_1(\psi) > y_*(\psi)$, то из (6) следует, что конечное приращение V_1 в силу преобразования

M_1 в P_1 (см. рис. 3), осуществляющегося траекторией $y_1(\psi)$, суть величина меньшая, чем приращение функции V_* в силу преобразования M в P . Согласно (3), такое же соотношение сохраняется для положительно определенной функции

$$v = \int_0^y y G_1(y) dy.$$

Заметим теперь, что левый конец траектории $y_*(\psi)$ принадлежит поверхности $v = \delta$, где $\delta > 0$, тогда как правый — поверхности $v = 0$. Таким образом, приращение функции v в силу точечного преобразования прямой $\psi = \psi_1 - 2\pi$ в прямую $\psi = \psi_1$

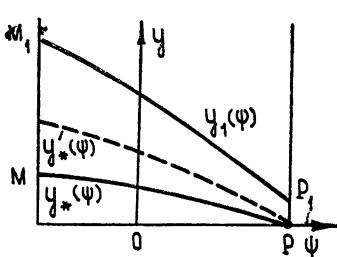


Рис. 3

удовлетворяет условию $\Delta v < c < 0$, так что (с учетом (4), (5) и того, что $\lambda > 0$) удовлетворяются все условия достаточного критерия устойчивости из [4]. Практическую проверку критерия (в том числе при реализации на ЭВМ) проще производить, если рассматривать уравнение (2) при обратном течении τ (соответственно, если в (2) принять ψ за новую независимую переменную,—при обратном течении ψ).

Рассмотрим далее две точки в пространстве параметров μ , λ . Пусть одной из них (λ_1, μ_1) соответствует траектория $y_*(\psi)$, тогда как другой (λ_2, μ_2) — траектория $y'_*(\psi)$ (рис. 3). Из факта положительной определенности функции v следует, что менее интенсивное убывание функции v имеет место для траектории $y_*(\psi)$. Иными словами, в рамках сформулированного выше критерия приближению к границе устойчивости в пространстве параметров соответствует перемещение конечной (при обратном течении τ) точки траектории $y_*(\psi)$ в направлении оси абсцисс. Тогда предыдущим рассмотрением применительно к системе (2) оправдывается известный прием определения критического соотношения параметров путем выделения особой траектории, идущей из седла $\psi_1 = 2\pi$, 0 в седло $\psi_1 = 0$.

Отметим в заключение, что система (2) может быть сведена к одному дифференциальному уравнению вида $\ddot{\psi} + R(\psi, \dot{\psi}) + \Phi(\psi) = 0$, которое обстоятельно изучалось Табуевой и Вдовиной (сводку результатов и ссылки см. в [3]), однако в рассматривающем здесь случае используемые ими условия $yR(x, y) > 0$ и $\frac{\partial R(x, y)}{\partial y} > 0$ могут нарушаться.

ЛИТЕРАТУРА

1. М. В. Капранов, В. А. Иванов, Н. Н. Иванова, Радиотехника и электроника, 5, № 11, 1774 (1960).
2. В. Н. Белых, В. Д. Шалфеев, Изв. высш. уч. зав—Радиофизика, 11, № 3, 407 (1968).
3. Е. А. Барбашин, В. А. Табуева, Динамические системы с цилиндрическим fazовым пространством, изд. Наука, М., 1969.
4. Ю. Н. Бакаев, Изв. АН СССР, ОТН, Энергетика и автоматика, № 2, 132 (1960).

Поступила в редакцию
30 декабря 1969 г.