

ВЕСТНИК

ЮЖНО-УРАЛЬСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА

№3 (12)
2002

СЕРИЯ

«МАТЕМАТИКА,
ФИЗИКА, ХИМИЯ»

Выпуск 2

Редакционная коллегия

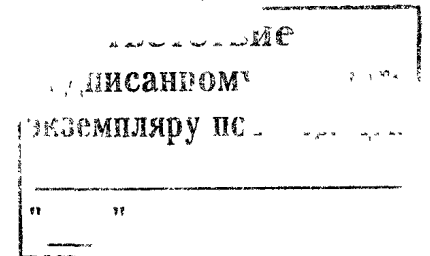
д.т.н., профессор Михайлов Г.Г. (отв. редактор),

д.ф.-м.н., профессор Бескачко В.П. (отв. секретарь),

д.т.н., профессор Гуревич С.Ю., д.ф.-м.н., профессор Менихес Л.Д.,

к.ф.-м.н., профессор Заляпин В.И., д.х.н., профессор Лыкасов А.А.,

д.х.н., профессор Рябухин А.Г.



СОДЕРЖАНИЕ

Математика

- ЗАЛЯПИН В.И., КРИВОЩАПОВ В.А. Численный анализ одной обратной задачи
прикладной биофизики 3
- ГРИГОРЬЕВ С.М. Асимптотические разложения функций Люстерника 11
- УХАНОВ М.В., ШИРЯЕВ В.И. Алгоритмы построения информационных множеств
для реализации минимаксного фильтра 19
- КУВШИНОВ Б.М., ШИРЯЕВ О.В. Метод комитетов в задачах распознавания
образцов в условиях неопределенности априорной информации 34

Физика

- ГУРЕВИЧ С.Ю., ГОЛУБЕВ Е.В., ХАБИРОВ К.Б. Лазерная генерация импульсов
поверхностных акустических волн в ферромагнетике при магнитном фазовом
переходе 44

ПРОКОПЬЕВ К.В., ШУЛЬГИНОВ А.А., ГУРЕВИЧ С.Ю. Структура электромагнитного поля, наведенного над поверхностью ферромагнитного полупространства, при падении сферического фронта акустической волны	50
ПЕТРОВ Ю.В., ШУШАРИН А.В., БОЖКО В.Ф. Экспериментальные исследования продольных ультразвуковых волн, возбуждаемых термоакустическим излучателем ...	56
КОРЕНЧЕНКО А.Е., БЕСКАЧКО В.П. Особенности установления колебаний в крутильном вискозиметре	62
БЕСКАЧКО В.П., КОРЕНЧЕНКО А.Е. Течение, возбуждаемое диском, вращающимся на поверхности жидкости	68
ГУРЕВИЧ С.Ю., ПЕТРОВ Ю.В., ГОЛУБЕВ Е.В., ХАБИРОВ К.Б. Исследование поверхностных акустических волн, возбуждаемых лазерным импульсом в ферромагнетиках при температуре магнитного фазового перехода	71

Химия

КАЧУРИНА О.И., ЛЕОНОВИЧ Б.И., ПЛАТОНОВА О.В. Термодинамика углерода и кислорода в железе	74
КАТКОВ А.Э., ЛЫКАСОВ А.А. Равновесие твердого шпинельного раствора $Fe_3O_4-CuFe_2O_4$ с делафосситом и гематитом	84
ДРОЗИН А.Д., РОЦИН В.Е. Термодинамика образования зародышей продуктов гетерофазных химических реакций в жидких растворах	90

ЧИСЛЕННЫЙ АНАЛИЗ ОДНОЙ ОБРАТНОЙ ЗАДАЧИ ПРИКЛАДНОЙ БИОФИЗИКИ

В.И. Заляпин, В.А. Кривошапов

Рассматривается проблема восстановления динамики поступления радионуклидов по результатам прижизненного измерения содержания ^{90}Sr в эмали передних зубов. Хорошо известное биофизикам соотношение, связывающее измерения поверхностной β -активности эмали передних постоянных зубов с характеристиками процесса обмена ^{90}Sr , сведено к системе нелинейных интегральных уравнений. Основным методом исследования последней – численная процедура нелинейной регуляризации с использованием априорной информации об отыскиваемых решениях.

Введение

Процесс наращивания производства плутония на ПО «Маяк», а также отсутствие надежных технологий переработки и хранения радиоактивных отходов привели к тому, что с 1949 г. часть сточных вод, содержащих радиоактивные вещества, удалялась в реку Теча, причем до июля 1951 г. сбросы были практически бесконтрольными. Это привело к сильному радиоактивному загрязнению речной системы (воды, донных отложений, пойменных почв), а жители прибрежных населенных пунктов подверглись как внешнему (за счет повышенного вблизи реки гамма фона), так и внутреннему облучению от смеси радионуклидов, поступающих в организм с водой и продуктами питания местного производства [1].

Жидкие радиоактивные отходы, удаляемые в реку Теча, представляли собой смесь радионуклидов стронция ($^{89,90}\text{Sr}$), цезия (^{137}Cs), ниобия (^{95}Nb), рутения ($^{103,106}\text{Ru}$), циркония (^{95}Zr) и изотопов редкоземельных элементов. Причем около 25 % суммарной активности приходилось на долю долгоживущих радионуклидов: ^{90}Sr и ^{137}Cs (период полураспада около 30 лет). Сбрасываемые с отходами радиоактивные элементы частично оседали на дне водоема, сорбировались донными отложениями и поглощались биомассой водных растений. Вследствие этих процессов основная активность задержалась в верхнем участке реки, а концентрация радионуклидов резко уменьшалась с увеличением расстояния от места сброса. После прекращения сброса жидких отходов и распада сравнительно короткоживущих радионуклидов основным компонентом радиоактивного загрязнения речной воды стали ^{90}Sr и ^{137}Cs . Цезий-137 достаточно быстро выводится из организма после прекращения поступления (период полувыведения около 100 суток). ^{90}Sr , напротив, прочно фиксируется в кальцифицированных тканях (кости, зубы) и продолжает облучать прилежащие ткани (красный костный мозг и костные поверхности) многие годы после прекращения поступления. Таким образом, основной вклад в дозу внутреннего облучения вносит ^{90}Sr . Знание уровней и динамики его поступления, позволяет рассчитать накопленную дозу облучения от стронция-90 на органы и ткани.

Как уже отмечалось, стронций-90 аккумулируется в костной ткани и тканях зуба, в том числе эмали, являющейся наиболее минерализованной частью скелета. Передние зубы (первые резцы) формируются достаточно быстро в период от 6 месяцев до 2 лет, а выведение ^{90}Sr из эмали происходит чрезвычайно медленно, со скоростью незначительно превышающей скорость радиоактивного распада (около 2,5 % в год).

Поэтому для восстановления динамики его поступления можно попробовать использовать результаты прижизненного измерения содержания ^{90}Sr в эмали передних зубов у лиц, проживавших весь период сбросов на загрязненной территории, и где река была основным источником питьевой воды.

Идея использовать для восстановления динамики поступления многочисленные (более 1000) измерения зубного детектора, выполненные в период 1960–1965 для людей различного возраста в населенном пункте Муслюмово, принадлежит Расину И.М. (1970 г.) [2]. Однако, полученное им решение не соответствует результатам прижизненных измерений содержания ^{90}Sr в организме, полученных после создания счетчика излучений человека – СИЧ-9.1 [3], начиная с 1974 г. В 1985 году эта обратная задача была вновь сформулирована и решена Кожеуровым В.П. [4], абсолютные уровни поступления были определены, исходя из результатов измерений СИЧ-9.1 и модели метаболизма стронция для взрослых (ICRP 20, 1973). Это решение также обладало значительными недостатками.

В последние годы в УНПЦ РМ была проведена ревизия данных измерений зубного детектора. Были пересмотрены и некоторые биологические параметры, используемые при решении обратной задачи. Вследствие этого возникла необходимость провести переоценку функции поступления ^{90}Sr на основе измерений зубного детектора, не меняя радикально методологические подходы к расчетам функции поступления

Восстановление динамики поступления ^{90}Sr в организм человека на основе измерений¹ $Y(T)$ – поверхностной β -активности эмали передних постоянных зубов, использует известную [5] связь между показаниями счетчика β -частиц и характеристиками процесса обмена ^{90}Sr :

$$Y(T) = \beta \int_{t_{\text{нач}}}^T \alpha(t-T, t) x(t) k(t-T) R(t-T, t_u - t) dt, \quad T_{\text{мин}} \leq T \leq T_{\text{макс}}. \quad (1)$$

Здесь β – масштабный коэффициент (количество отсчетов, приходящееся на единицу активности в зубах), $\alpha(t-T, t)$ – отношение поступления ^{90}Sr детям к поступлению взрослым; $x(t)$ – скорость поступления ^{90}Sr в момент времени t ; $k(t-T)$ – коэффициент перехода ^{90}Sr из желудочно-кишечного тракта в эмаль передних постоянных зубов; $R(t-T, t_u - t)$ – функция удержания, т.е. доля ^{90}Sr в эмали передних постоянных зубов по истечении некоторого времени после однократного поступления; t_u – момент измерения активности эмали зубов (в используемой базе данных $t_u = 1968$); $t_{\text{нач}}$ – момент начала поступления ^{90}Sr в организм.

Определению подлежат функции $x(t)$ и $k(\tau)$, описывающие динамику поступления ^{90}Sr в организм человека.

Отметим, что относительно основных параметров соотношения (1) известно следующее:

$$t_{\text{нач}} = \begin{cases} 1950, & T \leq 1950; \\ T, & T > 1950. \end{cases}$$

$$R(\tau, \Delta t) = \begin{cases} e^{-\lambda \Delta t}, & \Delta t > 0 \\ 0, & \Delta t \leq 0, \end{cases}$$

где λ – постоянная распада, для которой принято значение $\lambda = 0.025 \frac{1}{\text{год}}$.

Функция $\alpha(t-T, t)$ оценивалась экспертным путем и имеет вид, приведенный на рис. 1.

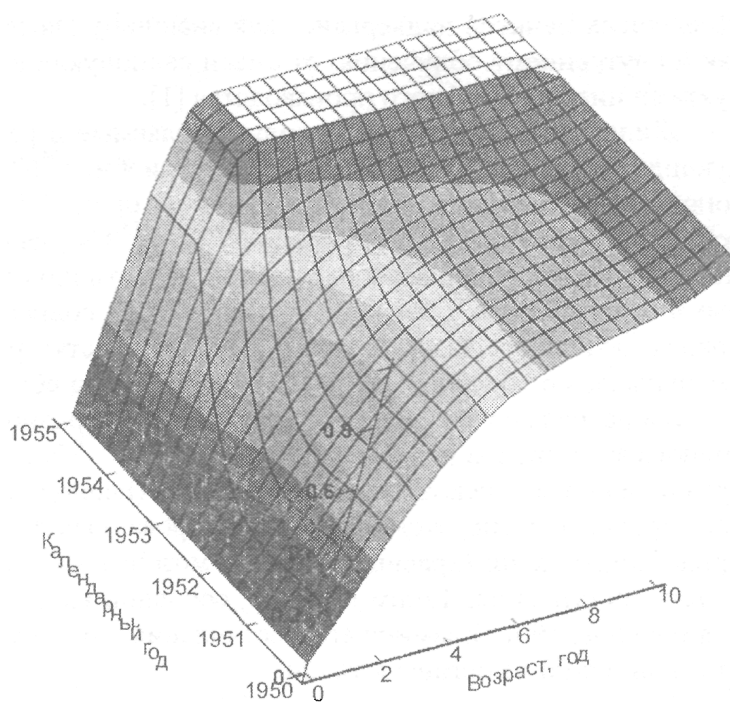


Рис. 1. Рельеф поверхности $\alpha(t-T, t)$

¹ С 1959 года в Уральском научно-практическом центре радиационной медицины проводились измерения β -активности эмали передних зубов с использованием счетчика Гейгера-Мюллера. За весь период наблюдений (с 1959 по 1998 гг.) было получено свыше 32 000 измерений для более чем 16 000 человек. Нами были использованы результаты измерений зубного детектора для жителей села Муслюмово, постоянно проживавших в указанном населенном пункте в период массовых сбросов.

1. Анализ основного интегрального соотношения

Область **D** изменения переменных в соотношении (1) описывается множеством точек (t, T) (рис. 2а) таких, что

$$D = \{(t, T): T \leq t \leq t_n, 1950 \leq T \leq T_{max}, 1950 \leq t \leq t_n, T_{min} \leq T \leq 1950\}.$$

Сменим начало отсчета, положив $t = t - 1950$ и $T = T - 1950$ (т.е. примем 1950 год за нулевой) и сделаем в интеграле (1) замену переменной интегрирования, приняв $t = t - T$. Для начального T_{min} и конечного T_{max} моментов рождения сохраним следующие обозначения:

$$T_{min} = T_{min} - 1950 = -T_{min}^0, T_{max} = T_{max} - 1950 = T_{max}^0.$$

Уравнение (1) примет вид

$$Y(T, t_n) = \beta \int_0^{t_n} \alpha(t-T, t)x(t)k(t-T)R(t-T, t_n-t)dt, \quad -T_{min} \leq T \leq 0; \quad (2)$$

$$Y(T, t_n) = \beta \int_T^{t_n} \alpha(t-T, t)x(t)k(t-T)R(t-T, t_n-t)dt, \quad 0 \leq T \leq T_{max}. \quad (2')$$

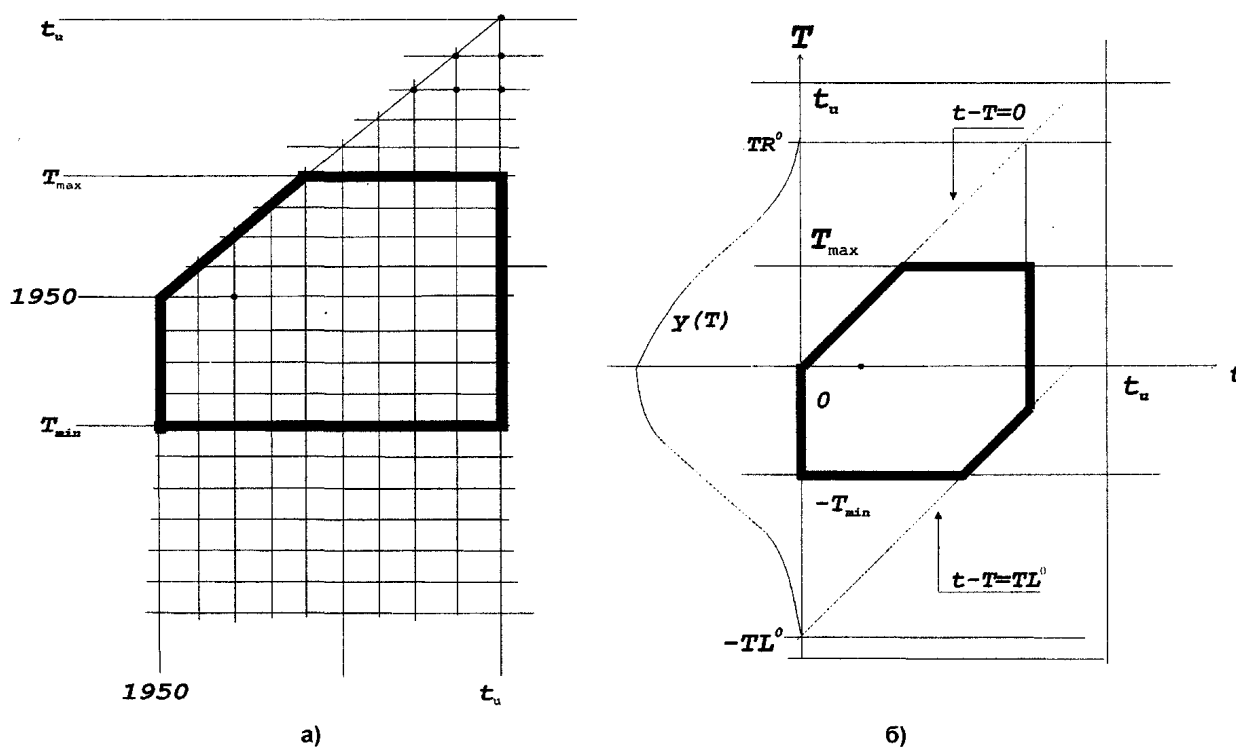


Рис. 2. Область изменения переменных

Известно [5], что функция экспериментальных наблюдений $Y(T, t_n)$ характеризуется быстрым убыванием при $T \rightarrow \pm\infty$ (рис. 3). Это позволяет указать такие два момента времени $-TL^0 > -\infty$ и $TR^0 \leq t_n$ – правый и левый соответственно, что

$$Y(T, t_n) = 0 \quad \forall T: T < -TL^0 \cap T > TR^0.$$

Но в силу неотрицательности подынтегральных функций в уравнениях (2) и убывания функции $Y(T, t_n)$ это влечет

$$x(t) = 0, \forall t: t \geq TR^0, k(\tau) = 0, \forall \tau \leq -TL^0.$$

Поэтому фактически интегрирование в уравнениях (2) осуществляется от 0 и до $\min\{TR^0, T + TL^0\}$ в первом и от T и до $\min\{TR^0, T + TL^0\}$ – во втором уравнении (рис. 2б).

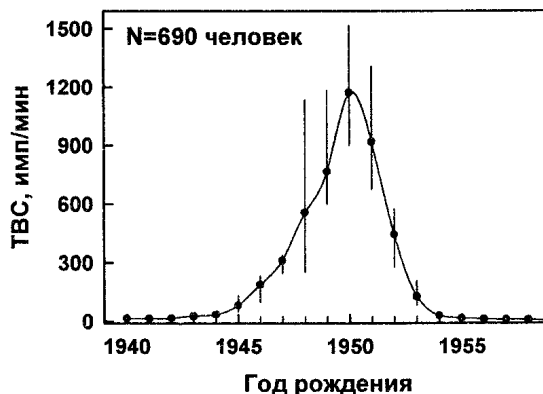


Рис. 3. Динамика изменения функции $Y(T)$

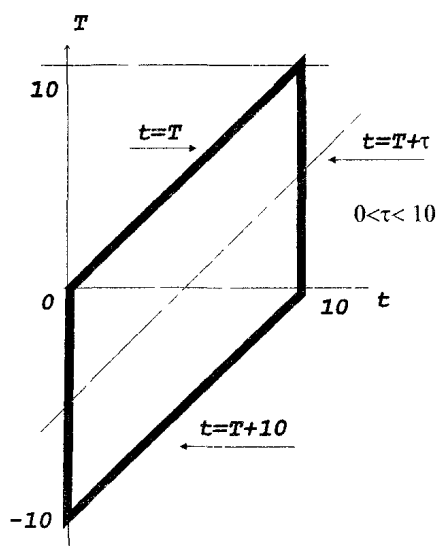


Рис. 4. Фактическая область изменения переменных

Имеющиеся экспериментальные данные таковы, что с учетом принятой оценки точности измерений значений функции $Y(T, t_n)$, можно считать $T_{\min} = TL^0 = 10$, $T_{\max} = TR^0 = 10$. Поэтому фактически область изменения переменных (t, T) описывается соотношениями (рис. 4):

$$D = \{(t, T):$$

$$T \leq t \leq 10; 0 \leq T \leq 10; 0 \leq t \leq T + 10; -10 \leq T \leq 0\}.$$

Уравнения (2) принимают вид

$$Y(T, t_n) = \begin{cases} \beta \int_0^{T+10} \alpha(t-T, t)x(t)k(t-T)R(t-T, t_n-t)dt, & -10 \leq T \leq 0; \\ \beta \int_T^{10} \alpha(t-T, t)x(t)k(t-T)R(t-T, t_n-t)dt, & 0 \leq T \leq 10. \end{cases} \quad (3)$$

Для дальнейшего нам будет удобно использовать другую форму записи этих уравнений.

Во-первых, напомним, что в силу принятых допущений,

$$R(t, \Delta t) = R(\Delta t) = e^{-\lambda \Delta t}.$$

В первом из уравнений (3) введем новую переменную U , положив $T + 10 = U$. Поскольку $-10 \leq T \leq 0$, то $0 \leq U \leq 10$. Уравнение примет вид

$$y(U-10) = \beta \int_0^U \alpha(t-U+10, t)x(t)k(t-U+10)e^{-\lambda(t_n-t)}dt, \quad 0 \leq U \leq 10.$$

Во втором уравнении заменим переменную интегрирования, положив $t = t - T$. Обозначая $10 - T = V$, приходим к уравнению

$$y(10-V) = \beta \int_0^V \alpha(t, t-V+10)x(t-V+10)k(t)e^{-\lambda(t_n-t+V-10)}dt, \quad 0 \leq V \leq 10.$$

Уравнения (3) теперь принимают вид

$$y(U-10) = \beta \int_0^U \alpha(t-U+10, t)x(t)k(t-U+10)e^{-\lambda(t_n-t)}dt, \quad 0 \leq U \leq 10; \quad (4)$$

$$y(10-V) = \beta \int_0^V \alpha(t, t-V+10)x(t-V+10)k(t)e^{-\lambda(t_n-t+V-10)}dt, \quad 0 \leq V \leq 10. \quad (4')$$

Введем следующие обозначения

$$e^{\lambda t_n} \cdot y(U-10) = \varphi(U), \quad e^{\lambda t_n} \cdot y(10-V) = \phi(V), \quad e^{\lambda t_n} \cdot x(t) = z(t).$$

В этих обозначениях предыдущие уравнения запишутся в форме

$$\varphi(U) = \beta \int_0^U \alpha(t-U+10, t)k(t-U+10)z(t)dt, \quad 0 \leq U \leq 10; \quad (5)$$

$$\phi(V) = \beta \int_0^V \alpha(t, t-V+10)k(t)z(t-V+10)dt, \quad 0 \leq V \leq 10. \quad (5')$$

каковую и будем считать окончательной.

2. Постановка задачи

Суммируя вышеизложенное, отметим, что задача состоит в определении из системы интегральных уравнений (5)–(5') неизвестных функций $z(t)$, $k(t)$, $0 \leq t \leq 10$ и неизвестной постоянной β .

Если предположить, что у такой задачи есть, по крайней мере, одно решение, то очевидно что для любых ненулевых постоянных c_1, c_2 набор $z_1 = c_1 \cdot z(t), k_1 = c_2 \cdot k(t), \beta_1 = \frac{1}{c_1 \cdot c_2} \cdot \beta$ также

будет решением.

С целью устранения отмеченной неоднозначности потребуем чтобы дополнительно выполнялось $\beta \cdot k(0) = 1$ и положив $k(t) = \frac{k(t)}{k(0)}$, придем к системе:

$$\varphi(U) = \int_0^U \alpha(t-U+10, t)k(t-U+10)z(t)dt, \quad 0 \leq U \leq 10; \quad (6)$$

$$\phi(V) = \int_0^V \alpha(t, t-V+10)k(t)z(t-V+10)dt, \quad 0 \leq V \leq 10; \quad (6')$$

$$k(0) = 1,$$

задача решения которой и есть основная задача настоящей работы.

3. Линеаризация

Введем в рассмотрение операторную форму записи задачи (6)–(6'):

$$Af = \Phi, \quad (7)$$

где $f = \begin{bmatrix} k(t) \\ z(t) \end{bmatrix}$ – векторная функция неизвестных; $\Phi = \begin{bmatrix} \varphi(U) \\ \phi(U) \end{bmatrix}$ – вектор правых частей системы

(6); $A = \begin{bmatrix} A_1(f) \\ A_2(f) \end{bmatrix}$ – нелинейный интегральный оператор.

Для решения задачи (7) применим итерационный метод Ньютона–Канторовича, позволяющий линеаризовать рассматриваемую задачу в окрестности некоторого начального приближения

$$f_0 = \begin{bmatrix} k_0 \\ z_0 \end{bmatrix}.$$

Подробнее, полагая

$$f = f_0 + \Delta f,$$

перейдем от нелинейного уравнения (7) относительно f к линейному уравнению относительно поправки Δf :

$$A'(f_0)\Delta f = -A(f_0) + \Phi. \quad (8)$$

Здесь $A'(f_0)$ – линейный оператор (производная оператора A). Фактически в рассматриваемой ситуации уравнение (8) представляет собой систему двух линейных интегральных уравнений относительно компонент $(\Delta k, \Delta z)$ векторной поправки Δf :

$$\begin{aligned} & \int_0^U \alpha(t-U+10, t)\Delta k(t-U+10)z_0(t)dt + \int_0^U \alpha(t-U+10, t)k_0(t-U+10)\Delta z(t)dt = \\ & = \varphi(U) - \int_0^U \alpha(t-U+10, t)k_0(t-U+10)z_0(t)dt; \end{aligned} \quad (9)$$

$$\begin{aligned} & \int_0^V \alpha(t, t-V+10)\Delta k(t)z_0(t-V+10)dt + \int_0^V \alpha(t, t-V+10)k_0(t)\Delta z(t-V+10)dt = \\ & = \phi(V) - \int_0^V \alpha(t, t-V+10)k_0(t)z_0(t-V+10)dt; \end{aligned} \quad (9')$$

$$\Delta k(0) = 0. \quad (9'')$$

Процедура решения системы (6)–(7) представляет собой итерационный процесс, начинающийся в некоторой начальной точке f_0 и реализуемый в соответствии с правилом $f_{N+1} = f_N + \Delta f$. Поправка Δf определяется на каждом шаге процедуры как решение линейной системы (9)–(9')–(9'').

4. Решение системы (9)–(9')–(9''). Регуляризация

Известно [6], что система (9) неустойчива по отношению к погрешностям задания правых частей и задача ее решения является некорректно поставленной по Тихонову А.Н. задачей.

Для ее решения используем процедуру регуляризации, заменив задачу решения системы (9) задачей минимизации [7] сглаживающего функционала

$$M^\alpha(\Delta f) = \|A' \Delta f + A - \Phi\|^2 + \alpha \cdot \Omega(\Delta f), \tag{10}$$

где $\|\bullet\|^2$ – L^2 -отклонение левых частей системы (9) от правых; Ω – стабилизатор первого порядка, задаваемый соотношением

$$\Omega(\Delta f) = \int_0^{10} [(\Delta k)^2 + (\Delta z)^2 + (\Delta k)^{\prime 2} + (\Delta z)^{\prime 2}] dt,$$

α – параметр регуляризации, который должен быть определен из уравнения

$$\|A' \Delta f_\alpha + A - \Phi\|^2 = \delta^2. \tag{10'}$$

В последнем соотношении Δf_α – решение задачи (10) как функция параметра регуляризации; δ – L^2 -точность задания правой части уравнения (8)–(9).

5. Основные расчетные соотношения

Численная процедура регуляризации (т.е. решение задачи минимизации функционала (10)) уравнения (9)–(9') на каждом шаге (N) итерационного процесса осуществлялась параллельно методом Хука–Дживса и методом Ритца, основанном на дискретизации задачи кусочно-линейными функциями с использованием локального базиса Лагранжа на промежутке $[0, 10]$.

Подробнее, зададим на промежутке $0 \leq t \leq 10$ сетку $(t, i = 0, 1, \dots, n)$ $0 = t_0 < t_1 < \dots < t_n = 10$ и определим $\Lambda_i(t)$ – кусочно-линейные функции, принимающие в узле с номером i значение 1, во всех прочих узлах – значение 0. Функции $\Lambda_i(t), i = 0, 1, \dots, n$ линейно независимы и образуют базис (базис Лагранжа) в пространстве кусочно-линейных функций на заданной сетке.

Положим

$$\Delta k(t) = \sum_0^n x_i \Lambda_i(t), \Delta z(t) = \sum_1^n y_i \Lambda_i(t).$$

Заметим, что при подобной дискретизации коэффициенты $x = (x_i), y = (y_i)$ представляют собой просто значение искомых функций в узлах сетки.

Функционал (10), скалярная форма которого

$$M^\alpha(\Delta k, \Delta z) =$$

$$\begin{aligned} &= \int_0^{10} \int_0^U \alpha(t-U+10, t) [\Delta k(t-U+10)z_0(t) + k_0(t-U+10)\Delta z(t) + k_0(t-U+10)z_0(t)] dt - \phi(U) \}^2 du + \\ &+ \int_0^{10} \int_0^U \alpha(t, t-U+10) [\Delta k(t)z_0(t-U+10) + k_0(t)\Delta z(t-U+10) + k_0(t)z_0(t-U+10)] dt - \phi(U) \}^2 du + \\ &+ \alpha \cdot \int_0^{10} [(\Delta k)^2 + (\Delta z)^2 + (\Delta k)^{\prime 2} + (\Delta z)^{\prime 2}] dU, \Delta k(0) = 0, \end{aligned}$$

становится квадратичной функцией относительно неизвестных (x, y)

$$M^\alpha(\Delta k, \Delta z) = M_n^\alpha(x, y) = Q_n^1(x, y) + Q_n^2(x, y) + \alpha \cdot \Omega_n(x, y). \tag{11}$$

Здесь

$$\begin{aligned} Q_n^1(x, y) &= \int_0^{10} [\sum_0^n m_i^1(U)x_i + \sum_1^n l_i^1(U)y_i + r^1(U)]^2 dU; \\ Q_n^2(x, y) &= \int_0^{10} [\sum_0^n m_i^2(U)x_i + \sum_1^n l_i^2(U)y_i + r^2(U)]^2 dU; \end{aligned}$$

$$\Omega_n(x, y) = \int_0^{10} [(\sum x_i \Lambda_i(U))^2 + (\sum y_i \Lambda_i(U))^2 + (\sum x_i \Lambda'_i)^2 + (\sum y_i \Lambda'_i)^2] dU;$$

$$m_i^1(U) = \int_0^U \alpha(t-U+10, t) z_0(t) \Lambda_i(t-U+10) dt, l_i^1(U) = \int_0^U \alpha(t-U+10, t) k_0(t-U+10) \Lambda_i(t) dt;$$

$$r^1(U) = \int_0^U \alpha(t-U+10, t) z_0(t) k_0(t-U+10) dt - \varphi(U).$$

Коэффициенты $m_i^2(U)$, $l_i^2(U)$ и $r^2(U)$ вычисляются аналогично.

Необходимые условия экстремума функции (11) по переменным $x_s, s=1, 2, \dots, n$; $y_j, j=0, 1, 2, \dots, n$, приводят к соотношениям:

$$\frac{\partial M_n^\alpha}{\partial x_s} = 0; \quad \frac{\partial M_n^\alpha}{\partial y_j} = 0.$$

Первые n из указанных соотношений дают:

$$\int_0^{10} [\sum m_i^1(U) x_i + \sum l_i^1(U) y_i + r^1(U)] \cdot m_i^1(U) dU + \int_0^{10} [\sum m_i^2(U) x_i + \sum l_i^2(U) y_i + r^2(U)] \cdot m_i^2(U) dU +$$

$$+ \alpha \cdot \int_0^{10} [\sum x_i \Lambda_i(U) \cdot \Lambda_i(U) + \sum x_i \Lambda'_i(U) \cdot \Lambda'_i(U)] dU = 0, \quad s=1, 2, \dots, n,$$

здесь суммирование по индексу i ведется от 1 до n (т.е. учтено, что $x_0 = 0$).

Остальные $(n+1)$ уравнение могут быть записаны в аналогичной форме

$$\int_0^{10} [\sum m_i^1(U) x_i + \sum l_i^1(U) y_i + r^1(U)] \cdot l_i^1(U) dU + \int_0^{10} [\sum m_i^2(U) x_i + \sum l_i^2(U) y_i + r^2(U)] \cdot l_i^2(U) dU +$$

$$+ \alpha \cdot \int_0^{10} [\sum y_i \Lambda_i(U) \cdot \Lambda_i(U) + \sum y_i \Lambda'_i(U) \cdot \Lambda'_i(U)] dU = 0, \quad s=0, 1, 2, \dots, n.$$

Таким образом, мы приходим к системе $(2n+1)$ линейных уравнений относительно неизвестных x, y , зависящих от параметра α :

$$\sum_{i=1}^n A_{vi}^1(\alpha) \cdot x_i + \sum_{i=0}^n A_{vi}^2 \cdot y_i = A_v^3, \quad s=1, 2, \dots, n;$$

$$\sum_{i=1}^n B_{vi}^1 \cdot x_i + \sum_{i=0}^n B_{vi}^2(\alpha) \cdot y_i = B_v^3, \quad s=0, 1, 2, \dots, n.$$

Коэффициенты этой системы даются соотношениями:

$$A_{vi}^1(\alpha) = \int_0^{10} [m_i^1(U) \cdot m_i^1(U) + m_i^2(U) \cdot m_i^2(U) + \alpha \cdot (\Lambda_i(U) \cdot \Lambda_i(U) + \Lambda'_i(U) \cdot \Lambda'_i(U))] dU;$$

$$A_{vi}^2 = \int_0^{10} [l_i^1(U) \cdot m_i^1(U) + l_i^2(U) \cdot m_i^2(U)] dU, \quad A_v^3 = - \int_0^{10} [r^1(U) \cdot m_i^1(U) + r^2(U) \cdot m_i^2(U)] dU;$$

$$B_{vi}^1 = \int_0^{10} [m_i^1(U) \cdot l_i^1(U) + m_i^2(U) \cdot l_i^2(U)] dU, \quad B_v^3 = - \int_0^{10} [r^1(U) \cdot l_i^1(U) + r^2(U) \cdot l_i^2(U)] dU;$$

$$B_{vi}^2(\alpha) = \int_0^{10} [l_i^1(U) \cdot l_i^1(U) + l_i^2(U) \cdot l_i^2(U) + \alpha \cdot (\Lambda_i(U) \cdot \Lambda_i(U) + \Lambda'_i(U) \cdot \Lambda'_i(U))] dU.$$

Решение интересующей нас задачи минимизации (11) – это решение системы (12), отвечающее значению параметра регуляризации α , определяемому из уравнения (10’):

$$Q_n^1(x(\alpha), y(\alpha)) + Q_n^2(x(\alpha), y(\alpha)) = \delta^2.$$

6. Результаты счета

Задача решалась на сетке 20×20 (шаг $\Delta t = 1/2$ года). Начальное приближение выбиралось исходя из эмпирических представлений о характере поступления. Восстановленная таким образом динамика поступления ^{90}Sr представлена на рис. 5.

Полученное решение достаточно хорошо описывает изучаемый процесс, при этом расчетное значение показаний счетчика β -частиц, получаемое при подстановке решения в исходный интеграл, хорошо согласуется с экспериментальными данными (рис. 6).

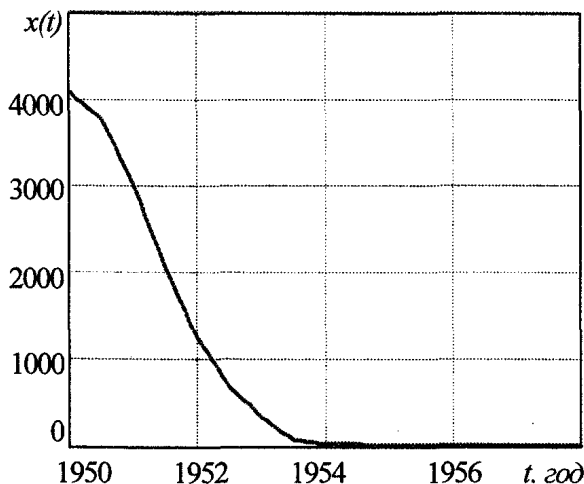


Рис. 5. Динамика поступления ^{90}Sr

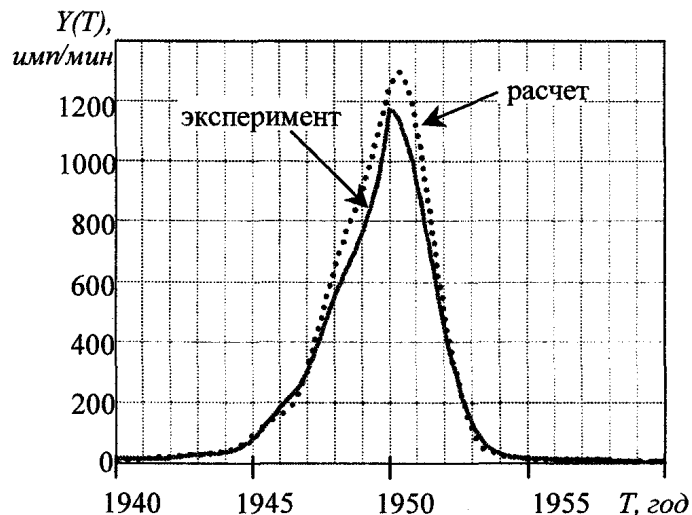


Рис. 6. Невязка

В заключение авторы приносят свою признательность коллективу лаборатории биофизики УНПЦ РМ за полезные и плодотворные обсуждения.

Работа была выполнена при финансовой поддержке Департамента по энергетике США и Федерального управления медико-биологических и экстремальных проблем при Министерстве здравоохранения РФ.

Литература

1. Медико-биологические и экологические последствия радиоактивного загрязнения реки Теча / Под ред. А.В. Аклеева, М.Ф. Киселева. – М.: Вторая типография ФУ «Медбиоэкстрем», 2000.
2. Расин И.М. Кинетика накопления стронция-90 и формирование тканевых доз в растущем организме. Дисс. ... канд. биол. наук. – М.: ИБФ, 1970.
3. Белле Ю.С., Ковтун А.Н., Кожеуров В.П., Козловская Л.А., Коровин В.И., Лебедев О.В., Пантелеев Л.И., Злобин Л.И. Спектрометр излучений человека СИЧ-9.1, позволяющий измерять низкие содержания инкорпорированного стронция-90 // *Мед. радиология*. – 1975. – Т. 20. – Вып. 6. – С. 52–58.
4. Kozheurov V.P.; Degteva M.O. Dietary Intake Evaluation and Dosimetric Modeling for the Techa River Residents Based on in vivo Measurements of Strontium-90 in Teeth and Skeleton // *Sci. Tot. Environ.* – 142: 63–72; 1994.
5. Degteva M.O., Kozheurov V.P., Tolstykh E.I. et al. The Techa River Dosimetry System: Methods for the reconstruction of internal dose // *Health Phys.* – 2000. – V. 79, № 1 – P. 24–35.
6. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. – М.: Наука, 1986.
7. Иванов В.К., Васин В.В., Танана В.П. Теория линейных некорректных задач и ее приложения. – М., Наука, 1978.

АСИМПТОТИЧЕСКИЕ РАЗЛОЖЕНИЯ ФУНКЦИЙ ЛЮСТЕРНИКА

С.М. Григорьев

Изучается предельное поведение пуассоновского случайного блуждания в R^n , порождающего широкий класс специальных функций. Для данного случайного блуждания получена предельная теорема и построен главный член асимптотического разложения функций Люстерника.

1. Обозначения

Нам будет удобно в дальнейшем для сокращения записи использовать следующие обозначения:

$$x, y, z, \dots \in R^l, x = \begin{pmatrix} x_1 \\ \dots \\ x_l \end{pmatrix}, \dots, x^y = \prod_i x_i^{y_i}, S(x) = \sum_i x_i, \Gamma(x) = \prod \Gamma(x_i).$$

Здесь $\Gamma(x)$ – гамма-функция Эйлера.

2. Пуассоновское блуждание в R^{n+1}

Рассмотрим $\{e_i\}_{i=1}^{n+1}$ – базис в R^{n+1} . Зададим случайное блуждание в R^{n+1} с помощью однородной переходной функции $P(x, y)$:

$$\forall x, y \in R^{n+1} : P(x, y) = P(0, y-x) = \begin{cases} p_j, & y-x = e_j, \\ 0, & y-x \neq e_j, \end{cases} \quad \sum_{i=1}^{n+1} p_i = 1. \quad (1)$$

Предполагается, что количество N_τ переходов за время τ , прошедшее с некоторого момента времени $\tau = 0$, регулируется пуассоновским процессом с параметром λ :

$$p\{N_\tau = s\} = e^{-\lambda\tau} \frac{(\lambda\tau)^s}{s!}. \quad (2)$$

Будем обозначать θ_τ – положение данного случайного блуждания в момент времени τ .

3. Орбиты в R^{n+1}

Векторы $x, y \in R^{n+1}$ назовем эквивалентными (и обозначим $x \equiv y$), если $y-x = (m, m, \dots, m)$, где $m \in Z$. Отношение $x \equiv y$, очевидно, является отношением эквивалентности и разбивает R^{n+1} на непересекающиеся классы: $R^{n+1} = \bigcup H_x$, где $H_x = \{y \in R^{n+1} | y \equiv x\}$. Для $\forall x, y \in R^{n+1}$ таких, что $x \not\equiv y$, выполнено $H_x \cap H_y = \emptyset$.

Обозначим Ω_{n+1} – целочисленную сеть в R^{n+1} , т.е. множество векторов с целочисленными в R^{n+1} координатами:

$$\Omega_{n+1} = \{x | x = (x_1, x_2, \dots, x_{n+1}), x_i \in Z, i = 1, 2, \dots, n+1\}.$$

Очевидно, что если $x \in \Omega_{n+1}$ и $x \equiv y$, то $y \in \Omega_{n+1}$. Таким образом, сеть Ω_{n+1} распадается в прямую сумму целочисленных классов эквивалентности: $\Omega_{n+1} = \sum_k H_k$, где классы

$$H_k = \{r \in \Omega_{n+1} | r = k + (m, m, \dots, m), m \in Z\}, \quad k - \text{целочисленный вектор.}$$

Далее всякий класс H_k будем называть *орбитой целочисленного вектора* $k \in R^{n+1}$.

4. Функции Люстерника

Случайное блуждание (1)–(2) индуцирует случайное блуждание по орбитам сети Ω_{n+1} , задаваемое однородной функцией перехода $P^*(H_k, H_r)$:

$$P^*(H_k, H_r) = \begin{cases} 0, & \exists m \in Z : k - r - (m, m, \dots, m) = e_i \\ p_i, & \exists m \in Z : k - r - (m, m, \dots, m) = e_i \end{cases} \quad (3)$$

Известно [2], что вероятность $P_\tau(H_k)$ для этого случайного блуждания за время τ попасть из начала координат на орбиту H_k точки k равна:

$$P_\tau(H_k) = P(\theta_\tau \in H_k) = e^{-\lambda\tau} G_k(\lambda\bar{p}\tau), \quad (4)$$

где $\bar{p} = (p_1, p_2, \dots, p_{n+1})$, а функции $G_k(\bar{z})$, называемые в дальнейшем функциями Люстерника, даются соотношением

$$G_k(\bar{z}) = \sum \frac{\bar{z}^{k+mA}}{(k+mA)!}, \quad A = (1, 1, \dots, 1). \quad (5)$$

(см. обозначения), в котором суммирование ведется по всем целым значениям параметра $m \in Z$. Функции $G_k(\bar{z})$, впервые определенные в [1], являются широким обобщением классических специальных функций математической физики – цилиндрических функций, в частности функций Бесселя, функций параболического цилиндра и др.

Целью настоящей заметки является получение главного члена асимптотического разложения функций (5).

Каждой орбите H_k поставим в соответствие вектор в R^n , координаты которого $\hat{k} = \{k_1 - k_{n+1}, \dots, k_n - k_{n+1}\}$.

При этом случайное блуждание (1)–(2) порождает случайное блуждание $\zeta_\tau = \sum_{i=1}^{N_\tau} \xi_i$ в R^n , где $\{\xi_i\}_{i=1}^\infty$ – независимые в совокупности, одинаково распределенные в R^n случайные векторы, ряд распределения которых

$$P\{\xi_i = x\} = \begin{cases} p_j, & x = e_j, & j = 1, 2, \dots, n \\ p_{n+1}, & x = (-1, -1, \dots, -1) = e_{n+1} \\ 0, & x \neq e_j, & j = 1, 2, \dots, n+1 \end{cases}.$$

Здесь вероятности p_j – такие же, как и в п. 2, $\{e_j\}_{j=1}^n$ – базис в R^n , $\sum_{i=1}^{n+1} e_i = 0$, N_τ – пуассоновская величина с параметром λ , не зависящая от ξ_i . Очевидно, имеет место

Лемма 1

Для случайного блуждания (3) по орбитам сети Ω_{n+1} справедливо:

$$P_\tau(H_k) = P(\zeta_\tau = \hat{k} = (k_1 - k_{n+1}, k_2 - k_{n+1}, \dots, k_n - k_{n+1})), \quad \forall k = (k_1, k_2, \dots, k_{n+1}) \in \Omega_{n+1}.$$

Доказательство

Рассмотрим события $C_m = \{\theta_\tau = k + (m, m, \dots, m)\}$. Событие C_m происходит, если за время τ произошло $k_i + m$ переходов в направлении e_i . Пусть C – событие $\{\theta_\tau \in H_k\}$. Тогда, если

$$k_0 = \min_{i=1-n+1} k_i, \text{ то } C = \sum_{m=-k_0}^{\infty} C_m.$$

Рассмотрим теперь события D_m , происходящие тогда, когда за время τ вектор ζ_τ образовался как сумма $\zeta_\tau = \sum_{i=1}^{N_\tau} \xi_i$, где среди слагаемых – по $k_j - k_{n+1} + m$ штук в направлении e_j ($j = 1 \div n + 1$), и событие $D = \{\zeta_\tau = \hat{k}\}$. Тогда, если $k'_0 = \max_i (k_{n+1} - k_i) = k_{n+1} - k_0$, то

$D = \sum_{m=k'_0}^{\infty} D_m = \sum_{m'=-k_0}^{\infty} D_{m'+k_{n+1}}$, где $m' = m - k_{n+1}$. Из описания случайного блуждания (1-2) и вектора ζ_{τ} очевидным образом вытекает, что $P(C_m) = P(D_{k_{n+1}+m})$. Отсюда:

$$P(C) = \sum_{m=-k_0}^{\infty} P(C_m) = \sum_{m=-k_0}^{\infty} P(D_{m+k_{n+1}}) = P(D),$$

что и требовалось доказать.

5. Некоторые предельные теоремы

Заметим, что описанный выше процесс удовлетворяет закону больших чисел и для него справедливы аналоги классических предельных теорем, важнейшими из которых для дальнейшего являются следующие.

Лемма 2 (предельная теорема для случайной суммы одинаково распределенных векторов)

Пусть случайные векторы $\xi_i \in R^n$ ($i=1,2,\dots$) независимы и одинаково распределены с рядом распределения $P\{\xi = e_i\} = p_i, i=1,2,\dots,n+1, \sum_i p_i = 1; e_i (i=1,\dots,n)$ – орты в $R^n, \sum_{i=1}^{n+1} e_i = 0, N_{\tau}$ – пуассоновская случайная величина с параметром $\lambda\tau$, не зависящая от $\bar{\xi}_i$. Обозначим

$\bar{p} = \sum_{i=1}^{n+1} p_i e_i = \sum_{i=1}^n (p_i - p_{n+1}) e_i$. Тогда при $\lambda\tau \rightarrow \infty$ распределение вектора $\bar{\zeta}_{\tau} = \frac{\sum_{i=1}^{N_{\tau}} \bar{\xi}_i - \lambda\bar{p}\tau}{\sqrt{\lambda\tau}}$ стремится к нормальному распределению с характеристической функцией

$$\omega(\bar{t}) \sim \exp\left[-\frac{1}{2} \sum_{i=1}^{n+1} p_i t_i^2\right]. \tag{6}$$

Доказательство

Пусть $\sum_{i=1}^{N_{\tau}} \bar{\xi}_i$ – характеристическая функция вектора ξ_i .

Из условия: $\varphi(\bar{t}) = \sum_{i=1}^{n+1} p_i e^{i\bar{t}e_i} = \sum_{i=1}^{n+1} p_i e^{t_i}$, где $t_{n+1} = -\sum_{i=1}^n t_i$.

Далее, характеристическая функция вектора $\sum_{i=1}^{N_{\tau}} \bar{\xi}_i$:

$$\psi(\bar{t}) = \sum_{s=0}^{\infty} \left(e^{-\lambda\tau} \frac{(\lambda\tau)^s}{s!} \right) \varphi^s(\bar{t}) = e^{-\lambda\tau} \cdot e^{\lambda\tau \varphi(\bar{t})} = \exp[\lambda\tau(\varphi(\bar{t}) - 1)].$$

Отсюда характеристическая функция вектора $\bar{\eta}_{\tau} = \left(\sum_{i=1}^{N_{\tau}} \bar{\xi}_i \right) - \lambda\tau\bar{p}$ такова:

$$\psi_0(\bar{t}) = e^{-\lambda\tau\bar{p}\bar{t}} \cdot \psi(\bar{t}) = \exp[\lambda\tau(\varphi(\bar{t}) - \bar{p}\bar{t} - 1)].$$

Здесь, очевидно, $\bar{p}\bar{t} = \sum_{i=1}^n (p_i - p_{n+1})t_i = \sum_{i=1}^{n+1} p_i t_i$.

Заметим, что $\varphi(t)$ раскладывается в ряд Тейлора:

$$\varphi(\bar{t}) = \sum_{i=1}^{n+1} p_i \left(1 + it_i - \frac{t_i^2}{2} + \dots \right) = \sum_{i=1}^{n+1} p_i + i \sum_{i=1}^{n+1} p_i t_i - \frac{1}{2} \sum_{i=1}^{n+1} p_i t_i^2 + \dots = 1 + \bar{p}\bar{t} - \frac{1}{2} \sum_{i=1}^{n+1} p_i t_i^2 + \dots$$

Отсюда

$$\varphi(\bar{t}) - \bar{p}\bar{t} - 1 = \sum_{s=2}^{\infty} \frac{i^s}{s!} \sum_{i=1}^{n+1} p_i t_i^s.$$

Но тогда для характеристической функции случайного вектора $\bar{\xi}_\tau = \frac{\left(\sum_{i=1}^{N_\tau} \xi_{s_i}\right) - \lambda\tau\bar{p}}{\sqrt{\lambda\tau}} = \frac{\eta_\tau}{\sqrt{\lambda\tau}}$ получаем:

$$\omega(\bar{t}) = \psi_0\left(\frac{\bar{t}}{\sqrt{\lambda\tau}}\right) = \exp\left[\lambda\tau\left(\sum_{s=2}^{\infty} \frac{i^s}{s!} \sum_{i=1}^{n+1} p_i \left(\frac{t_i}{\sqrt{\lambda\tau}}\right)^s\right)\right] = \exp\left[-\sum_{s=0}^{\infty} \left(\frac{i^s}{(s+2)!} \sum_{i=1}^{n+1} p_i \frac{t_i^{s+2}}{(\lambda\tau)^{s/2}}\right)\right].$$

При $\lambda\tau \rightarrow \infty$ имеем $\omega(\bar{t}) \sim \exp\left[-\frac{1}{2} \sum_{i=1}^{n+1} p_i t_i^2\right]$, что и требовалось.

Следующая лемма есть многомерный аналог известной леммы о выражении вероятностей дискретного распределения через характеристические функции (например, [3], с. 573).

Лемма 3 (формула обращения)

Пусть $\bar{\xi}$ – случайный n -мерный вектор, имеющий решетчатое дискретное распределение с шагами по осям h_1, h_2, \dots, h_n и одним из узлов $\bar{b}_0 = (b_1, b_2, \dots, b_n)$. Тогда вероятности $p_{\bar{k}} = P_{k_1, k_2, \dots, k_n}$ того, что вектор $\bar{\xi}$ примет значение $\bar{b}_{\bar{k}} = (b_1 + k_1 h_1, b_2 + k_2 h_2, \dots, b_n + k_n h_n)$, могут быть получены из соотношения

$$p_{\bar{k}} = \frac{h_1 h_2 \dots h_n}{(2\pi)^n} \int_{-\pi/h_1}^{\pi/h_1} \int_{-\pi/h_2}^{\pi/h_2} \dots \int_{-\pi/h_n}^{\pi/h_n} \varphi(\bar{t}) e^{-i\bar{b}_{\bar{k}} \bar{t}} d\bar{t}, \tag{7}$$

где $\varphi(\bar{t})$ – характеристическая функция вектора $\bar{\xi}$.

Доказательство

Характеристическую функцию вектора $\bar{\xi}$ можно записать в виде $\varphi(\bar{t}) = \sum_{\bar{k}} p_{\bar{k}} e^{-i\bar{b}_{\bar{k}} \bar{t}}$.

Отсюда:

$$\begin{aligned} \frac{h_1 h_2 \dots h_n}{(2\pi)^n} \int_{-\pi/h_1}^{\pi/h_1} \int_{-\pi/h_2}^{\pi/h_2} \dots \int_{-\pi/h_n}^{\pi/h_n} \varphi(\bar{t}) e^{-i\bar{b}_{\bar{k}} \bar{t}} d\bar{t} &= \frac{h_1 h_2 \dots h_n}{(2\pi)^n} \int_{-\pi/h_1}^{\pi/h_1} \int_{-\pi/h_2}^{\pi/h_2} \dots \int_{-\pi/h_n}^{\pi/h_n} \sum_{\bar{r}} \left[p_{\bar{r}} e^{i(\bar{b}_{\bar{r}} - \bar{b}_{\bar{k}}) \bar{t}} \right] d\bar{t} = \\ &= \frac{h_1 h_2 \dots h_n}{(2\pi)^n} \sum_{\bar{r}} p_{\bar{r}} I(\bar{k}, \bar{r}), \end{aligned}$$

где

$$I(\bar{k}, \bar{r}) = \int_{-\pi/h_1}^{\pi/h_1} \int_{-\pi/h_2}^{\pi/h_2} \dots \int_{-\pi/h_n}^{\pi/h_n} e^{i(\bar{b}_{\bar{k}} - \bar{b}_{\bar{r}}) \bar{t}} d\bar{t} = \int_{-\pi/h_1}^{\pi/h_1} e^{ih_1(k_1 - r_1)t_1} dt_1 \int_{-\pi/h_2}^{\pi/h_2} e^{ih_2(k_2 - r_2)t_2} dt_2 \dots \int_{-\pi/h_n}^{\pi/h_n} e^{ih_n(k_n - r_n)t_n} dt_n.$$

Заметим, что

$$\int_{-\pi/h_j}^{\pi/h_j} e^{ih_j(k_j - r_j)t_j} dt_j = \frac{1}{ih_j(k_j - r_j)} \left(e^{i\pi(k_j - r_j)} - e^{-i\pi(k_j - r_j)} \right) = \begin{cases} 0, & k_j \neq r_j, \\ 2\pi/h_j, & k_j = r_j, \end{cases}$$

что влечет за собой

$$I(\bar{k}, \bar{r}) = \begin{cases} 0, & \bar{k} \neq \bar{r} \\ (2\pi)^n, & \bar{k} = \bar{r} \\ h_1 h_2 \dots h_n, & \bar{k} = \bar{r} \end{cases}$$

Поэтому

$$\frac{h_1 h_2 \dots h_n}{(2\pi)^n} \sum_{\bar{r}} p_{\bar{r}} I(\bar{k}, \bar{r}) = p_{\bar{k}},$$

что и требовалось.

Теорема 1 (вариант предельной теоремы для случайной суммы в R^n)

Пусть случайные векторы $\bar{\xi}_i \in R^n$ ($i=1,2,\dots$) независимы и одинаково распределены с рядом распределения; $P\{\bar{\xi} = e_i\} = p_i$, $i=1,2,\dots,n+1$, $\sum_{i=1}^{n+1} p_i = 1$, e_i ($i=1,\dots,n$) – базис в R^n , $\sum_{i=1}^{n+1} e_i = 0$, N_τ – пуассоновская величина с параметром $\lambda\tau$, не зависящая от $\bar{\xi}_i$. Обозначим:

$$a = \frac{1}{d} - \frac{1}{d\sqrt{1+p_{n+1}d}}; \quad d = \sum_{i=1}^n \frac{1}{p_i}. \tag{8}$$

Тогда при $\lambda\tau \rightarrow \infty$ для вероятности $P_\tau(\bar{k}) = P\left\{\sum_{i=1}^{N_\tau} \bar{\xi}_i = \bar{k} = (k_1, k_2, \dots, k_n)\right\}$ справедливо соотношение

$$P_\tau(\bar{k}) \sim \frac{1}{\sqrt{(2\pi\lambda\tau)^n p_1 p_2 \dots p_{n+1} \left(\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_{n+1}}\right)}} \times \exp\left[-\frac{1}{2} \sum_{j=1}^n \frac{\left(\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{i=1}^n \frac{\lambda(p_i - p_{n+1})\tau - k_i}{p_i}\right)^2}{\lambda p_j \tau}\right]. \tag{9}$$

Доказательство

Пусть $\bar{p} = \sum_{i=1}^n (p_i - p_{n+1})e_i = \sum_{i=1}^{n+1} p_i e_i$. Рассмотрим вектор $\bar{\zeta}_\tau = \frac{\sum_{i=1}^{N_\tau} \bar{\xi}_i - \lambda\bar{p}\tau}{\sqrt{\lambda\tau}}$. Из условия теоремы очевидным образом вытекает, что вектор $\bar{\eta}_\tau$ имеет решетчатое распределение с узлом $(0,0,\dots,0)$ и периодами по каждой оси $h_i = 1$, поэтому $\bar{\zeta}_\tau$ имеет решетчатое распределение с узлом $\bar{b}_0 = -\bar{p}\sqrt{\lambda\tau}$ и периодами по каждой оси $h_i = \frac{1}{\sqrt{\lambda\tau}}$. В соответствии с леммой 3 заключаем:

$$P_{\bar{k}} = P\left\{\sum_{i=1}^{N_\tau} \bar{\xi}_i = \bar{k}\right\} = P\left\{\frac{\sum_{i=1}^{N_\tau} \bar{\xi}_i - \lambda\bar{p}\tau}{\sqrt{\lambda\tau}} = \bar{b}_{\bar{k}} = \frac{\bar{k} - \lambda\bar{p}\tau}{\sqrt{\lambda\tau}}\right\} = \frac{1}{(2\pi\sqrt{\lambda\tau})^n} \int_{-\pi\sqrt{\lambda\tau}}^{\pi\sqrt{\lambda\tau}} \dots \int_{-\pi\sqrt{\lambda\tau}}^{\pi\sqrt{\lambda\tau}} \omega(\bar{t}) e^{-i\bar{b}_{\bar{k}}\bar{t}} dt_1 \dots dt_n,$$

где $\omega(\bar{t})$ – характеристическая функция случайного вектора $\bar{\zeta}_\tau$. Тогда из леммы 2 следует:

$$P_\tau(\bar{k}) \sim \frac{1}{(2\pi\sqrt{\lambda\tau})^n} \int_{-\pi\sqrt{\lambda\tau}}^{\pi\sqrt{\lambda\tau}} \dots \int_{-\pi\sqrt{\lambda\tau}}^{\pi\sqrt{\lambda\tau}} \exp\left[-\sum_{i=1}^{n+1} \frac{p_i t_i^2}{2} + \frac{i\bar{t}}{\sqrt{\lambda\tau}} (\lambda\bar{p}\tau - \bar{k})\right] dt_1 \dots dt_n = \frac{1}{(2\pi\sqrt{\lambda\tau})^2} \iint_{D_\tau} \dots \int \exp\left[-\frac{1}{2} \left(\sum_{i=1}^{n+1} p_i t_i^2 - \frac{2i\bar{t}}{\sqrt{\lambda\tau}} (\lambda\bar{p}\tau - \bar{k})\right)\right] d\bar{t} = I, \tag{10}$$

где

$$D_\tau = \bigcap_{i=1}^n \left\{-\pi\sqrt{\lambda\tau} < x_i < \pi\sqrt{\lambda\tau}\right\}.$$

Для вычисления интеграла (10) перейдем к новым переменным:

$$\bar{t} = Q\bar{T}, \text{ где } Q = \|q_{ij}\|_{i,j=1}^n, \quad q_{ij} = \frac{\delta_{ij}}{\sqrt{p_i}} - \frac{a}{p_i\sqrt{p_i}}.$$

Таким образом, переменные t_1, t_2, \dots, t_n выражаются через переменные T_1, T_2, \dots, T_n в следующем виде:

$$t_i = \frac{T_i}{\sqrt{p_i}} - \frac{aS}{p_i},$$

где для краткости обозначено:

$$S = S(\bar{T}) = \sum_{i=1}^n \frac{T_i}{\sqrt{p_i}}, \quad t_{n+1} = -\sum_{i=1}^n t_i = -S(1-ad). \quad (11)$$

Отсюда:

$$\begin{aligned} \sum_{i=1}^{n+1} p_i t_i^2 &= \sum_{i=1}^n (t_i \sqrt{p_i})^2 + p_{n+1} t_{n+1}^2 = \sum_{i=1}^n \left(T_i - \frac{aS}{\sqrt{p_i}} \right)^2 + p_{n+1} S^2 (1-ad)^2 = \\ &= \sum_{i=1}^n \left(T_i^2 - \frac{2aS T_i}{\sqrt{p_i}} + \frac{a^2 S^2}{p_i} \right) + p_{n+1} S^2 (1-ad)^2 = \left(\sum_{i=1}^n T_i^2 \right) - 2aS \cdot S + a^2 S^2 \cdot d + p_{n+1} S^2 (1-ad)^2 = \\ &= \left(\sum_{i=1}^n T_i^2 \right) + S^2 \left[a^2 d(1+p_{n+1}d) - 2a(1+p_{n+1}d) + p_{n+1} \right] = \left(\sum_{i=1}^n T_i^2 \right), \end{aligned}$$

т.к. выражение в квадратных скобках в силу (8) равно 0. Далее:

$$\begin{aligned} -\frac{2i\bar{t}}{\sqrt{\lambda\tau}} (\lambda\bar{p}\tau - \bar{k}) &= -\frac{2i}{\sqrt{\lambda\tau}} \sum_{i=1}^n t_i (\lambda(p_i - p_{n+1})\tau - k_i) = -2i\sqrt{\lambda\tau} \left[\sum_{j=1}^n t_j p_j - p_{n+1} \sum_{j=1}^n t_j - \frac{1}{\lambda\tau} \sum_{j=1}^n k_j t_j \right] = \\ &= -2i\sqrt{\lambda\tau} \left[\sum_{j=1}^n (t_j \sqrt{p_j} - aS) - p_{n+1} S(1-ad) - \frac{1}{\lambda\tau} \sum_{j=1}^n \frac{k_j t_j}{\sqrt{p_j}} + \frac{aS}{\lambda\tau} \sum_{j=1}^n \frac{k_j}{p_j} \right] = \\ &= -2i \sum_{i=1}^n \frac{T_i}{\sqrt{\lambda p_i \tau}} \left[\lambda p_i \tau - k_i - a n \lambda \tau - \lambda p_{n+1} \tau (1-ad) + a \sum_{j=1}^n \frac{k_j}{p_j} \right] = \\ &= -2i \sum_{j=1}^n \frac{T_j}{\sqrt{\lambda p_j \tau}} \left[\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{j=1}^n \frac{\lambda(p_j - p_{n+1})\tau - k_j}{p_j} \right]. \end{aligned}$$

Отсюда:

$$\begin{aligned} \sum_{j=1}^{n+1} p_j t_j^2 - \frac{2i\bar{t}}{\sqrt{\lambda\tau}} (\lambda\bar{p}\tau - \bar{k}) &= \sum_{j=1}^n \left(T_j - \frac{i}{\sqrt{\lambda p_j \tau}} \left(\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{j=1}^n \frac{\lambda(p_j - p_{n+1})\tau - k_j}{p_j} \right) \right)^2 + \\ &+ \sum_{j=1}^n \frac{\left(\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{j=1}^n \frac{\lambda(p_j - p_{n+1})\tau - k_j}{p_j} \right)^2}{\lambda p_j \tau}. \end{aligned}$$

Тогда интеграл (10) равен:

$$\begin{aligned} I &= \frac{\det Q}{\sqrt{(2\pi\sqrt{\lambda\tau})}^n} \left(\iint \dots \int_{\Delta_\tau} \exp \left[-\frac{1}{2} \left(\sum_{j=1}^n \left(T_j - \frac{i}{\sqrt{\lambda p_j \tau}} \left(\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{j=1}^n \frac{\lambda(p_j - p_{n+1})\tau - k_j}{p_j} \right) \right) \right)^2 \right] d\bar{T} \right) \times \\ &\times e^{-\frac{1}{2} \sum_{j=1}^n \frac{\left(\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{j=1}^n \frac{\lambda(p_j - p_{n+1})\tau - k_j}{p_j} \right)^2}{\lambda p_j \tau}}. \end{aligned}$$

Интеграл, стоящий в скобках, есть интеграл Пуассона по области Δ_τ , которая является образом области D_τ при линейном невырожденном ($\det Q \neq 0$, см. ниже) преобразовании $\bar{t} = Q\bar{T}$. Но $\lim_{\lambda\tau \rightarrow \infty} D_\tau = R^n$, откуда аналогично $\lim_{\lambda\tau \rightarrow \infty} \Delta_\tau = R^n$. Поэтому интеграл в скобках при $\lambda\tau \rightarrow \infty$ стремится к $(\sqrt{2\pi})^n$. Отсюда

$$I \sim \frac{\det Q}{\sqrt{(2\pi\lambda\tau)^n}} \exp \left(-\frac{1}{2} \sum_{j=1}^n \frac{\left(\lambda(p_j - p_{n+1})\tau - k_j - a \sum_{l=1}^n \frac{\lambda(p_l - p_{n+1})\tau - k_l}{p_l} \right)^2}{\lambda p_j \tau} \right). \tag{12}$$

Осталось найти $\det Q$:

$$\begin{aligned} \det Q &= \det \left\| \frac{\delta_{ij}}{\sqrt{p_i}} - \frac{a}{p_i \sqrt{p_j}} \right\| = \frac{1}{\sqrt{p_1 p_2 \dots p_{n+1}}} \det \left\| \delta_{ij} - \frac{a}{p_i} \right\| = \\ &= \frac{a^n}{(p_1 p_2 \dots p_{n+1})^{3/2}} \det \left\| \frac{p_i \delta_{ij}}{a} - 1 \right\| = \frac{a^n}{(p_1 p_2 \dots p_{n+1})^{3/2}} \det \left\| \frac{p_i \delta_{ij}}{a} - \frac{p_n}{a} \delta_{nj} (1 - \delta_m) - \delta_m \right\| = \\ &= \frac{a^n}{(p_1 p_2 \dots p_{n+1})^{3/2}} \det \left\| \frac{p_i \delta_{ij}}{a} - \frac{p_n}{a} \delta_{nj} (1 - \delta_m) - \delta_m \delta_{nj} \sum_{i=1}^n \frac{p_n}{p_i} \right\| = \frac{a^n \cdot p_1 p_2 \dots p_{n-1} \cdot \left(\frac{p_n}{a} - \sum_{i=1}^n \frac{p_n}{p_i} \right)}{(p_1 p_2 \dots p_n)^{3/2}} = \\ &= \frac{a \cdot \left(\frac{1}{a} - d \right)}{\sqrt{p_1 p_2 \dots p_n}} = \frac{1 - ad}{\sqrt{p_1 p_2 \dots p_n}} = \frac{1}{\sqrt{p_1 p_2 \dots p_n (1 + p_{n+1} d)}} = \frac{1}{\sqrt{p_1 p_2 \dots p_{n+1} \left(\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_{n+1}} \right)}}, \end{aligned}$$

откуда автоматически следует (9).

Полученное соотношение может служить источником для получения главного члена асимптотического разложения функций Люстерника. Точнее, имеет место

Следствие (асимптотическая формула для функций Люстерника).

Для функций Люстерника $G_{\bar{k}}(\bar{x})$ (5) при $S(\bar{x}) \rightarrow \infty$ справедливо соотношение:

$$G_{\bar{k}}(\bar{x}) \sim \frac{1}{\sqrt{(2\pi)^n x_1 x_2 \dots x_{n+1} \left(\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_{n+1}} \right)}} \times \exp \left(S(\bar{x}) - \frac{1}{2} \sum_{j=1}^n \frac{\left(x_j - x_{n+1} - k'_j - a(\bar{x}) \sum_{l=1}^n \frac{x_l - x_{n+1} - k'_l}{x_l} \right)^2}{x_j} \right), \tag{9'}$$

где обозначено $k'_j = k_j - k_{n+1}$; $j = 1, 2, \dots, n$; $a(\bar{x}) = \frac{1}{\sum_{i=1}^n \frac{1}{x_i}} \left(1 - \frac{1}{\sqrt{\sum_{i=1}^{n+1} \frac{x_{n+1}}{x_i}}} \right)$. (8')

Доказательство

Формула (9') немедленно вытекает из (9) и из леммы 1, если учесть, что $S(\lambda\bar{p}\tau) = \lambda\tau$ и подставить в формулу (4) \bar{x} вместо $\lambda\bar{p}\tau$, тогда $p_i = x_i/S(\bar{x})$.

6. Пример

1. **Функции Люстерника одного переменного.** Пусть $\bar{x} = (x, x, \dots, x)$. Тогда

$$G_{\bar{k}}(\bar{x}) = \sum_{m \in \mathbb{Z}} \frac{x^{S(\bar{k}+mA)}}{(k+mA)!} = U_{\bar{k}}(x) - \text{функция Люстерника одного переменного, } A = [1, 1, \dots, 1]. \text{ Из (9')}$$

для $U_{\bar{k}}(x)$ имеем:

$$x_1 = x_2 = \dots = x_{n+1} = x; S(\bar{x}) = (n+1)x; a(\bar{x}) = \frac{x}{n} \left(1 - \frac{1}{\sqrt{n+1}} \right); \sum_{j=1}^n \frac{x_j - x_{n+1} - k'_j}{x_j} = -\frac{S(\bar{k}')}{x};$$

где $\bar{k}' = (k'_1, k'_2, \dots, k'_n)$. Отсюда:

$$\begin{aligned} \sum_{j=1}^n \frac{\left(x_j - x_{n+1} - k'_j - a(\bar{x}) \sum_{j=1}^n \frac{x_j - x_{n+1} - k'_j}{x_j} \right)^2}{x_j} &= \frac{1}{x} \sum_{j=1}^n \left(-k'_j + \frac{S(\bar{k}')}{n} \left(1 - \frac{1}{\sqrt{n+1}} \right) \right)^2 = \\ &= \frac{1}{x} \sum_{j=1}^n \left(k_j'^2 - 2 \frac{k'_j S(\bar{k}')}{n \sqrt{n+1}} (\sqrt{n+1} - 1) + \frac{(S(\bar{k}'))^2 (\sqrt{n+1} - 1)^2}{n^2 (n+1)} \right) = \\ &= \frac{1}{x} \sum_{j=1}^n k_j'^2 + \frac{1}{x} \sum_{j=1}^n (S(\bar{k}'))^2 \left(\frac{2 - 2\sqrt{n+1}}{n \sqrt{n+1}} + \frac{(\sqrt{n+1} - 1)^2}{n^2 (n+1)} \cdot n \right) = \\ &= \frac{1}{x} \left(|\bar{k}'|^2 + (S(\bar{k}'))^2 \left(\frac{2\sqrt{n+1} - 2n - 2 + n + 1 - 2\sqrt{n+1} + 1}{n(n+1)} \right) \right) = \frac{1}{x} \left(|\bar{k}'|^2 - \frac{(S(\bar{k}'))^2}{n+1} \right). \end{aligned}$$

Итак:

$$U_{\bar{k}}(x) \sim \frac{1}{\sqrt{(2\pi x)^n (n+1)}} \exp \left((n+1)x - \frac{1}{2x} \left(|\bar{k}'|^2 - \frac{(S(\bar{k}'))^2}{n+1} \right) \right) \quad (x \rightarrow \infty). \quad (13)$$

2. **Модифицированные функции Бесселя.** Для модифицированных функций Бесселя

$$I_k(x) = \sum_{m=-k}^{\infty} \frac{\left(\frac{x}{2}\right)^{k+2m}}{(k+m)!m!} = U_{[k,0]} \left(\frac{x}{2} \right) \text{ из (13) имеем: } n=1; |\bar{k}'| = k = S(\bar{k}'). \text{ Отсюда:}$$

$$I_k(x) = U_{[k,0]} \left(\frac{x}{2} \right) \sim \frac{1}{\sqrt{\left(2\pi \cdot \frac{x}{2}\right) \cdot 2}} \exp \left(2 \cdot \frac{x}{2} - \frac{1}{2 \cdot \frac{x}{2}} \left(k^2 - \frac{k^2}{2} \right) \right) = \frac{1}{\sqrt{2\pi x}} \exp \left(x - \frac{k^2}{2x} \right) \quad (x \rightarrow \infty). \quad (14)$$

Аналогичным образом из формулы (9') можно получить асимптотики для многих других функций.

В заключение автор выражает благодарность своему научному руководителю профессору В.И. Заляпину за постановку задачи, многочисленные обсуждения возникавших вопросов и результатов, а также за предоставленное техническое и информационное обеспечение.

Литература

1. Люстерник Л.А. Об одной задаче теории массового обслуживания и связанном с ней обобщении цилиндрических функций // ДАН СССР. – 1967. – Т. 177, № 5.
2. Заляпин В.И., Люстерник Л.А. О системе функций пуассоновского блуждания// ДАН СССР. – 1972. – Т. 207. – Вып. 1.
3. Феллер В. Введение в теорию вероятностей и ее приложения. – М.: Мир, 1984. – Т. II. – С. 738.

АЛГОРИТМЫ ПОСТРОЕНИЯ ИНФОРМАЦИОННЫХ МНОЖЕСТВ ДЛЯ РЕАЛИЗАЦИИ МИНИМАКСНОГО ФИЛЬТРА

М.В. Уханов, В.И. Ширяев

Статья посвящена разработке алгоритмов построения информационных множеств, заданных системами линейных неравенств, для нахождения в темпе реального времени минимаксной оценки состояния линейной динамической системы в условиях неопределенности возмущений.

Введение

Задача оценивания состояния динамических систем при наличии неопределенных, но ограниченных мешающих факторов, занимает важное место в теории идентификации. Начиная с пионерской работы F.C. Schweppe [1], этой задаче посвящается все возрастающее количество исследований, включая фундаментальные работы А.Б. Куржанского [2] и Ф.Л. Черноусько [3].

Первоначально выдвинутая F.C. Schweppe идея оценивания заключалась в построении компактного множества, гарантированно включающего оцениваемое состояние динамической системы. В качестве такого множества предлагалось использовать многомерный эллипсоид. Преимущество такого выбора состояло в том, что вычислительная сложность соответствующих алгоритмов оценивания не зависела от числа наблюдений (например, см. [4]). В работе [5] Г.М. Бакан предложил использовать метод динамического программирования для эллипсоидального оценивания.

Исследовались и другие варианты оценивающих множеств. Так, исходя из требований повышения точности оценивания, в [6] было предложено использовать многогранники, заданные своими вершинами. В работе М.М. Лычака [7] многогранник задан методом двойного описания (вершинами и гранями). Учитывая, что при больших объемах измерительной информации число вершин многогранников может значительно возрасти, что соответственно увеличивает вычислительные затраты, в [8] Е.К. Костоусовой был приведен вариант аппроксимации оценивающих множеств параллелотопами.

При реализации минимаксного фильтра на ЭВМ основную трудность представляет выполнение операций над множествами в темпе реального времени. Необходим компромисс между объемом данных для описания аппроксимации (объемом памяти ЭВМ), скоростью выполнения преобразований (временем выполнения на ЭВМ) с одной стороны и точностью построения аппроксимации (погрешностью вычислений) с другой. Более точные аппроксимации требуют больших вычислительных ресурсов и, следовательно, больше времени на выполнение операций.

Чем больше вычислительные мощности, тем с более сложной структурой оценок в виде многогранников можно работать, — и тогда тем меньшая степень закругления используется. В алгоритме эллипсоидальных оценок величина закругления не зависит от мощности вычислителя.

Вычислительная техника постоянно совершенствуется, что дает возможность решать все более трудоемкие задачи и, следовательно, позволяет сдвинуть компромисс между простотой и точностью в сторону точности. Поэтому в настоящей работе, развивающей идеи оценивания с использованием многогранников [9–11], информационные множества аппроксимированы системами линейных неравенств (гранями многогранников). А именно, рассматривается задача построения аппроксимации информационного множества в виде системы линейных неравенств. Такой выбор класса и способа задания аппроксимирующих областей по сравнению с представлением вершинами дает экономию памяти, а по сравнению с эллипсоидами — повышает точность. Когда множества заданы вершинами, алгоритмы преобразования информационных множеств получены И.С. Панченко [10]. При задании информационных множеств эллипсоидами результаты были получены Ф.Л. Черноусько [3], Г.М. Баканом [4, 5] и др.

1. Постановка задачи

Рассматривается задача теоретико-множественного оценивания состояния наблюдаемой линейной дискретной динамической системы, действующей в условиях неопределенности, в которой объект описывается линейным разностным уравнением

$$x_{k+1} = A_k x_k + w_k, \quad k = 0, 1, \dots, N-1. \quad (1)$$

Состояние системы на k -ом шаге характеризуется вектором состояния $x_k \in \mathfrak{R}^n$. Известные матрицы перехода состояний системы $A_k \in \mathfrak{R}^{n \times n}$ являются обратимыми.

В каждый дискретный момент времени состояние наблюдается по результатам зашумленных измерений, описываемых уравнением

$$y_k = G_k x_k + v_k, \quad k = 1, 2, \dots, N. \quad (2)$$

Здесь $y_k \in \mathfrak{R}^m$ – вектор измерения; $G_k \in \mathfrak{R}^{m \times n}$ – матрица измерений.

На начальное состояние $x_0 \in \mathfrak{R}^n$ и неопределенные воздействия $w_k \in \mathfrak{R}^n$, $v_k \in \mathfrak{R}^m$ наложены ограничения

$$x_0 \in X_0; \quad w_k \in W_k, \quad k = 0, \dots, N-1; \quad v_k \in V_k, \quad k = 1, \dots, N, \quad (3)$$

где X_0, W_k, V_k – заданные выпуклые компакты в соответствующих пространствах. Таким образом, информация о неопределенных возмущениях и начальном состоянии исчерпывается заданием выпуклых компактов.

Определение 1. Множеством достижимости X_k системы (1) называется множество всех точек $x_k \in \mathfrak{R}^n$, в которые эту систему можно перевести из X_0 за k шагов.

Определение 2. Информационным множеством \bar{X}_k состояния системы (1) при наличии измерений $y_k(\cdot) = \{y_1, \dots, y_k\}$ называется множество таких точек из множества достижимости X_k , которые по уравнениям (2) совместимы с измерениями при выполнении ограничений (3).

Информационное множество характеризует неопределенность знания вектора состояния на основе полученных данных наблюдений.

Многошаговая задача гарантированного оценивания текущего состояния системы (1)–(3) состоит в построении последовательности информационных множеств \bar{X}_{k+1} по рекуррентным соотношениям, полученным И.Я. Кацом и А.Б. Куржанским [12]

$$\begin{aligned} X_{k+1/k} &= A_k \bar{X}_k + W_k, \quad \bar{X}_0 = X_0, \\ X[y_{k+1}] &= \{x \in \mathfrak{R}^n \mid y_{k+1} = G_{k+1}x + v, v \in V_{k+1}\}, \\ \bar{X}_{k+1} &= X_{k+1/k} \cap X[y_{k+1}]. \end{aligned} \quad (4)$$

Здесь $X_{k+1/k}$ – множество прогнозов состояния системы на $k+1$ шаге (или, по-другому, область достижимости системы за один шаг). $X[y_{k+1}]$ – множество состояний, совместимых с измерением y_{k+1} .

Таким образом, задача гарантированного оценивания состоит в нахождении информационных множеств или их аппроксимаций.

Задача минимаксной фильтрации состоит в нахождении по известной реализации измерений $y_k(\cdot) = \{y_1, \dots, y_k\}$ для каждого $k \geq 1$ оценки x_k^* , удовлетворяющей условию (минимум ошибки в наихудшем случае)

$$\delta_k^* = \max_{x \in \bar{X}_k} \|x - x_k^*\| = \min_{c \in \mathfrak{R}^n} \max_{x \in \bar{X}_k} \|x - c\|. \quad (5)$$

Таким образом, по постановке задачи нам надлежит найти чебышевский центр информационного множества \bar{X}_k (δ_k^* – чебышевский радиус множества \bar{X}_k).

Для реализации на ЭВМ минимаксного фильтра системы (1)–(3) по уравнениям (4)–(5) необходимо выполнить следующие шаги:

- 1) выбрать класс областей и способ их задания для аппроксимации информационных множеств;
- 2) построить алгоритмы преобразования этих областей (операции сложения, пересечения и линейного преобразования) с хорошими временными и точностными характеристиками;
- 3) найти минимаксную точечную оценку x_k^* состояния системы по множественной оценке.

2. Алгоритм построения суммы многогранников в задачах минимаксной фильтрации

Определение 3. Суммой по Минковскому [13, с. 76] двух множеств P_1 и P_2 из n -мерного евклидова пространства \mathcal{R}^n называется множество $P \in \mathcal{R}^n$ такое, что

$$P = P_1 + P_2 \triangleq \{x \in \mathcal{R}^n \mid x = x_1 + x_2, x_1 \in P_1, x_2 \in P_2\}.$$

Будем рассматривать случай, когда суммируемые множества заданы системами линейных неравенств (многогранниками)

$$P_1 = \{x_1 \in \mathcal{R}^n \mid A_1 x_1 \leq b_1, A_1 \in \mathcal{R}^{m_1 \times n}, b_1 \in \mathcal{R}^{m_1}\},$$

$$P_2 = \{x_2 \in \mathcal{R}^n \mid A_2 x_2 \leq b_2, A_2 \in \mathcal{R}^{m_2 \times n}, b_2 \in \mathcal{R}^{m_2}\}.$$

Необходимо получить сумму P многогранников P_1 и P_2 также в виде системы линейных неравенств

$$P = \{x \in \mathcal{R}^n \mid Ax \leq b, A \in \mathcal{R}^{m \times n}, b \in \mathcal{R}^m\}.$$

Сумма по Минковскому имеет большое значение для геометрии многогранников [14], потому что применение этой операции к двум многогранникам вновь приводит к многограннику (в отличие от эллипсоидов, относительно суммы класс многогранников замкнут).

Построение суммы многогранников является самой трудоемкой по времени вычисления операцией при решении задачи минимаксной фильтрации состояния динамической системы [6, 9]. В случае, когда множества заданы эллипсоидами, алгоритм построения множеств, которые являются оценками сверху и снизу суммы, разработан Ф.Л. Черноусько [3].

В настоящей работе разработаны алгоритмы построения внешней и внутренней аппроксимации суммы в виде системы линейных неравенств.

Нахождение оценки снизу (алгоритм последовательных приближений). Учитывая, что суммируемые множества заданы системами линейных неравенств, из определения суммы получаем систему линейных неравенств и равенств относительно переменных $x \in \mathcal{R}^n$, $x_1 \in \mathcal{R}^n$ и $x_2 \in \mathcal{R}^n$

$$\begin{cases} x - x_1 - x_2 = 0; \\ A_1 x_1 \leq b_1; \\ A_2 x_2 \leq b_2. \end{cases} \quad (6)$$

Заменяем в системе (6) $x_1 \in \mathcal{R}^n$ на разность $(x - x_2)$ и получаем эквивалентную ей систему относительно переменных x и x_2

$$\begin{cases} A_1 x - A_1 x_2 \leq b_1; \\ A_2 x_2 \leq b_2. \end{cases}$$

Введем обозначения

$$G = \begin{pmatrix} A_1 & -A_1 \\ 0 & A_2 \end{pmatrix}, G_2 = \begin{pmatrix} -A_1 \\ A_2 \end{pmatrix}, g = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Матрица G имеет размер $(m_1 + m_2) \times 2n$; матрица G_2 – $(m_1 + m_2) \times n$; вектор g – $m_1 + m_2$, где m_1 и m_2 – количество граней в многогранниках P_1 и P_2 .

Запишем в матричной форме

$$A_1 x + G_2 x_2 \leq g. \quad (7)$$

Описание суммы P в виде системы линейных неравенств может быть получено путем исключения n компонент вектора x_2 из системы (7). Таким образом, способ задания множества его гранями (системой линейных неравенств) позволил заменить задачу построения суммы в \mathcal{R}^n на задачу построения ортогональной проекции многогранника из \mathcal{R}^{2n} в \mathcal{R}^n .

Исключение компонент вектора x_2 из системы (7) можно выполнить методом фундаментального сокращенного свертывания Н.Б. Черниковой [15]. При решении задачи минимаксной или гарантированной фильтрации важно получить информационное множество к заданному времени. Существенным недостатком методов свертывания является то, что в том случае, когда

процесс свертывания не удастся провести до конца, затраченные усилия не приводят даже к построению оценки информационного множества.

Наиболее приемлем для исключения вектора x_2 из системы (7) итерационный алгоритм последовательных приближений (АПП) для построения ортогональной проекции многогранника В.А. Бушенкова [16]. Опишем алгоритм [16] применительно к задаче построения суммы многогранников.

Шаг 1. Пусть задано неотрицательное число $\varepsilon(P)$, характеризующее требуемую точность описания суммы. Решим на системе ограничений (7) ряд задач линейного программирования. Причем, минимальное число таких задач – $n + 1$ (тогда можно получить выпуклый многогранник ненулевого объема). Это могут быть, например, задачи максимизации и минимизации по каждой координате вектора $x \in \mathcal{R}^n$. Проекция решений этих задач в пространство переменных x образуют множество $V^0 = \{v_1, v_2, \dots, v_{n+1}\}$. При проецировании возможно получение одинаковых вершин $v_i = v_j, i \neq j$. В этом случае решить еще ряд задач линейного программирования.

Шаг 2. Построим описание выпуклой оболочки H^0 векторов из V^0 в виде системы линейных неравенств. В том случае, когда число векторов в V^0 равно $n + 1$, легко построить выпуклую оболочку. Для этого отбрасываем одну точку и проводим гиперплоскость через n точек. Далее определяем знак неравенства, подставляя отброшенную точку. И так n раз. Положим индекс k равным нулю.

Шаг 3. Для всех неравенств системы

$$(c_i^k, z) \leq d_i^k, i = 1, 2, \dots, q^k, \quad (8)$$

описывающей H^k , вычислим значения сдвигов

$$\rho_i^k = [(c_i^k, x^*) - d_i^k] / \|c_i^k\|,$$

где (x^*, x_2^*) – решение задачи линейного программирования

$$(c_i^k, x) + (0, x_2) \rightarrow \max \text{ при ограничения (7)}. \quad (9)$$

Если неравенство было перенесено в систему из системы (8), описывающей H^{k-1} , и значение сдвига для него известно, то повторно находить это значение не требуется. Для неравенств с ненулевыми значениями сдвигов запоем проекции решений задач (9) в пространство переменных x .

Шаг 4. Величина $H(H^k, P) = \rho_{\max} = \max \rho_i^k, i = 1, 2, \dots, q^k$ характеризует достигнутую точность построения суммы. Если $H(H^k, P) \leq \varepsilon(P)$, то нижней оценкой суммы будет H^k и работа алгоритма заканчивается. Получаем описание суммы в виде системы линейных неравенств $Cx \leq d$. В противном случае вычисления продолжаются.

Шаг 5. Выберем из (8) неравенство с максимальным значением сдвига. Добавим к V^k проекцию решения задачи (9), соответствующую выбранному неравенству, и получим совокупность V^{k+1} .

Шаг 6. Найти выпуклую оболочку H^{k+1} совокупности векторов V^{k+1} в виде системы линейных неравенств. Увеличим значение k на единицу и вернемся к шагу 3 для выполнения следующей итерации алгоритма.

При решении задачи минимаксной или гарантированной фильтрации необходимо получить оценку информационного множества за какое-то определенное время (часто весьма ограниченное). Выполнение АПП можно прервать в любое время (если выполнена хотя бы одна итерация алгоритма). При этом будет известна внутренняя аппроксимация суммы H^k с минимальным числом вершин при заданной точности $\varepsilon(P)$. Если ограничений по времени нет, то за конечное число шагов при $\varepsilon(P) = 0$ АПП строит точную сумму многогранников, причем, для ее нахождения необходимо найти все ее вершины.

Замечание 1. Скорость построения суммы АПП зависит от двух факторов: размерности n пространства и количества неравенств системы (7). От первого фактора зависит число итераций алгоритма, необходимых для достижения заданной точности описания суммы, и объем исполь-

зуемой памяти. От второго фактора зависит время решения вспомогательных задач линейного программирования.

Замечание 2. Если необходимо построить сумму $P = \sum_{i=1}^S P_i$, то матрица G будет иметь специальный вид

$$G = \begin{pmatrix} A_1 & -A_1 & -A_1 & \cdots & -A_1 \\ 0 & A_2 & 0 & \cdots & 0 \\ 0 & 0 & A_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & A_S \end{pmatrix},$$

то есть, это блочно-диагональная матрица. В этом случае решение задачи линейного программирования можно значительно ускорить, если применять методы разложения.

Замечание 3. При достаточно большой размерности n число вершин для аппроксимации (количество элементов множества V^k) должно быть весьма большим. Например, гиперкуб в R^{10} имеет 1024 вершины и требует хранения 10240 вещественных чисел в оперативной памяти.

Замечание 4. Основной операцией в АПП является решение задачи линейного программирования. От количества решенных задач линейного программирования в конечном итоге зависит точность алгоритма. Причем, задачи линейного программирования могут решаться независимо друг от друга, то есть алгоритм допускает распараллеливание.

Точность построения суммы алгоритмом последовательных приближений. Расстоянием по Хаусдорфу [3, с. 21] между двумя множествами $A, B \in \mathcal{R}^n$ называется величина

$$H(A, B) = \min r : r \geq 0, A \subset B + S_r(0), B \subset A + S_r(0),$$

где $S_r(0)$ – шар радиуса r с центром в 0 . Каждая точка многогранника H^k находится на расстоянии 0 от некоторой точки многогранника P (так как $H^k \subseteq P$), и каждая точка многогранника P находится не далее, чем на расстоянии $H(H^k, P)$ от некоторой точки многогранника H^k . Таким образом, $H(H^k, P)$ есть расстояние по Хаусдорфу между оценкой снизу H^k и точной суммой P .

Мерой величины множества A является максимальное расстояние между двумя его произвольными точками, называемое диаметром множества (обозначается $Diam(A)$) [17]. При решении практических задач необходимо задать точность построения суммы ε_1 для случая, когда $Diam(P) = 1$. Чем больше диаметр результата суммирования, тем больше возможно отклонение оценки суммы снизу от точной суммы. Поэтому точность построения суммы любого диаметра задается выражением $\varepsilon(P) = Diam(P)\varepsilon_1$. В случае, когда суммируемые множества известны с точностью $\varepsilon(P_1)$ и $\varepsilon(P_2)$, то следует задавать

$$\varepsilon(P) = \max \{Diam(P)\varepsilon_1, \varepsilon(P_1) + \varepsilon(P_2)\}.$$

В n -мерном пространстве диаметр множества из N точек всегда можно определить за время $O(nN^2)$, вычисляя расстояние между каждой парой точек. В качестве оценки диаметра многогранника P можно взять диаметр многогранника H^k , при этом $Diam(P) \leq Diam(H^k) + 2H(H^k, P)$. Диаметр множества H^k легко найти, поскольку известны все его вершины (количество вершин $n+k+1$ не велико, так как строится внутренняя аппроксимация суммы с минимальным числом вершин). Если известны диаметры многогранников P_1 и P_2 , то в качестве приближенной оценки $Diam(P)$ можно использовать сумму диаметров

$$Diam(P_1) + Diam(P_2) \geq Diam(P_1 + P_2) = Diam(P).$$

Пример 1. Для иллюстрации работы АПП достаточно рассмотреть суммирование двух многогранников P_1 и P_2 в \mathcal{R}^2 (рис. 1). Пусть P_1 задан системой

$$\begin{cases} x_2 \leq 1; \\ x_1 \leq 3; \\ -x_1 - x_2 \leq 1; \end{cases}$$

и P_2 системой

$$\begin{cases} 2x_1 + x_2 \leq 6; \\ -2x_1 + x_2 \leq 4; \\ x_1 - x_2 \leq 4; \\ -x_1 \leq 3. \end{cases}$$

Решая задачи линейного программирования на максимум и минимум вдоль координатных осей, на первом шаге алгоритма находим координаты вершин $v_1 = (-5, -1)$, $v_2 = (6.33, 0.33)$, $v_3 = (0, -11)$. Далее находим начальное приближение снизу к сумме – выпуклую оболочку H^0 векторов v_1, v_2, v_3 . При этом достигнутая точность будет $\rho_{\max} = 6.54$, а диаметр построенного множества $Diam(H^0) = 12.98$. Находим последовательно грани с максимальными величинами сдвига (рис. 2) и последовательно добавляем соответствующие вершины в описание оценки суммы (вершины пронумерованы в порядке их включения в описание суммы)

$v_4 = (-1.5, 6)$,	$\rho_{\max} = 2.93$,	$Diam(H^1) = 17.06$;
$v_5 = (3.5, 6)$,	$\rho_{\max} = 2.43$,	$Diam(H^2) = 17.36$;
$v_6 = (6.33, -4.67)$,	$\rho_{\max} = 2.23$,	$Diam(H^3) = 17.36$;
$v_7 = (-5, -6)$,	$\rho_{\max} = 0$,	$Diam(H^4) = 17.36$.

После добавления вершины v_7 и построения выпуклой оболочки задача суммирования решена точно, так как $\rho_{\max} = 0$. Для построения точной суммы пришлось решить 7 задач линейного программирования (одну для каждой вершины суммы).

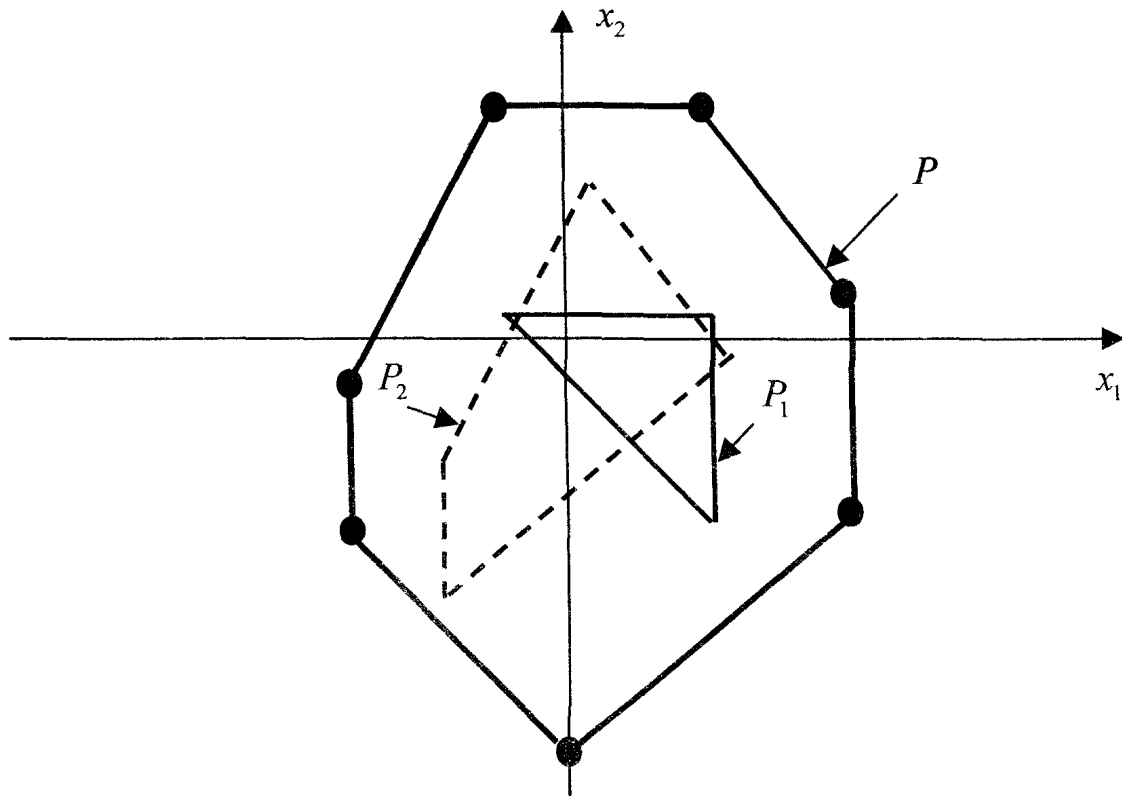


Рис. 1. Точная сумма многогранников в \mathbb{R}^2

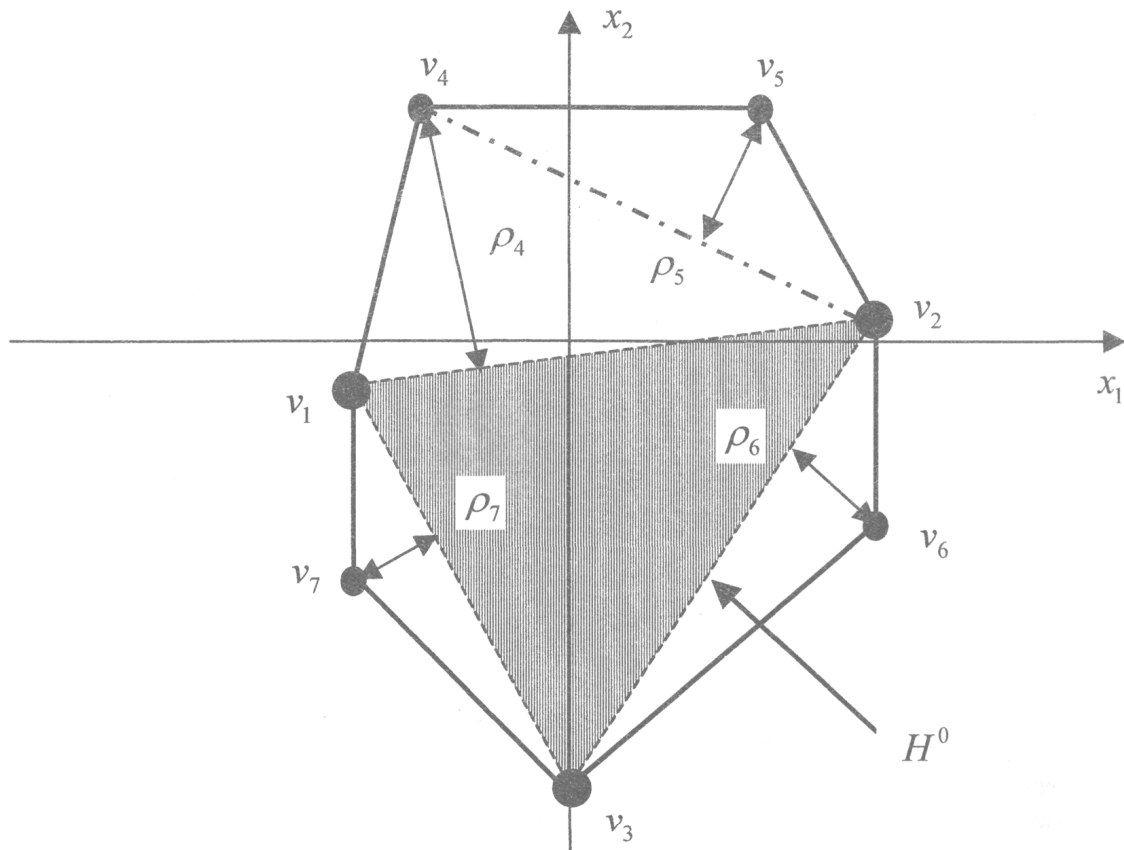


Рис. 2. Построение АПП точной суммы многогранников в \mathbb{R}^2

Найденная сумма описывается системой

$$P = H^4 = \begin{cases} -0.89x_1 & 0.45x_2 \leq 4.02; \\ & x_2 \leq 6.00; \\ 0.89x_1 & 0.45x_2 \leq 5.81; \\ 0.71x_1 & -0.71x_2 \leq 7.78; \\ x_1 & \leq 6.33; \\ -0.71x_1 & -0.71x_2 \leq 7.78; \\ -x_1 & \leq 5.00. \end{cases}$$

Нахождение оценки сверху (алгоритм сдвига граней). При решении задачи гарантированного оценивания необходимо получить оценку сверху информационного множества. Следовательно, алгоритм суммирования должен строить аппроксимацию сверху. В задаче же минимаксной фильтрации находится точечная оценка и не важно из какой множественной оценки (сверху или снизу) она получена. Поэтому оценка H^k с расстоянием по Хаусдорфу не больше $\varepsilon(P)$ достаточна для алгоритмов минимаксной фильтрации.

Пусть АПП построена H^k – нижняя оценка суммы. Тогда, \hat{P} – оценка суммы сверху, задается системой линейных неравенств

$$(c_i^k, x) \leq d_i^k + \|c_i^k\| \rho_i^k$$

или, используя определение для ρ_i^k , получаем описание многогранника \hat{P}

$$(c_i^k, x) \leq (c_i^k, x^*).$$

В результате такого сдвига граней многогранника H^k сдвигаются и вершины. Если сдвинутые грани образуют вершину n -го гиперкуба в \mathbb{R}^n , то точность построения верхней оценки будет удовлетворять неравенству

$$H(\hat{P}, P) \leq \sqrt{n}H(H^k, P).$$

В случае появления острых углов в вершинах верхняя оценка \hat{P} может сколь угодно сильно (в смысле хаусдорфова расстояния) отличаться от точной суммы.

Пример 2. Если задать точность $\varepsilon(P) = 0.25$ в примере 1, то две вершины суммы не будут найдены (рис. 3). Тогда расстояние по Хаусдорфу $H(\hat{P}, P) = 3.34$.

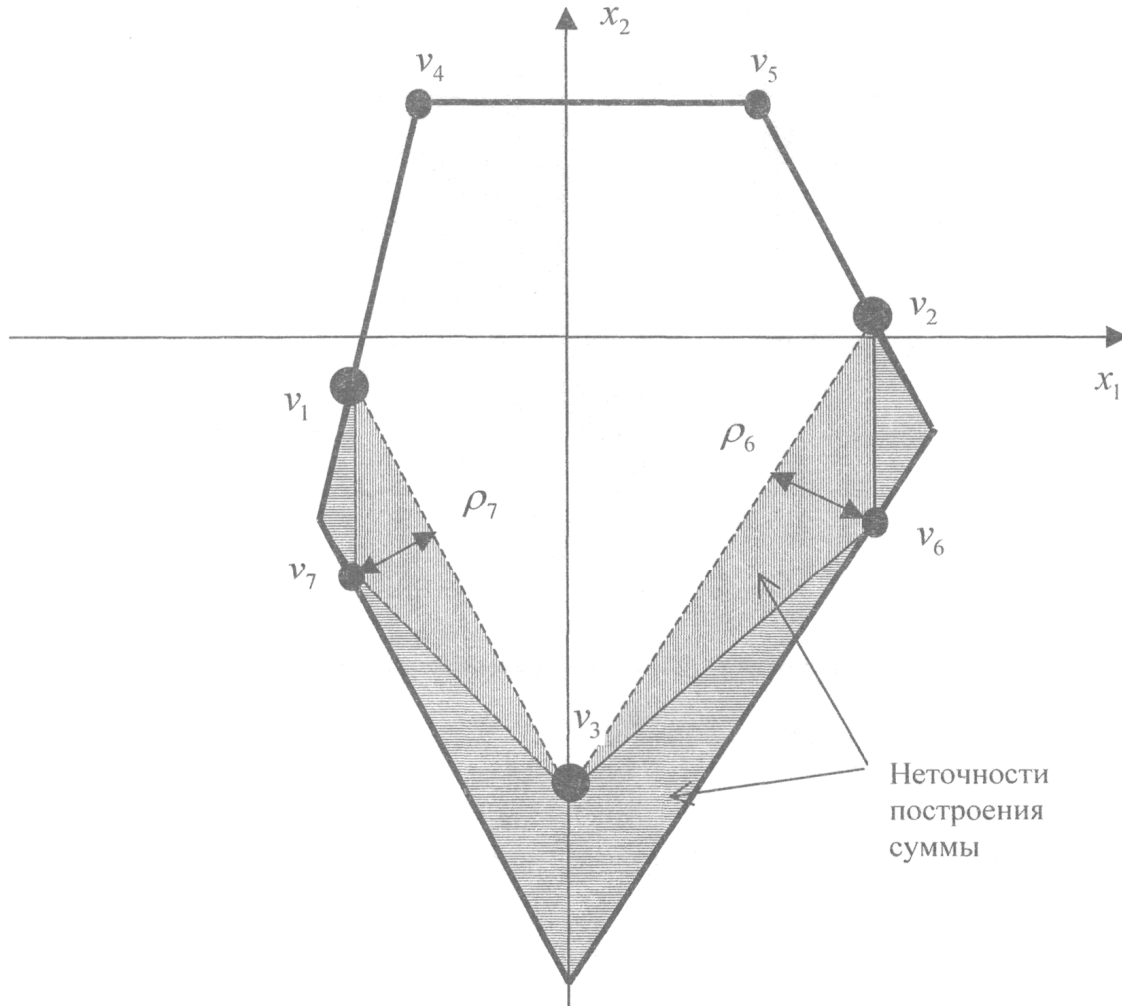


Рис. 3. Построение верхней оценки \hat{P} суммы по нижней H^k

Оценку \hat{P} суммы сверху можно получить, не выполняя трудоемкую операцию построения выпуклой оболочки, то есть не строя оценку H^k суммы снизу. Что приводит к алгоритму сдвига граней (АСГ) для нахождения суммы.

В АСГ используется тот факт, что сумма двух многогранников может быть аппроксимирована сверху пересечением исходных многогранников с передвинутыми гранями (рис. 4). Для построения суммы нужно найти сдвиги граней (например, симплекс-методом) обоих многогранников. Решаем задачи линейного программирования $c'_1 = \max(a'_1, x)$ при ограничениях $A_2x \leq b_2$, где a'_1 – i -ая строка матрицы A_1 ($i = \overline{1; m_1}$). Получаем c'_1 – i -ую компоненту вектора $c_1 \in \mathfrak{R}^{m_1}$. Далее решаем задачи линейного программирования $c'_2 = \max(a'_2, x)$ при ограничениях $A_1x \leq b_1$, где a'_2 – j -ая строка матрицы A_2 ($j = \overline{1; m_2}$). Получаем c'_2 – j -ую компоненту вектора $c_2 \in \mathfrak{R}^{m_2}$. Всего необходимо решить $m_1 + m_2$ задач линейного программирования. Тогда сумма будет описываться системой

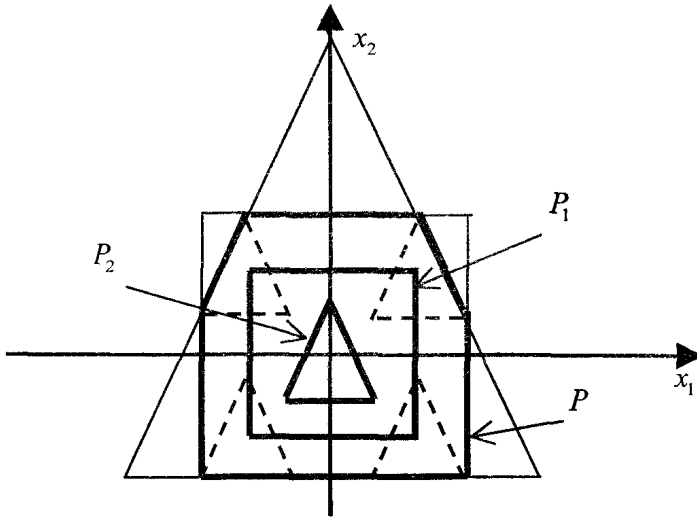


Рис. 4. Сумма квадрата и треугольника

$$\begin{cases} A_1x \leq b_1 + c_1; \\ A_2x \leq b_2 + c_2. \end{cases} \quad (10)$$

Исключаем из многогранника описываемого системой (10) параллельные грани.

АСГ использует только грани многогранников и не хранит информацию о вершинах суммы. Построенные алгоритмы (АПП и АСГ) обладают свойством коммутативности: при перестановке многогранников P_1 и P_2 результирующий многогранник не меняется. Однако ассоциативность АПП при $\varepsilon(P) > 0$ не имеет места, то есть при сложении более двух многогранников внутренняя аппроксимация суммы в общем случае зависит от

порядка выполнения сложения. АСГ обладает свойством ассоциативности.

АСГ строит аппроксимацию сверху суммы в \mathbb{R}^n (рис. 5). Рассмотрим случаи, когда АСГ строит точную сумму, то есть когда в сумме не возникают грани с новыми направлениями, кроме тех, которые есть в исходных многогранниках. Для этого докажем следующие теоремы.

Теорема 1. Алгоритм сдвига граней обеспечивает точное построение суммы многогранников в \mathbb{R}^2 .

Доказательство. Справедливость теоремы следует из того, что сумму многогранников в \mathbb{R}^2 можно получить, передвигая P_1 вдоль граней P_2 . Теорема не верна уже для \mathbb{R}^3 , так как для получения граней суммы нужно передвигать P_1 еще и вдоль ребер P_2 (ребро задается сочетанием граней P_2).

Теорема 2. Алгоритм сдвига граней строит точную сумму многогранников P_1 и P_2 , если многогранник P_1 представим в виде $P_1 = P_2 + Z$, где Z – произвольный многогранник. Причем, для построения суммы необходимо найти сдвиги граней многогранника P_1 относительно P_2 .

Доказательство. Покажем, что задача построения точной суммы $P_1 + P_2$ эквивалентна задаче сдвига граней многогранника P_1 относительно P_2 , то есть сумма $P_1 + P_2$ содержит грани с теми же направлениями что и многогранник P_1 . Для суммы по Минковскому справедливы свойства коммутативности и ассоциативности, поэтому получаем

$$P_1 + P_2 = (P_2 + Z) + P_2 = (Z + P_2) + P_2 = Z + (P_2 + P_2). \quad (11)$$

Сумма двух равных многогранников $P_2 = \{x_2 \in \mathbb{R}^n \mid A_2x_2 \leq b_2, A_2 \in \mathbb{R}^{m_2 \times n}, b_2 \in \mathbb{R}^{m_2}\}$ описывается системой $P = P_2 + P_2 = 2P_2 = \{x \in \mathbb{R}^n \mid A_2x \leq 2b_2, A_2 \in \mathbb{R}^{m_2 \times n}, b_2 \in \mathbb{R}^{m_2}\}$. Пусть m_1 – количество граней в P_1 , а m_2 – количество граней в P_2 . Тогда количество граней в многограннике $P_2 + P_2$ равно m_2 и их направления такие же, как и у многогранника P_2 . Суммы $Z + P_2$ и $Z + (P_2 + P_2)$ отличаются только столбцом свободных членов второго слагаемого. Операцию суммирования можно свести к операции исключения переменных из системы линейных неравенств. Формирование направлений неравенств свертки и их количество полностью зависит от матрицы системы и не зависит от столбца свободных членов. Следовательно, количество и направление граней суммы двух множеств, заданных системами линейных неравенств, не зависит от столбцов свободных членов (конечно, учитывая лишние неравенства). Лишние неравенства в P_1 не станут гранями $Z + (P_2 + P_2)$. Поэтому многогранники $Z + P_2$ и $Z + (P_2 + P_2)$ имеют по m_1 граней, направления которых совпадают. Используя свойство ассоциативности сложения по Минковскому (11), получаем, что количество граней в сумме $P_1 + P_2$ совпадает с количеством граней в много-

граннике $P_1 = P_2 + Z$, и направления граней в сумме $P_1 + P_2$ совпадают с направлениями граней в P_1 . Описание многогранника P_1 в виде системы линейных неравенств задано. Следовательно, при построении суммы $P_1 + P_2$ необходимо найти сдвиги граней многогранника P_1 относительно P_2 , а эту задачу АСГ решает точно. Теорема доказана.

Замечание 5. «Почти точно» АСГ будет работать для процесса

$$X_k = A_{k-1}X_{k-1} + W, \quad k = 1, \dots, N,$$

если $A_k \approx E$ (единичная матрица) и $X_1 = A_0X_0 + W$ найдено точно, например, алгоритмом последовательных приближений при $\varepsilon(P) = 0$.

Таким образом, результаты теоремы 2 можно применить к построению множества достижимости в гарантированном подходе, когда $A_k \approx E, k = 0, 1, \dots, N$.

Пример 3. Пусть суммируемые множества $P_1 \in \mathbb{R}^3$ – куб, $P_2 \in \mathbb{R}^3$ – пирамида

$$P_1 = \begin{cases} x_1 & \leq 1; \\ x_2 & \leq 1; \\ x_3 & \leq 1; \\ -x_1 & \leq 0; \\ -x_2 & \leq 0; \\ -x_3 & \leq 0; \end{cases} \quad \text{и} \quad P_2 = \begin{cases} x_1 + x_2 + x_3 & \leq 1; \\ -x_1 & \leq 0; \\ -x_2 & \leq 0; \\ -x_3 & \leq 0. \end{cases}$$

Оценка суммы сверху (рис. 5), найденная с помощью АСГ, описывается системой неравенств

$$\hat{P} = \begin{cases} -x_1 & \leq 0; \\ -x_2 & \leq 0; \\ -x_3 & \leq 0; \\ x_1 & \leq 2; \\ x_2 & \leq 2; \\ x_3 & \leq 2; \\ x_1 + x_2 + x_3 & \leq 4. \end{cases}$$

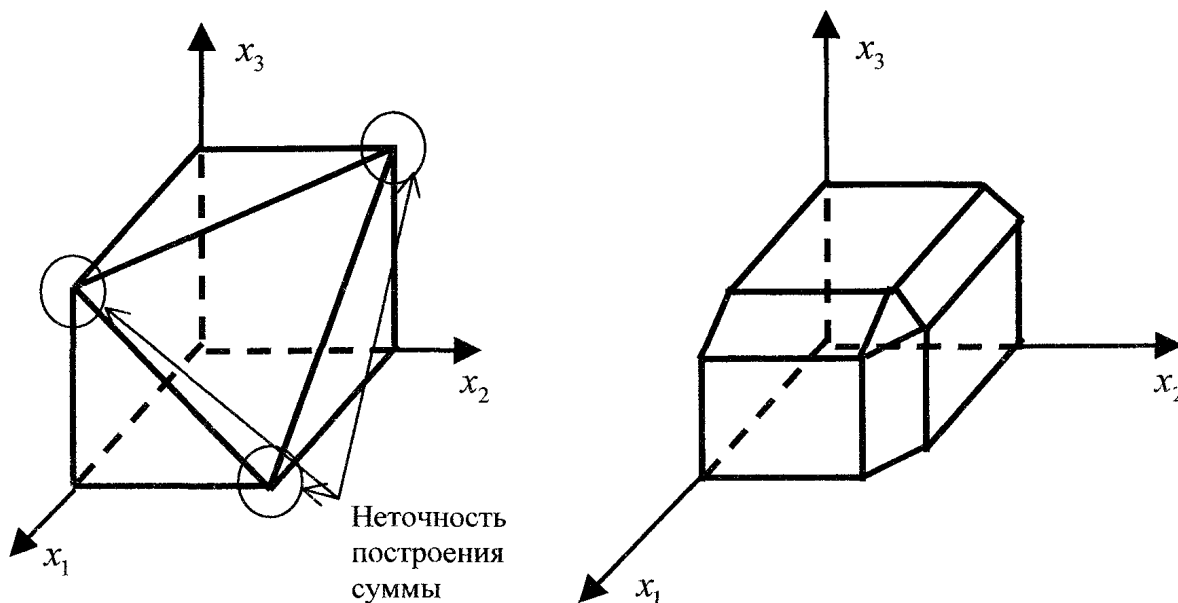


Рис. 5. Сумма куба и пирамиды: оценка сверху и точная сумма

Точное описание суммы задается системой

$$P = \begin{cases} x \in \hat{P}; \\ x_1 + x_2 \leq 3; \\ x_1 + x_3 \leq 3; \\ x_2 + x_3 \leq 3. \end{cases}$$

По сравнению с оценкой сверху в точной сумме содержится дополнительно еще три грани (рис. 5).

3. Другие теоретико-множественные операции над многогранниками

Пересечение двух многогранников. Рассматривается система (1)–(3), подверженная действию неопределенных возмущений. Производятся измерения вектора состояния, причем ошибка измерения также является неопределенным вектором, о котором известно лишь, что он принадлежит некоторому заданному множеству. В результате, после выполнения каждого очередного измерения оказывается, что вектор состояния системы принадлежит двум множествам: одно из них есть результат обработки всей прошлой информации, а другое соответствует последнему произведенному измерению. Следовательно, вектор состояния принадлежит пересечению указанных двух множеств. Поэтому ясно, что в процессе реализации методов гарантированного и минимаксного оценивания необходимо многократно выполнять операцию пересечения множеств или строить аппроксимации этого пересечения. Следовательно, операция пересечения множеств играет важную роль при реализации минимаксного фильтра.

Определение 4. Пересечением [3, с. 20] двух множеств P_1 и P_2 из n -мерного евклидова пространства \mathfrak{R}^n называется множество $P \in \mathfrak{R}^n$ такое, что

$$P = P_1 \cap P_2 \triangleq \{x \in \mathfrak{R}^n \mid x \in P_1, x \in P_2\}.$$

Как и при построении суммы будем рассматривать случай, когда пересекаемые множества заданы системами линейных неравенств (многогранниками)

$$P_1 = \{x_1 \in \mathfrak{R}^n \mid A_1 x_1 \leq b_1, A_1 \in \mathfrak{R}^{m_1 \times n}, b_1 \in \mathfrak{R}^{m_1}\},$$

$$P_2 = \{x_2 \in \mathfrak{R}^n \mid A_2 x_2 \leq b_2, A_2 \in \mathfrak{R}^{m_2 \times n}, b_2 \in \mathfrak{R}^{m_2}\}.$$

Приходим к следующей задаче: необходимо получить пересечение P многогранников P_1 и P_2 также в виде системы линейных неравенств

$$P = \{x \in \mathfrak{R}^n \mid Ax \leq b, A \in \mathfrak{R}^{m \times n}, b \in \mathfrak{R}^m\}.$$

При аппроксимации информационного множества гранями пересечение многогранников само по себе не требует вычислительных ресурсов и сводится к простому объединению систем линейных неравенств

$$P = P_1 \cap P_2 = \{x \in \mathfrak{R}^n \mid A_1 x \leq b_1, A_2 x \leq b_2\}.$$

Следовательно, операция пересечения выполняется точно, но возникает задача удаления зависимых неравенств, и эта задача объединяется с задачей аппроксимации. При задании множеств эллипсоидами операция пересечения двух эллипсоидов производится с большой потерей точности, поскольку пересечение эллипсоидов не является эллипсоидом.

Линейное преобразование. Линейным преобразованием $F \in \mathfrak{R}^{n \times m}$ множества $P_1 \in \mathfrak{R}^m$ называется множество $P \in \mathfrak{R}^n$ такое, что

$$P = FP_1 \triangleq \{x \in \mathfrak{R}^n \mid x = Fx_1, x_1 \in P_1\}.$$

Будем рассматривать случай, когда множество, подвергаемое линейному преобразованию, задано системой линейных неравенств (многогранником)

$$P_1 = \{x_1 \in \mathfrak{R}^m \mid A_1 x_1 \leq b_1, A_1 \in \mathfrak{R}^{m_1 \times m}, b_1 \in \mathfrak{R}^{m_1}\}.$$

Необходимо получить линейное преобразование P многогранника P_1 также в виде системы линейных неравенств

$$P = \{x \in \mathfrak{R}^n \mid Ax \leq b, A \in \mathfrak{R}^{m \times n}, b \in \mathfrak{R}^m\}.$$

В уравнениях (1)–(2) присутствуют вырожденное (в уравнении измерений, так как матрица измерений обычно вырожденная – не все координаты доступны непосредственному измерению)

и невырожденное преобразование (в уравнении движения объекта, так как матрицы перехода состояний являются обратимыми).

Легко получить аналитические формулы нахождения результата невырожденного линейного преобразования

$$P = FP_1 = \{x \in \mathfrak{R}^n \mid A_1 F^{-1} x \leq b_1\}.$$

Для вычисления результата невырожденного линейного преобразования необходимо выполнять операции над матрицами (умножение матриц и нахождение обратной матрицы). То есть невырожденное линейное преобразование многогранника вычисляется точно, быстро и дополнительная память не требуется.

Построение множества $X[y_k]$ состояний, совместных с измерением. Рассмотрим множество состояний, совместимых с измерением y_k

$$X[y_k] \triangleq \{x \in R^n \mid Gx + v = y_k, v \in V_k\}.$$

Будем рассматривать случай наиболее часто встречающийся на практике, когда данное преобразование представляет собой отображение из пространства \mathfrak{R}^m меньшей размерности в пространство \mathfrak{R}^n большей размерности.

Пусть множество возмущений V_k описывается системой линейных неравенств

$$v_k \in V_k = \{v \in \mathfrak{R}^m \mid U_2 v \leq u_2\}.$$

Тогда

$$X[y_k] = \{x \in R^n \mid -U_2 G_k x \leq u_2 - U_2 y_k\}. \quad (12)$$

Система линейных неравенств (12) задает в пространстве R^n многогранник, который в общем случае является неограниченным. Отсюда следует невозможность задания множества $X[y_k]$ с помощью вершин (так как есть еще лучи). Появляется вырожденный многогранник, но вырожденность уничтожается, так как далее в уравнениях гарантированной фильтрации следует операция пересечения с невырожденным многогранником (множеством прогнозов вектора состояния).

Полученная формула (12) справедлива и для случая, когда измеряем точно, то есть, когда $V_k = \emptyset$ и $y_k = G_k x_k$. В этом случае, учитывая что, как правило, матрица G_k – вырождена, x_k принадлежит линейному подпространству пространства \mathfrak{R}^n .

Пример 4. Рассмотрим пример неограниченности множества $X[y_k]$. Пусть измеряем сумму двух переменных пространства состояний

$$y_k = x_k^1 + x_k^2 + v_k, G_k = (1, 1), V_k = [-1; 1].$$

Здесь нижний индекс – номер шага, а верхний индекс – номер координаты. Тогда

$$X[y_k] = \{(x_k^1, x_k^2) \in R^2 \mid x_k^1 + x_k^2 \leq 1.5, -x_k^1 - x_k^2 \leq 0.5\}. \quad (13)$$

Множество (13) является неограниченным.

В процессе фильтрации происходит рост количества граней («закругление» многогранника) в описании информационного множества. Отсюда – увеличение времени вычислений, которое ограничено (необходимо к заданному времени получить оценку состояния для вычисления управления объектом). Следовательно, необходимо проводить аппроксимацию информационного множества многогранником с меньшим числом граней.

4. Нахождения точечной оценки

Для того чтобы управлять объектом, нужно знать точечную оценку, а не гарантированную, поэтому появляется задача нахождения точечной оценки по информационному множеству.

Если известна гарантированная оценка состояния – информационное множество, то в качестве точечной оценки выберем точку, принадлежащую этому информационному множеству. При решении задачи минимаксной фильтрации из всех точек выбирают одну – чебышевский центр информационного множества.

Определение 5. Чебышевским центром x^* многогранника P называется точка

$$x^* = \min_{x \in P} \max \{\|x - y\| \mid y \in P\}.$$

Чебышевский центр множества – это точка максимальное расстояние от которой до любой другой точки этого множества минимально.

Так как разрабатываются приближенные алгоритмы преобразования множеств, то отсутствует необходимость точного нахождения чебышевского центра информационного множества. Точное нахождение чебышевского центра весьма затруднительно.

Для построения чебышевского центра информационного множества $\bar{X}_k = \{x \in \mathfrak{R}^n \mid Ax \leq b, A \in \mathfrak{R}^{m \times n}, b \in \mathfrak{R}^m\}$ решаем задачи линейного программирования $c'_i = \max(a'_i, x)$ при ограничениях $Ax \leq b$, где $a'_i = (0, \dots, 1, \dots, 0) \in \mathfrak{R}^n$ – стоит единица на i -ом месте ($i = \overline{1; n}$). Получаем c'_i – i -ую компоненту вектора $c_1 \in \mathfrak{R}^n$. Далее решаем задачи линейного программирования $c'_{-j} = \min(a'_{-j}, x)$ при ограничениях $Ax \leq b$, где $a'_{-j} = (0, \dots, -1, \dots, 0) \in \mathfrak{R}^n$ – стоит минус единица на j -ом месте ($j = \overline{1; n}$). Получаем c'_{-j} – j -ую компоненту вектора $c_{-1} \in \mathfrak{R}^n$. Всего необходимо решить $2n$ задач линейного программирования. Таким образом, вписываем информационное множество \bar{X}_k в гиперпараллелепипед ориентированный по осям

$$\begin{cases} x \leq c_1; \\ x \geq c_{-1}. \end{cases} \quad (14)$$

За приближение к чебышевскому центру информационного множества \bar{X}_k принимаем чебышевский центр гиперпараллелепипеда (14)

$$\hat{x}^* = \frac{(c_1 + c_{-1})}{2}.$$

5. Результаты численного моделирования минимаксного фильтра на ЭВМ

Для реализации метода многогранников, изложенного в данной статье, был разработан комплекс программ на языке Borland C++. Комплекс программ предназначен для решения задач гарантированного и минимаксного оценивания состояния линейных дискретных динамических систем (1)–(2) с ограничениями (3), заданных системами линейных неравенств.

Так, аналогичный комплекс программ (но для метода эллипсоидов), разработанный В.Л. Трущенко на языке Fortran в институте проблем механики АН СССР позволяет решать задачи для системы (1)–(3), когда размерность вектора состояния $n \leq 10$ при числе измеряемых параметров $m \leq n$ в (2) [3, с. 279].

Близкая задача – задача построения множества достижимости для линейной динамической системы (1) решается пакетом прикладных программ ПОТЕНЦИАЛ, разработанным В.А. Бушенковым и А.В. Лотовым на языке АЛГОЛ 60 [18, 19]. На основе методов теории линейных неравенств [20, 21] и разработанного ими подхода удастся строить множества достижимости с размерностью вектора состояния n порядка 4 – 6 на несколько сот шагов по времени [3, с. 40–41].

Программа для построения эллипсоидальных аппроксимаций областей достижимости динамических систем, описываемых обыкновенными дифференциальными уравнениями, разработана В.А. Хониным в Институте математики и механики УрО АН СССР [22]. Размерность вектора состояния не должна превосходить 10 [22, с. 127].

Результаты измерений y_k должны либо поступать в комплекс программ непосредственно с измерительного устройства, либо моделироваться в ЭВМ. Во втором случае моделирование результатов измерений производится отдельной программой комплекса следующим образом. Задаются тем или иным образом внешнее возмущение w_k при $k \geq 0$, удовлетворяющее ограничению

$$w_k \in W_k = \{x \in \mathfrak{R}^n \mid U_1 x \leq u_1\},$$

ошибки измерения v_k при $k \geq 1$, удовлетворяющие ограничению

$$v_k \in V_k = \{x \in \mathfrak{R}^m \mid U_2 x \leq u_2\},$$

а также начальное условие x_0 , удовлетворяющее ограничению

$$x_0 \in X_0 = \{x \in \mathfrak{R}^n \mid A_0 x \leq b_0\}.$$

Вычисляется прогноз вектора состояния, в результате чего в любой момент времени определяется истинный вектор состояния x_k . На основе x_k и v_k согласно (2) вычисляется вектор результата измерения y_k , который используется в комплексе в процессе оценивания. Заметим, что вектор x_k не поступает в процедуру обработки и служит лишь для вычисления y_k ; при обработке он читается неизвестным.

В результате работы комплекса вычисляются грани многогранника \bar{X}_k , содержащего неизвестный текущий вектор состояния $x_k \in \bar{X}_k$. При этом для момента дискретного времени k учитываются результаты измерений $y_k(\cdot) = \{y_1, \dots, y_k\}$.

Рассмотрим операцию суммирования как самую трудоемкую операцию при реализации минимаксного фильтра. Грани суммируемых многогранников задавались датчиком случайных чисел так, чтобы многогранники стремились к шарам при стремлении числа граней к бесконечности. В алгоритмах используется двойственный симплекс-метод с нахождением начального псевдоплана M – методом. Для АПП была выбрана точность построения $\varepsilon_1 = 0.2$. Результаты приведены для Pentium-233 в табл. 1 и 2.

Таблица 1

Суммирование многогранников в \mathbb{R}^3 АПП

Число граней в P_1	10	10	20	40	40	100
Число граней в P_2	10	20	20	20	40	100
Время решения, с	2.20	14.50	7.52	33.23	17.36	87.93

Таблица 2

Суммирование многогранников АСГ ($m_1 = 100, m_2 = 100$)

Размерность пространства n	3	4	5	8	15	20
Время решения, с	1.54	2.01	2.63	5.33	16.98	30.16

Например, при суммировании двух многогранников в \mathbb{R}^3 с числом граней 100 и 100 АПП получено, что сумма содержит 628 неравенств (граней) и 64 вершины. Проведена 61 итерация АПП (каждая итерация прибавляет по одной вершине). В результате работы АПП получаем многогранник, который имеет в несколько раз больше граней, чем вершин.

Из табл. 1 и 2 можно сделать вывод о том, что АСГ работает быстрее АПП, т.к. не использует в своей работе построение выпуклой оболочки. Разница в скорости вычисления суммы увеличивается по мере увеличения размерности пространства.

Результаты численного моделирования в \mathbb{R}^n показали, что разработанный комплекс программ позволяет решать задачи для системы (1)–(3), когда размерность вектора состояния $n \leq 20$ при числе измеряемых параметров $m \leq n$.

Заключение

Способ задания информационных множеств системами линейных неравенств (то есть только гранями, не используя вершины – метод двойного описания [20]) позволил реализовать алгоритм минимаксного оценивания на ЭВМ в реальном времени для размерности вектора состояния не более 20. При этом разработаны эффективные алгоритмы построения внутренней и внешней аппроксимации суммы многогранников при решении задач минимаксной (или гарантированной) фильтрации. Для построения внутренней аппроксимации суммы операция суммирования была сведена к классической задаче теории линейных неравенств – задаче построения ортогональной проекции многогранника. Для построения внешней аппроксимации суммы разработан алгоритм сдвига граней, доказаны теоремы о случаях, когда алгоритм строит точную сумму.

Работа поддержана грантами РФФИ – УРАЛ № 01-01-96419 и МО № ТОО/13.2/2647.

Литература

1. Schweppe F.C. Recursive state estimation: unknown but bounded errors and systems inputs // *IEEE Trans. Autom. Control.* – 1968. – V. 13, № 1. – P. 22–28.
2. Куржанский А.Б. Управление и наблюдение в условиях неопределенности. – М.: Наука, 1977. – 392 с.
3. Черноушко Ф.Л. Оценивание фазового состояния динамических систем. Метод эллипсоидов. – М.: Наука, 1988. – 320 с.
4. Бакан Г.М. Нечеткий эллипсоидальный наблюдатель состояний многомерного статического объекта // *Проблемы управления и информатики.* – 1996. – № 1, 2. – С. 133–144.
5. Бакан Г.М. Гарантированное оценивание состояний динамических систем методом двустороннего динамического программирования // *Кибернетика и системный анализ.* – 2001. – № 1. – С. 83–96.
6. Кунцевич В.М., Лычак М.М. Синтез оптимальных и адаптивных систем управления: игровой подход. – Киев: Наукова думка, 1985. – 245 с.
7. Лычак М.М. Идентификация и оценивание состояния объектов управления на основе множественного подхода // *Проблемы управления и информатики.* – 1999. – № 5. – С. 34–41.
8. Костоусова Е.К. Внешнее и внутреннее оценивание областей достижимости при помощи параллелотопов // *Вычислительные технологии.* – 1998. – № 2. – С. 11–20.
9. Ширяев В. И. Минимаксная фильтрация в реальном времени многошаговых систем // *Проблемы управления и теории информации.* – 1991. – № 5. – С. 805–812.
10. Панченко И.С. Алгоритмы оценивания аддитивных скачкообразных возмущений в линейных динамических системах в условиях статистической неопределенности: Дис...канд. техн. наук / ЧГТУ. – Челябинск: Изд-во ЧГТУ, 1997. – 141 с.
11. Уханов М.В. Решение задачи гарантированного оценивания в условиях неопределенности // *Проблемы теоретической и прикладной математики. Труды 31-й Региональной молодежной конференции.* – Екатеринбург: УрО РАН, 2000. – С. 100–101.
12. Кац И.Я., Куржанский А.Б. Минимаксная многошаговая фильтрация в статистически неопределенных ситуациях // *Изв. РАН. Техн. киберн.* – 1978. – № 11. – С. 79–87.
13. Лейхтвейс К. Выпуклые множества. – М.: Наука, 1985. – 336 с.
14. Бренстед А. Введение в теорию выпуклых многогранников. – М.: Мир, 1988. – 240 с.
15. Черникова Н.Б. Алгоритм для нахождения общей формулы неотрицательных решений системы линейных неравенств // *Ж. вычисл. матем. и матем. физ.* – 1965. – Т. 5, № 2. – С. 334–337.
16. Бушенков В.А. Итерационный метод построения ортогональных проекций выпуклых многогранных множеств // *Ж. вычисл. матем. и матем. физ.* – 1985. – Т. 25, № 9. – С. 1285–1292.
17. Препарата Ф., Шеймос М. Вычислительная геометрия: Введение. – М.: Мир, 1989. – 478 с.
18. Компьютер и поиск компромисса. Метод достижимых целей / Лотов А.В, Бушенков В.А., Каменев Г.К., Черных О.Л. – М.: Наука, 1997. – 239 с. (Серия «Кибернетика: неограниченные возможности и возможные ограничения»).
19. Бушенков В.А., Лотов А.В. Методы и алгоритмы анализа линейных систем на основе построения обобщенных множеств достижимости // *Ж. вычисл. матем. и матем. физ.* – 1980. – Т. 20, № 5. – С. 1130–1141.
20. Моцкин Т.С., Райфа Х., Томсон Дж.Л., Тролл Р.М. Метод двойного описания // *Матричные игры.* – М.: Физматгиз, 1961. – С. 81–109.
21. Черников С.Н. Линейные неравенства. – М.: Наука, 1968. – 488 с.
22. Хонин В.А. О программах, реализующих алгоритмы аппроксимации области достижимости управляемой системы // *Динамические задачи оценивания в условиях неопределенности.* Свердловск: УрО АН СССР, 1989. – С. 125–129.

МЕТОД КОМИТЕТОВ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ОБРАЗОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ АПРИОРНОЙ ИНФОРМАЦИИ

Б.М. Кувшинов, О.В. Ширяев

Предложен подход к решению задачи автоматической классификации многопараметрических объектов методами распознавания образов, учитывающий неполноту знаний экспертов на этапе построения системы распознавания. Он позволяет совместить использование экспертных суждений и возможности математических методов анализа данных в рамках единого процесса обучения системы распознавания. В основу системы классификации положены комитетные решающие правила, для которых предложен эффективный алгоритм обучения, учитывающий неточность априорной информации.

Введение

Системы автоматической классификации, основанные на математических методах распознавания образов, получили широкое применение в самых разнообразных прикладных областях в качестве систем поддержки принятия решений [8, 9, 20]. Задачу выбора стратегии управления тем или иным объектом можно рассматривать как вопрос определения класса, к которому принадлежит данный объект. Априорной информацией для системы автоматической классификации являются результаты некоторых инструментальных измерений, проводимых над объектами исследования, а также суждения экспертов относительно классов некоторых из обследованных объектов. Такая конкретизация задачи управления позволяет применять для ее решения аппарат теории распознавания образов [5, 14, 17].

В традиционной постановке задачи распознавания четко разграничиваются случаи обучения и самообучения (дискриминантный и кластерный анализ соответственно) [2, 5, 17]. В первом случае имеется некоторый набор объектов (обучающая выборка), заданных значениями своих измеряемых характеристик. Для каждого объекта априорно, на основании суждений экспертов, задан его класс. Необходимо, опираясь на эти данные, построить решающее правило, т.е. правило отнесения произвольного объекта с известными значениями измеряемых параметров к тому или иному классу. Задача кластерного анализа отличается тем, что для объектов обучающей выборки значения их принадлежности к классам заранее не известны. При этом используется гипотеза о компактности, т.е. предположение о том, что объекты одного и того же класса имеют близкие, в некотором смысле, характеристики, а объекты разных классов существенно различаются с точки зрения значений своих параметров.

Большой интерес представляет разработка методов, позволяющих решать задачи классификации при недостатке информации об объектах исследования [3, 13, 23]. Во многих случаях эксперты не обладают достаточным количеством информации для принятия гарантированно безошибочных решений. Задачи такого рода постоянно возникают при внедрении систем распознавания в те прикладные области, в которых продолжают вестись научные исследования и окончательные представления об исследуемых объектах еще не сформировались. При этом сведения, получаемые от специалистов, уже нельзя рассматривать как некоторое абсолютное знание. Эта информация может содержать ошибки, быть неточной или неполной.

Достаточно распространенной является ситуация, когда эксперты в состоянии априорно задать классификацию для объектов обучающей выборки, однако допускают возможность наличия в ней большого количества ошибок. Решающее правило, построенное по данным такой выборки, будет обеспечивать весьма низкое качество классификации (т.к. ошибочные сведения, полученные от экспертов, будут использоваться в качестве материала для обучения). Таким образом, применение методов дискриминантного анализа оказывается неэффективным. С другой стороны, отказ от использования имеющихся неточных экспертных суждений, т.е. применение методов кластерного анализа, существенно снижает возможности системы классификации по анализу данных, т.к. происходит потеря существенной части априорной информации.

В подобных случаях разработчиками систем распознавания, как правило, делаются некоторые предположения относительно статистических законов распределения значений параметров, характеризующих состояние объектов [3, 6, 7]. При этом исходная обучающая выборка подвергается фильтрации: объекты, априорные сведения о классах которых не укладываются в сделанные вероятностные допущения, исключаются из рассмотрения, после чего используются методы дискриминантного анализа. Однако существует довольно широкий класс практических задач, в которых никаких предположений о статистических характеристиках распределения значений параметров объектов сделать невозможно. Это обуславливает необходимость разработки соответствующих непараметрических методов [4, 13].

1. Постановка задачи

Проблема недостатка и неточности исходных данных может быть решена за счет интеграции идей обучения и самообучения, используемых в системах распознавания образов. Если считать, что значения параметров объектов определяют точки в многомерном пространстве измеряемых характеристик объектов, то решающая функция задает некоторую поверхность в этом пространстве – границу между классами. Задача обучения – обобщить локальные свойства отдельных объектов обучающей выборки на целые классы, к которым они принадлежат. С геометрической точки зрения эта задача, в случае дискриминантного анализа, состоит в следующем: расположить решающую функцию в пространстве параметров объектов так, чтобы она:

- а) была как можно ближе к истинной границе между классами;
- б) правильно классифицировала все объекты обучающей выборки.

В рассматриваемом случае задачу обучения необходимо поставить следующим образом: имеется обучающая выборка, для объектов которой известны значения принадлежности к классам. В то же время некоторые из этих значений могут быть ошибочными. Необходимо построить решающее правило, которое:

- а) было бы как можно ближе к истинной границе между классами;
- б) классифицировало объекты обучающей выборки в соответствии с априорной классификацией в тех случаях, когда это не противоречит гипотезе о компактном расположении классов в пространстве характеристик объектов.

На рис. 1 показан случай решения задачи классификации для объектов с двумя измеряемыми характеристиками (x_1 и x_2). Закрашенные и незакрашенные точки соответствуют объектам двух классов (согласно априорной классификации). Решающее правило представляет собой некоторую кривую на плоскости параметров объектов. Апостериорные значения принадлежности объектов к классам определяются тем, с какой стороны от разделяющей кривой находится соответствующая объекту точка. Рис. 1а соответствует решению «чистой» задачи обучения – все объекты обучающей выборки классифицированы согласно априорным значениям их классов. Разделяющая функция на рис. 1б классифицирует один из объектов не так, как это было указано экспертом. В то же время ее положение гораздо ближе к действительной границе между классами. То есть способность системы классификации к обобщению информации во втором случае выше.

После того, как построено решающее правило, для всех объектов обучающей выборки можно вновь определить их классы. При этом априорная и апостериорная классификации будут частично различаться. Как показывает опыт, эти значения могут нести дополнительную информацию для специалиста из прикладной области исследований. Результаты автоматической классификации не могут быть абсолютно надежными, поэтому эксперт может соглашаться или не соглашаться с отдельными коррективами в классификации, сделанными системой распознавания. Таким образом, можно организовать итерационный процесс взаимодействия между экспертом и системой распознавания. В ходе этого процесса они обмениваются информацией, выраженной в форме классификаций объектов обучающей выборки. При этом решающее правило постепенно приближается к истинной границе между классами, и качество работы системы распознавания возрастает.

Следует отметить, что поставленная задача не сводится к вопросу подавления помех в обучающей выборке. Начальное количество правильно классифицированных объектов может быть очень мало и вообще не обеспечивать обучения системы классификации. Задача математических методов – обеспечить извлечение максимально возможного количества информации из предоставленных данных. Таким образом, задачи обучения и самообучения сводятся в единую задачу извлечения информации из обучающей выборки.

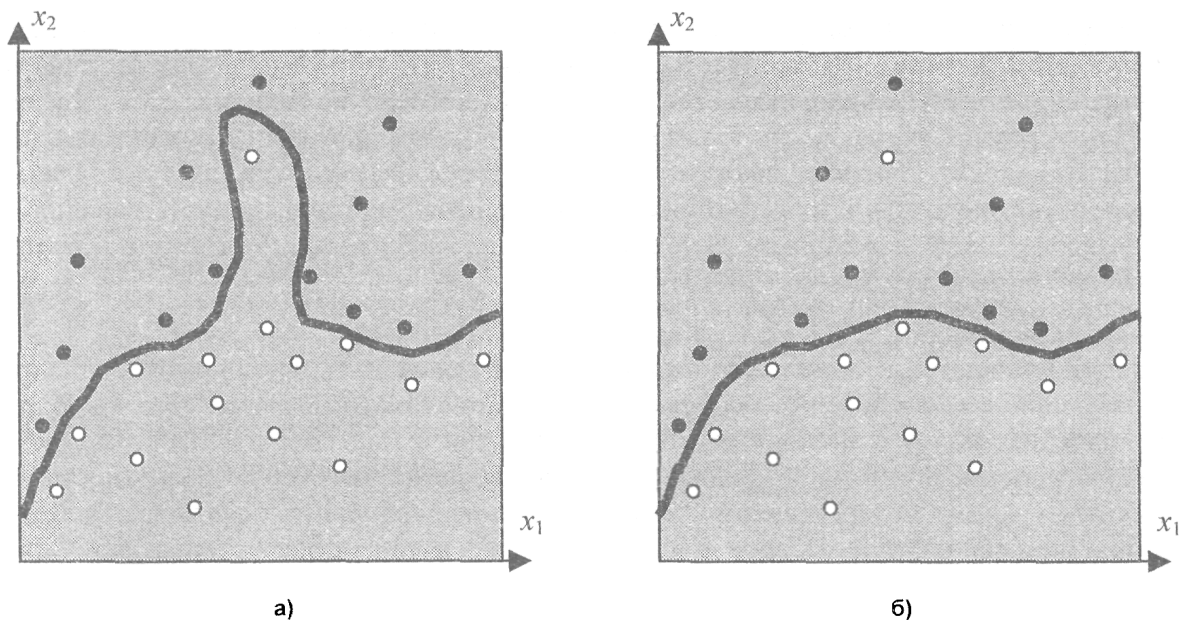


Рис. 1. Решающее правило при неточной априорной классификации

2. Выбор решающих правил

Для решения поставленной задачи необходимые средства предоставляет использование комитетных решающих правил (комитетов) [1, 14, 15, 18, 19, 21, 22]. Наибольшее распространение на сегодняшний день получили комитеты большинства на базе кусочно-линейных решающих правил [14, 15, 18, 19].

Определение [14]. Комитетом большинства системы линейных неравенств

$$\langle c; \tilde{x}^j \rangle < 0, \quad j = \overline{1, m}, \tag{2.1}$$

где $\tilde{x}^j \in \mathfrak{R}^n$, $c \in \mathfrak{R}^n$ называется такое множество $\{c^1, \dots, c^q\}$, $c^i \in \mathfrak{R}^n$, $i = \overline{1, q}$, что каждому неравенству системы удовлетворяет более половины элементов этого множества (членов комитета).

Комитеты такого вида обладают некоторыми свойствами, обуславливающими их эффективность при решении задач с рассмотренными выше особенностями:

- комитеты могут задавать сколь угодно сложную границу между классами;
- комитеты позволяют выбрать определенный класс функций (например, линейных) и на его основе наращивать сложность решающего правила до такой степени, чтобы она была адекватна сложности решаемой задачи;
- для элементов кусочно-линейного комитета – гиперплоскостей – существуют эвристические процедуры оптимизационного поиска, временная сложность которых линейно зависит от размерности задачи.

Дадим геометрическую интерпретацию комитета в задаче классификации. Будем считать, что система (2.1) содержит $(n + 1)$ неизвестных. Разложим каждый член комитета на составляющие

$$c^i = (a', b') = (a'_1, \dots, a'_n, b') \in \mathfrak{R}^{n+1}, \quad i = \overline{1, q},$$

и выберем коэффициенты неравенств специального вида:

$$\begin{cases} \tilde{x}^j = (x^j, -1), \quad \forall x^j \in A, \\ \tilde{x}^j = (-x^j, 1), \quad \forall x^j \in B, \end{cases} \quad j = \overline{1, m},$$

где $A = \{x^1\}$, $B = \{x^m\}$, $A \cup B = \{x^j, j = \overline{1, m}\}$, $A \cap B = \emptyset$ – обучающая выборка, состоящая из объектов двух классов (A и B). Вектора a' представляют собой нормальные вектора гиперплоскостей и вместе с числами b' задают q гиперплоскостей в пространстве параметров объектов. Вектора x^j определяют соответствующие объектам точки в том же пространстве. Разделяющий комитет $\{c^1, \dots, c^q\}$ обладает тем свойством, что для $\forall x \in A \cup B$ справедливы неравенства

$$\begin{cases} \sum_{i=1}^q \text{sign}[\langle a^i; x \rangle - b^i] \geq 0 \quad \forall x \in A, \\ \sum_{i=1}^q \text{sign}[\langle a^i; x \rangle - b^i] < 0 \quad \forall x \in B, \end{cases} \quad (2.2)$$

где $\text{sign}(y) = 1$, если $y \geq 0$, $\text{sign}(y) = -1$, если $y < 0$.

То есть комитет обладает свойствами решающей функции. Граница между образами, соответствующими классам A и B в пространстве параметров объектов – это поверхность, при переходе через которую значение функции

$$\sum_{i=1}^q \text{sign}[\langle a^i; x \rangle - b^i], \quad x \in \mathbb{R}^n$$

переходит через 0. Эту поверхность формируют пересекающиеся части гиперплоскостей, соответствующих членам комитета. Таким образом, граница между классами задается многогранной (кусочно-линейной) поверхностью. На рис. 2а показана геометрическая интерпретация комитета для случая двумерного пространства признаков. Отдельные гиперплоскости, соответствующие членам комитета, представляют собой в этом случае прямые. Штрихи на концах прямых обозначают направление векторов их нормалей, т.е., практически, указывают те полуплоскости (полупространства для многомерного случая), в которых выражения $\langle a^i; x \rangle - b^i$ имеют отрицательный знак. Жирной линией показана граница между образами, задаваемая этим комитетом. Представленный набор гиперплоскостей действительно является комитетом, поскольку он удовлетворяет условию (2.2). Геометрически это означает, что точки, соответствующие разным классам находятся по разные стороны от границы между образами, задаваемой комитетом. Далее для иллюстрации свойств разделяющих комитетов в двумерном случае используется та же система обозначений.

Важным частным случаем разделяющего комитета для заданной обучающей выборки является комитет, минимальный по количеству членов. Он не только задает решающую функцию, корректную с точки зрения обучающей выборки, но и обеспечивает максимальное обобщение информации, т.е. собственно обучение системы распознавания. Чем меньше количество членов комитета, тем ближе они должны быть расположены к своим оптимальным положениям, чтобы обеспечить сохранение комитетом свойства (2.2) решающей функции. Важность этой задачи иллюстрирует рис. 2. Выбранные в качестве комитетов совокупности гиперплоскостей в обоих случаях обеспечивают корректное разделение на классы объектов обучающей выборки. В то же время комитет на рис. 2б описывает границу между образами гораздо более точно, чем комитет на рис. 2а.

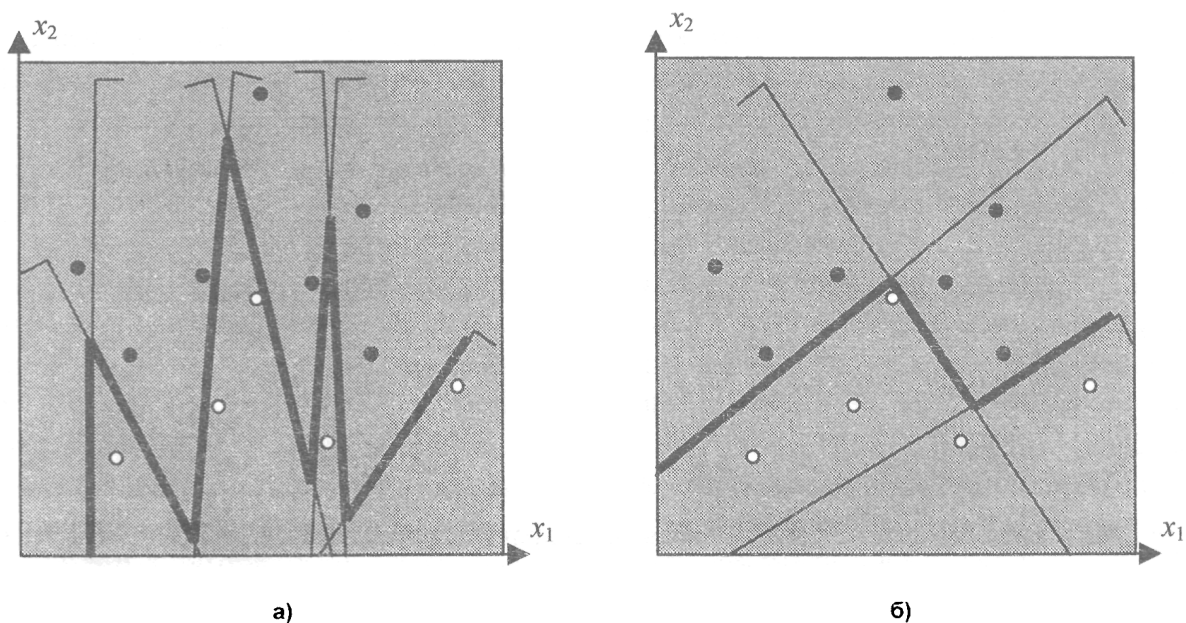


Рис. 2. Способность разных комитетов к обобщению информации

Для решения поставленной задачи извлечения информации из обучающей выборки необходимо модифицировать понятие комитета. Как было указано выше, решающее правило должно классифицировать объекты так же, как в априорной классификации, если это не противоречит положению других объектов обучающей выборки. Решение комитета большинства можно рассматривать как результат голосования отдельных членов комитета (гиперплоскостей) «за» или «против» принадлежности объекта тому или иному классу. То есть, если задан комитет, то для любого объекта может быть определен не только его класс, но и ранг (степень уверенности комитета в принятом решении):

$$r = \begin{cases} -\sum_{i=1}^q \text{sign} [\langle a'; x \rangle - b'] , & \text{если } x \in A, \\ \sum_{i=1}^q \text{sign} [\langle a'; x \rangle - b'] , & \text{если } x \in B. \end{cases} \quad (2.3)$$

Тогда задачу извлечения информации из обучающей выборки можно рассматривать как задачу построения комитета, обеспечивающего максимизацию некоторой функции от рангов объектов. В качестве такой функции можно выбрать, например, суммарный ранг всех объектов обучающей выборки. Необходимость предлагаемой модификации обусловлена двумя причинами:

1) реально существующие задачи распознавания и классификации, в силу указанных выше особенностей, как правило, не позволяют четко разграничить случаи обучения и самообучения системы распознавания;

2) для комитетных решающих правил стандартного вида, несмотря на их высокую эффективность в задачах большой размерности, пока не найдено эффективных алгоритмов обучения, сочетающих достоинства малой вычислительной сложности и высокой степени обобщения информации (т.е. приближения к минимальному комитету).

3. Алгоритм построения комитета

Для построения комитета, обладающего указанными свойствами, необходимо организовать итерационный процесс. На каждом шаге этого процесса положение гиперплоскостей, составляющих комитет, должно изменяться так, чтобы обеспечить возрастание, или хотя бы неубывание, значения целевой функции

$$R = \sum_{i=1}^m r_i ,$$

где r_i – ранг i -го объекта обучающей выборки, вычисленный по формуле (2.3), m – размер обучающей выборки.

Пусть на некотором шаге получен комитет $\{c^1, \dots, c^q\}$, $c^i \in \mathcal{R}^{n+1}$, $i = \overline{1, q}$. Вычислим ранги всех точек обучающей выборки по формуле (2.3). Для какой-нибудь из q гиперплоскостей, составляющих комитет, найдем объект $x^v \in A \cup B$, правильно классифицированный ею (с точки зрения априорной классификации), и объект $x^w \in A \cup B$, классифицированный данной гиперплоскостью неправильно. При этом необходимо, чтобы выполнялось условие:

$$r_v \leq r_w . \quad (3.1)$$

Стоит подчеркнуть, что речь идет о правильности классификации объектов именно отдельной гиперплоскостью, а не комитетом в целом. Пусть $I \subset \{1, \dots, m\}$ – множество номеров объектов обучающей выборки, правильно классифицированных данной гиперплоскостью. Согласно введенным обозначениям $v \in I$, $w \notin I$. Попытаемся изменить положение гиперплоскости так, чтобы обеспечить $w \in I$, а требование $v \in I$ снимем. Если это удалось, то новое множество номеров правильно классифицированных объектов будет иметь вид $I' \supset \{w\} \cup I \setminus \{v\}$ или $I' \supset \{w\} \cup I$, в зависимости от того, удалось ли сохранить принадлежность объекта x^v к числу правильно классифицированных. При этом новые значения рангов объектов:

$$\begin{aligned}
 r'_i &= r_i + 1, \quad i = w, \\
 r'_i &\geq r_i - 1, \quad i = v, \\
 r'_i &= r_i, \quad i \in I \setminus \{v\}, \\
 r'_i &\geq r_i, \quad i \in \{1, \dots, m\} \setminus [I \cup \{w\}],
 \end{aligned}
 \tag{3.2}$$

т.е. новое значение суммарного ранга всех точек по крайней мере не меньше его старого значения. Если же при изменении положения гиперплоскости удастся сохранить правильную классификацию ею объекта x^v или добавить к числу правильно классифицированных какие-либо из объектов $x^i, i \notin I \cup \{w\}$, то значение целевой функции увеличивается. На рис. 3 показан пример изменения положения комитета в результате рассмотренной процедуры. Числа около точек показывают ранги соответствующих объектов обучающей выборки до и после перемещения гиперплоскости.

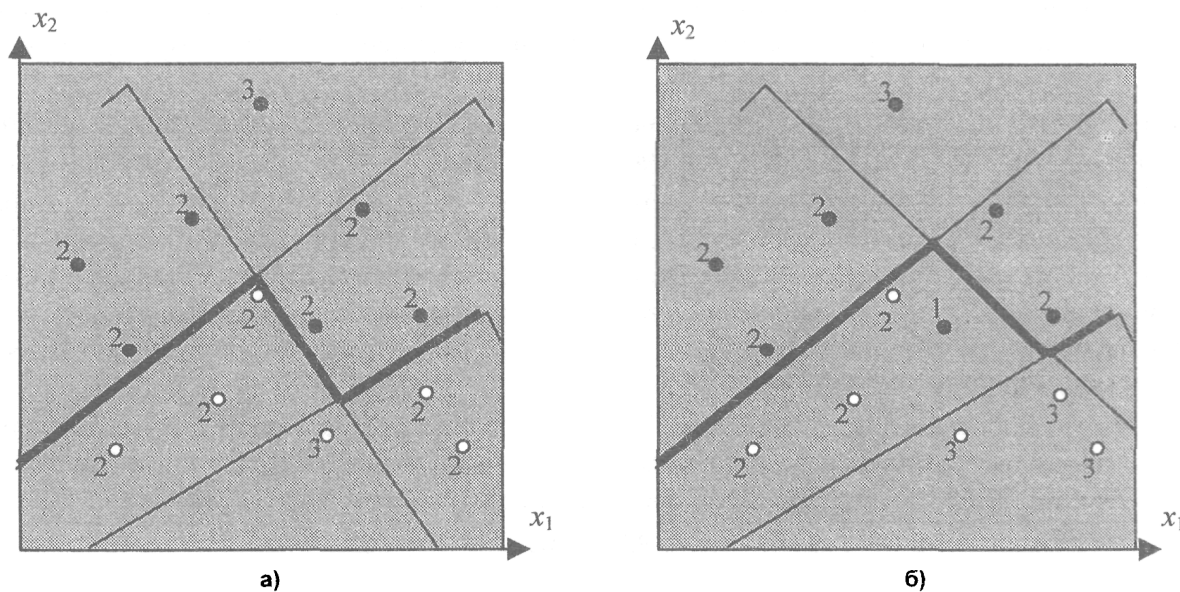


Рис. 3. Перемещение гиперплоскости-члена комитета

Таким образом, задавшись некоторым начальным положением комитета, можно организовать итерационный процесс коррекции положения гиперплоскостей, ведущий к увеличению значения целевой функции. Рассмотрим функцию следующего вида:

$$\tilde{R} = \sum_{i=1}^m \left[\left(\sum_{j=1}^q s_{ij} \right)^2 \right],$$

где $s_{ij} = 1$, если j -я гиперплоскость правильно классифицирует i -й объект и $s_{ij} = 0$ – в противном случае. Согласно выражениям (3.2), с учетом требования (3.1), значение функции \tilde{R} строго возрастает на каждом шаге итерационного процесса. Поскольку существует точная верхняя граница значения этой функции:

$$\tilde{R} = \sum_{i=1}^m \left[\left(\sum_{j=1}^q s_{ij} \right)^2 \right], \quad s_{ij} \in \{0, 1\} \Rightarrow \tilde{R} \leq mq^2,$$

процесс завершится за конечное число шагов.

Если пересчитывать ранги объектов после перемещения каждой гиперплоскости, то изменение положения комитета будет в значительной степени зависеть от порядка выбора гиперплоскостей, подлежащих перемещению. Поэтому целесообразно на каждой итерации перемещать все плоскости, а уже потом рассчитывать новые значения рангов объектов. Исходя из вышесказанного, алгоритм построения комитета, извлекающего максимум информации из обучающей выборки, должен выглядеть следующим образом.

1. Задать начальное положение комитета случайным образом.
2. Рассчитать ранги всех объектов обучающей выборки по формуле (2.3).
3. Выбрать для обработки очередную гиперплоскость.
4. Попытаться найти еще не рассмотренную пару правильно и неправильно классифицированных ею объектов, удовлетворяющих условию (3.1).
5. Если объекты с требуемыми свойствами не найдены, то перейти к шагу 7, иначе попытаться изменить положение гиперплоскости в соответствии с вышеописанными правилами.
6. Если положение гиперплоскости удалось изменить, то перейти к шагу 7, иначе перейти к шагу 4.
7. Если все гиперплоскости уже рассмотрены, то перейти к шагу 8, иначе перейти к шагу 3.
8. Если хотя бы одна из гиперплоскостей изменила свое положение, то перейти к шагу 2, иначе искомым комитет найден.

В приведенном алгоритме имеется две операции, вычислительная сложность которых определяет аналогичный показатель для всего алгоритма. Это поиск объектов x^v и x^w – кандидатов на изменение рангов, и проверка возможности перемещения гиперплоскости, такого, что будут выполняться условия (3.2).

Для поиска пар объектов x^v и x^w необходимо организовать перебор всех возможных сочетаний (v, w) , $v \in I$, $w \in \{1, \dots, m\} \setminus I$, удовлетворяющих условию (3.1). С целью сокращения перебора объекты правильно и неправильно классифицированные данной гиперплоскостью необходимо представить в виде двух упорядоченных списков: первые – в порядке возрастания рангов, вторые – в порядке убывания. Сначала из первого списка выбирается объект-кандидат на исключение из множества I . Затем из второго списка последовательно выбираются объекты, номер которых мог бы быть присоединен к множеству I . При этом просмотр второго списка выполняется лишь до тех пор, пока выполнено условие (3.1).

После того, как были выбраны объекты x^v и x^w , необходимо найти такое положение гиперплоскости, которое обеспечит выполнение условий (3.2), либо определить, что у этой задачи решения нет. Это означает, что необходимо проверить линейную разделимость множеств:

$$\begin{aligned} \tilde{A} &= \{x^j : x^j \in A, j \in \{w\} \cup I \setminus \{v\}\} \quad \text{и} \\ \tilde{B} &= \{x^j : x^j \in B, j \in \{w\} \cup I \setminus \{v\}\}. \end{aligned}$$

То есть нужно решить систему линейных неравенств:

$$\begin{cases} \langle x^j; a^j \rangle - b^j < 0, & x^j \in \tilde{A}, \\ \langle x^j; a^j \rangle - b^j > 0, & x^j \in \tilde{B}, \end{cases} \quad (3.3)$$

относительно $c^j = (a_1^j, \dots, a_n^j, b^j) \in \mathfrak{R}^{n+1}$ либо определить, что решения нет. С геометрической точки зрения линейная разделимость множеств означает, что существует гиперплоскость c^j такая, что точки, соответствующие объектам из разных множеств, находятся по разные стороны от нее.

Переобозначим

$$\begin{cases} d^i = (x^i, -1), & i : x^i \in \tilde{A}, \\ d^i = (-x^i, 1), & i : x^i \in \tilde{B}. \end{cases}$$

Тогда систему неравенств (3.3) можно представить в виде:

$$\langle d^i; c^j \rangle < 0, \quad i = 1, |\tilde{A} \cup \tilde{B}|. \quad (3.4)$$

Каждый вектор d^j , $j = 1, |\tilde{A} \cup \tilde{B}|$ будем рассматривать как нормальный вектор некоторой гиперплоскости в пространстве размерности $(n + 1)$, проходящей через начало координат. Тогда решением системы неравенств (3.4) является такая точка $(c_1^j, \dots, c_{n+1}^j) \in \mathfrak{R}^{n+1}$, которая лежит с нужной стороны от всех гиперплоскостей (т.е. задается $(n + 1)$ -мерным радиус-вектором, составляющим острый угол с нормальным вектором d^j , $j = 1, |\tilde{A} \cup \tilde{B}|$ каждой гиперплоскости). Таким образом, каждому объекту исходного n -мерного пространства соответствует одна из гиперплоскостей, образующих в совокупности $(n + 1)$ -мерный конус, а искомая разделяющая гиперплоскость c^j – любая точка внутри этого конуса.

Для поиска решения воспользуемся модификацией известного метода линейной коррекции [16]. Выберем некоторую исходную точку $c^j(0) \subset R^{n+1}$, и будем перемещаться по пространству посредством итерационного процесса:

$$c^j(t+1) = c^j(t) - \frac{\lambda}{\sigma} \sum_{i=1}^m f_i \cdot d^i, \quad t = 0, 1, \dots, T, \quad (3.5)$$

где λ – коэффициент ускорения движения; $\sigma = \sum_{i=1}^m \langle d^i; d^i \rangle$ – нормирующий коэффициент, обеспечивающий масштабирование итерационного шага; $f_i = \langle c^j(t); d^i \rangle$ – расстояние от текущей точки пространства до i -й гиперплоскости; T – максимально допустимое количество итераций, по достижении которого задача считается неразрешимой.

Фактически, формула (3.5) описывает частный случай метода градиентного спуска, где целевая функция – это сумма расстояний от точки пространства до всех гиперплоскостей:

$F = \sum_{i=1}^m f_i$. В то же время, в задаче проверки линейной разделимости решение трактуется как раз-

деляющая гиперплоскость. Отсюда следует, что вектора x^j , $j = 1, |\bar{A} \cup \bar{B}|$ задают большое число избыточных ограничений, т.е. таких, которые не уменьшают объем n -мерной области, в пределах которой может находиться разделяющая гиперплоскость. Соответственно, вектора d^j , $j = 1, |\bar{A} \cup \bar{B}|$ задают большое число гиперплоскостей в $(n+1)$ -мерном пространстве, которые не уменьшают объем конуса решений системы (3.4). Поэтому применение выражения (3.5) может не дать решения – в случае, если центр (ось) масс совокупности гиперплоскостей d^j : $x \subset R^{n+1} : \min_{|x|=const} F$, лежит вне конуса.

Во избежание рассмотренного эффекта модифицируем целевую функцию следующим образом:

$$\tilde{F} = \sum_{i=1}^m [\text{sign}'(f_i) f_i],$$

где $\text{sign}'(y) = \begin{cases} 1, & y \geq 0, \\ 0, & y < 0. \end{cases}$ – пороговая функция. Тогда соотношение (3.5) принимает вид:

$$c^j(t+1) = c^j(t) - \frac{\lambda}{\sigma} \sum_{i=1}^m [\text{sign}'(f_i) f_i \cdot d^i], \quad t = 0, 1, \dots, T.$$

То есть при вычислении направления перемещения на очередном итерационном шаге будут учитываться только те ограничения, которые еще не выполнены. Использование переменного параметра λ позволяет снизить количество итераций. Будем модифицировать его по следующему правилу:

– если при совершении очередного итерационного шага не начало выполняться ни одно ранее не выполнявшееся ограничение и не перестало выполняться ни одно ранее выполнявшееся ограничение, то $\lambda_{t+1} = 2\lambda_t$;

– в противном случае $\lambda_{t+1} = \frac{1}{2}\lambda_t$.

Такой выбор изменения скорости перемещения, с одной стороны, предотвращает «зависание» алгоритма в области, достаточно удаленной от решения, а, с другой стороны, подавляет колебания процесса вокруг решения в случае, когда размер последнего шага оказался шире сечения конуса в направлении $c^j(t) - c^j(t-1)$.

Рассмотренный алгоритм проверки линейной разделимости является эвристическим и, в принципе, не гарантирует нахождения решения системы (3.3). В то же время практическое его применение полностью оправдано в силу двух причин.

1. Временная сложность этого алгоритма линейно зависит от размерности задачи, т.е. он пригоден для использования в реальных задачах распознавания большой размерности.

2. Итерационный процесс поиска решения начинается от некоторого начального приближения. В качестве такого приближения можно выбрать текущее положение гиперплоскости. При очередном перемещении гиперплоскости в число правильно классифицированных ею входит достаточно небольшое количество новых объектов (в предельном случае только один объект – x^w). Поэтому большинство ограничений, налагаемых системой (3.3), выполняется уже для начального приближения. То есть обнаружение решения, если оно есть, следует ожидать уже после небольшого количества итераций. Таким образом, максимальное количество итераций алгоритма линейной коррекции можно ограничить очень небольшим числом. Если решение после их выполнения не найдено, то можно прекратить поиск.

Для завершения описания алгоритма поиска решающего правила необходимо сделать еще одно замечание. Комитетный характер решающего правила приводит к тому, что в ходе работы алгоритма его построения некоторые гиперплоскости начинают объединяться. Т.е. в результате итерационного перемещения комитета появляются такие его члены, которые одинаково классифицируют все объекты обучающей выборки. Если множества \tilde{A} и \tilde{B} для гиперплоскостей $c', c^i, i, j \in \overline{1, q}$ на некоторой итерации совпали, то, согласно алгоритму, следует ожидать, что и на последующих итерациях эти множества будут изменяться одинаково. Таким образом, дальнейшие расчеты можно проводить только для одной из объединившихся гиперплоскостей, а остальные перемещать так же, как и данную. Кроме того, можно считать, что объединение гиперплоскостей – это результат изменения сложности комитета под действием структуры предоставленных данных. Это означает, что объединившиеся гиперплоскости можно рассматривать как одну, т.е. в ходе итерационного процесса сложность комитета (число его членов) может уменьшаться.

4. Пример использования алгоритма

Для проверки эффективности работы рассмотренного алгоритма использовались результаты обследования больных гипертрофической кардиомиопатией [10–12]. Ставилась задача оценивания тяжести заболевания. В качестве обучающей выборки рассматривалась группа из 75 больных. Каждый больной описывался вектором числовых и логических параметров, которые являются результатами инструментальных измерений и простых экспертных суждений, не требующих высокой квалификации медицинского персонала. Размерность вектора параметров составила 120. Больные были разделены по степени тяжести заболевания на три группы: легкие, средние и тяжелые. Относительно некоторых больных (30 % от общего числа) экспертами-медиками был достоверно определен класс тяжести. Для остальных больных класс тяжести также задавался априорно, однако эксперты допускали возможность его ошибочного определения.

В результате автоматической классификации система распознавания изменила значения классов тяжести более чем у 40 % больных. В половине из этих случаев эксперты согласились с выводами системы распознавания. После второго обмена классификациями между экспертами и системой распознавания удалось довести степень их согласованности до 95 %. Для проверки обобщающей способности полученных решающих правил в рассмотрение было введено еще 25 больных с известными классами тяжести (тестовая выборка). В 90 % случаев система распознавания правильно определила тяжесть кардиомиопатии. Количество итераций алгоритма построения комитета не превышало 20.

Заключение

1. Предложенная методика построения системы распознавания опирается на два основных положения:

– объединение идей обучения и самообучения и построение единой процедуры автоматической классификации, обеспечивающей извлечение максимума информации из предоставленных данных;

– включение экспертов в состав единой автоматизированной системы принятия решений и использование их способностей в качестве элемента системы классификации.

2. В качестве основного элемента системы классификации использованы комитетные решающие правила, для которых предложен эффективный алгоритм обучения, использующий особенности постановки задачи.

3. Разработанные алгоритмы применены для решения практической задачи оценивания тяжести кардиологических заболеваний. Полученные показатели качества работы системы распо-

знавания: 95 % – для решения единой задачи извлечения информации из обучающей выборки и 90 % – при классификации объектов тестовой выборки, говорят о высокой эффективности применения метода к решению разнообразных прикладных задач.

Работа поддержана грантами РФФИ – УРАЛ № 01-01-96419 и МО № ТОО/13.2/2647.

Литература

1. Алескеров Ф.Т., Дуган Д. Процедуры голосования в пространстве функций выбора: Немонотонный случай // *Автоматика и телемеханика*. – 1994. – № 6. – С. 115–134.
2. Бархушин А.П. Обучение и самообучение систем распознавания образов // *Сб. науч. тр. НИИ информационных технологий*. – 1999. – № 7. – С. 14–20.
3. Буркин В.С. Синтез оптимальных алгоритмов портретного распознавания при различном уровне априорной осведомленности // *Известия АН. Теория и системы управления*. – 1997. – № 6. – С. 152–163.
4. Викентьев А.А., Лбов Г.С. О метризациях булевой алгебры предложений и информативности высказываний экспертов // *Доклады РАН*. – 1998. – Т. 361, № 2. – С. 174–176.
5. Горелик А.Л., Скрипкин В.А. Методы распознавания. – М.: Высшая школа, 1977. – 222 с.
6. Жук Е.Е. Робастность в кластер-анализе при наличии аномальных наблюдений // *Автоматика и телемеханика*. – 1998. – № 9. – С. 72–87.
7. Журавлев Ю.И. и др. Параметрические логические модели анализа данных // *Математические методы распознавания образов*. – М.: ВЦ РАН, 1995. – С. 95–96.
8. Захаров В.Н. Интеллектуальные системы управления: основные понятия и определения // *Известия АН. Теория и системы управления*. – 1997. – № 3. – С. 138–145.
9. Компьютер и поиск компромисса: Метод достижимых целей / А.В. Лотов, В.А. Бушенков, Г.К. Каменев, О.Л. Черных. – М.: Наука, 1997. – 239 с.
10. Кувшинов Б.М., Ширяев О.В., Шапошник И.И. Диагностика заболеваний методами распознавания образов и классификации в n-мерном пространстве // *Информационные технологии*. – 2000. – № 6. – С. 43–47.
11. Кувшинов Б.М., Ширяев О.В., Богданов Д.В., Шапошник И.И., Ширяев В.И. Система классификации многопараметрических объектов для задач распознавания образов с неточной априорной информацией // *Информационные технологии*. – 2001. – № 11. – С. 37–43.
12. Кувшинов Б.М., Ширяев О.В., Богданов Д.В., Шапошник И.И., Ширяев В.И. Использование комитетов в задачах распознавания образов с неточными экспертными оценками // *Известия Челябинского научного центра*. – 2001. – № 2. – С. 12–17.
13. Лапко А.В., Лапко В.А., Ченцов С.В. Непараметрические модели распознавания образов в условиях малых выборок // *Автоматика*. – 1999. – № 6. – С. 105–113.
14. Мазуров В.Д. Метод комитетов в задачах оптимизации и классификации. – М.: Наука, 1990. – 248 с.
15. Мазуров В.Д., Хачай М.Ю. Комитетные конструкции // *Известия УрГУ. Математика и механика*. – 1999. – № 14. – С. 77–108.
16. Нильсон Н. Обучающиеся машины. – М.: Мир, 1968. – 225 с.
17. Ту Дж., Гонсалес Р. Принципы распознавания образов / Пер с англ. – М.: Мир, 1978. – 416 с.
18. Хачай М.Ю. О существовании комитета большинства // *Дискретная математика*. – 1997. – Т. 9, № 3. – С. 82–95.
19. Khachai M.Yu. Classification of committee solutions of majority // *Pattern Recognition and Image Analysis*. – 1997. – V. 7, № 2. – P. 222–230.
20. Mazurov V.I.D. Recognition and choice in a multistage procedure of modeling complex systems // *Pattern Recognition and Image Analysis*. – 1994. – V. 4, № 2. – P. 87–92.
21. Mazurov V.I.D. Generalized existence in nonequilibrium models of choice in modeling complex systems // *Pattern Recognition and Image Analysis*. – 1995. – V. 5, № 1. – P. 7–12.
22. Sven B., Jacob P. Collective decision making in hierarchies // *Math. Soc. Sci.* – 1998. – V. 35, № 3. – P. 232–244.
23. Zixing C. Typical structures for learning control // *J. Cent. S. Univ. Technol.* – 1998. – V. 5, № 1. – P. 60–63.

ЛАЗЕРНАЯ ГЕНЕРАЦИЯ ИМПУЛЬСОВ ПОВЕРХНОСТНЫХ АКУСТИЧЕСКИХ ВОЛН В ФЕРРОМАГНЕТИКЕ ПРИ МАГНИТНОМ ФАЗОВОМ ПЕРЕХОДЕ

С.Ю. Гуревич, Е.В. Голубев, К.Б. Хабиров

Данная работа посвящена теоретическому исследованию влияния магнитного фазового перехода на процесс генерации импульсов поверхностных акустических волн (волн Рэлея) лазерным импульсом в ферромагнитном материале. Получено выражение для компонент вектора смещения в термоупругом приближении. Учтена зависимость коэффициента объемного расширения от температуры.

Введение

Решение задачи о термоупругом возбуждении поверхностных акустических волн (ПАВ) в металлах хорошо известно [1]. Считалось, что при нагревании металлов лазерным импульсом, физические свойства металла не изменяются. Особенность задачи о термоупругом возбуждении ПАВ лазерным импульсом в ферромагнетике состоит в том, что при температуре магнитного фазового перехода (т. Кюри) происходит резкое изменение физических свойств ферромагнетика. Наиболее существенно изменяется коэффициент объемного расширения [2]. Так, например, в инваре при переходе через точку Кюри ($t = 207$ °С) коэффициент объемного расширения увеличивается более чем в 5 раз, а в железе ($t = 770$ °С) более чем в 4 раза. Это обстоятельство существенно усложняет задачу. Будем пренебрегать изменением других физических свойств ферромагнетика при изменении температуры. Будем считать, что вся энергия лазерного импульса поглощается поверхностью ферромагнетика (среда оптически не прозрачна) и при этом поглощаемая энергия выделяется в виде тепла. При небольшой интенсивности светового потока (до 30 МВт/см²) решение поставленной задачи сводится к решению динамической задачи термоупругости для теплового источника, действующего на поверхности полупространства. Первый этап решения задачи сводится к расчету температурного поля в ферромагнетике. Будем считать, что мы решаем несвязанную задачу, т.е. поле механических деформаций не влияет на температурное поле.

Теория

А. Распределение температуры. Лазерный импульс будем характеризовать распределением интенсивности $I(r, t)$, энергией W , временной $I_t(t)$ и пространственной $I_r(r)$ формами:

$$I(r, t) = WI_r(r)I_t(t), \quad (1)$$

Здесь r – расстояние от оси цилиндрической системы координат, t – время.

Распределение температур в полупространстве удовлетворяет уравнению теплопроводности [3]

$$\frac{\partial^2 T}{\partial r^2} + \frac{1}{r} \frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2} = \frac{1}{a} \frac{\partial T}{\partial t}, \quad (2)$$

где a – коэффициент температуропроводности; $z \geq 0$ – расстояние от поверхности до точки наблюдения внутри полупространства.

Решаем это уравнение методом интегральных преобразований Фурье–Бесселя с учетом граничного условия

$$-\lambda_q \left. \frac{\partial T}{\partial z} \right|_{z=0} = I(r, t), \quad (3)$$

где λ_q – коэффициент теплопроводности. В результате прямого преобразования получаем трансформанту температуры в виде

$$\tilde{T}^*(\lambda, z, \omega) = \frac{\tilde{T}^*(\lambda, \omega)}{2\pi\lambda_q\sqrt{\lambda^2 + i\omega/a}} \exp\left(-z\sqrt{\lambda^2 + i\omega/a}\right), \quad (4)$$

где знак «~» означает преобразование Бесселя по пространственной координате r ; а знак «*» – преобразование Фурье по времени t ; λ, ω – параметры преобразований.

Обратное преобразование Фурье–Бесселя дает распределение температуры в ферромагнетике в виде квадратур:

$$T(r, z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_0^{\infty} \tilde{T}^*(\lambda, z, \omega) J_0(\lambda r) d\lambda \right] \exp(i\omega t) d\omega, \quad (5)$$

где $J_0(\lambda r)$ – функция Бесселя первого рода нулевого порядка. Выражение (5) описывает распределение температуры в полупространстве ($z \geq 0$) для всех t и $r \geq 0$.

В. Компоненты вектора смещений в поверхностной волне. Уравнение движения упругой среды можно записать через потенциалы для продольных и поперечных волн соответственно [4]:

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} + \frac{\partial^2 \Phi}{\partial z^2} - \frac{1}{c_1^2} \frac{\partial^2 \Phi}{\partial t^2} = mf(r, z, t), \quad (6)$$

$$\frac{\partial^2 \Psi}{\partial r^2} + \frac{1}{r} \frac{\partial \Psi}{\partial r} - \frac{\Psi}{r^2} + \frac{\partial^2 \Psi}{\partial z^2} - \frac{1}{c_2^2} \frac{\partial^2 \Psi}{\partial t^2} = 0, \quad (7)$$

где $m = 3 - 4\gamma^2$; $\gamma = c_1/c_2$; $f(r, z, t) = \alpha_T(T) \cdot T(r, z, t)$; $\alpha_T(T)$ – коэффициент объемного расширения; c_1, c_2 – скорости распространения продольной и поперечной волн соответственно.

Решаем уравнения (6) и (7) методом интегральных преобразований Фурье–Бесселя. Потенциалы Φ и Ψ преобразуются по функциям Бесселя первого рода нулевого и первого порядков соответственно.

Выражения для трансформант потенциалов получаем в виде

$$\tilde{\Phi}^*(\lambda, z, \omega) = \left[B_1 + \int_0^{\infty} m\tilde{f}^* \exp(-\beta_1 z) dz \right] \exp(\beta_1 z) + \left[B_2 + \int_0^{\infty} m\tilde{f}^* \exp(\beta_1 z) dz \right] \exp(-\beta_1 z), \quad (8)$$

$$\tilde{\Psi}^*(\lambda, z, \omega) = B_3 \exp(-\beta_2 z) + B_4 \exp(\beta_2 z), \quad (9)$$

где

$$\beta_1 = \sqrt{\lambda^2 - \omega^2/c_1^2}, \quad \beta_2 = \sqrt{\lambda^2 - \omega^2/c_2^2}.$$

Коэффициенты B_1, B_2, B_3, B_4 определяются граничными условиями для механических напряжений и условием ограниченности потенциалов на бесконечности

$$\sigma_{zz}|_{z=0} = \sigma_{rz}|_{z=0} = 0, \quad \Phi|_{z=0} = \Psi|_{z=0} = 0. \quad (10)$$

Находим

$$\begin{aligned} B_1 &= -\frac{1}{2\beta_1} \int_0^{\infty} m\tilde{f}^* \exp(-\beta_1 z) dz; \\ B_2 &= \frac{1}{2\beta_1} \frac{(\lambda^2 + \beta_2^2)^2 + 4\lambda^2\beta_1\beta_2}{(\lambda^2 + \beta_2^2)^2 - 4\lambda^2\beta_1\beta_2} \int_0^{\infty} m\tilde{f}^* \exp(-\beta_1 z) dz, \\ B_3 &= \frac{1}{2\beta_1} \frac{4\lambda\beta_1(\lambda^2 + \beta_2^2)}{(\lambda^2 + \beta_2^2)^2 - 4\lambda^2\beta_1\beta_2} \int_0^{\infty} m\tilde{f}^* \exp(-\beta_1 z) dz. \end{aligned} \quad (11)$$

Выражения для смещения получают в соответствии с формулой

$$\vec{U} = \text{grad } \Phi + \text{rot } \vec{\Psi}, \quad (12)$$

откуда

$$U_r = \frac{\partial \Phi}{\partial r} - \frac{\partial \Psi}{\partial z}, \quad U_z = \frac{\partial \Phi}{\partial z} + \frac{1}{r} \frac{\partial(r\Psi)}{\partial r}. \quad (13)$$

Следовательно, трансформанты проекций вектора смещения

$$\tilde{U}_r^* = -\lambda \tilde{\Phi}^* - \frac{\partial \tilde{\Psi}^*}{\partial z}, \quad \tilde{U}_z^* = \frac{\partial \tilde{\Phi}^*}{\partial z} + \lambda \tilde{\Psi}^*. \quad (14)$$

Поверхностные волны определяются смещениями при $z = 0$:

$$\tilde{U}_r^*(\lambda, 0, \omega) = -\frac{2\lambda\beta_2(\lambda^2 - \beta_2^2)}{(\lambda^2 + \beta_2^2)^2 - 4\lambda^2\beta_1\beta_2} \int_0^\infty \tilde{m}\tilde{f}^* \exp(-\beta_1 z) dz, \quad (15)$$

$$\tilde{U}_z^*(\lambda, 0, \omega) = -\frac{\beta_2^4 - \lambda^4}{(\lambda^2 + \beta_2^2)^2 - 4\lambda^2\beta_1\beta_2} \int_0^\infty \tilde{m}\tilde{f}^* \exp(-\beta_1 z) dz. \quad (16)$$

Переход к смещениям с помощью формул обратного преобразования

$$U_r(r, 0, t) = \frac{1}{2\pi} \int_{-\infty}^\infty \left[\int_0^\infty \tilde{U}_r^*(\lambda, 0, \omega) J_1(\lambda r) d\lambda \right] \exp(i\omega t) d\omega, \quad (17)$$

$$U_z(r, 0, t) = \frac{1}{2\pi} \int_{-\infty}^\infty \left[\int_0^\infty \tilde{U}_z^*(\lambda, 0, \omega) J_0(\lambda r) d\lambda \right] \exp(i\omega t) d\omega. \quad (18)$$

В точке $\lambda = \omega/c_R$, где c_R – скорость распространения волн Рэлея, трансформанты компонент вектора смещения имеют особенность – полюс первого порядка. Вычет в этой точке определяет смещения в волне Рэлея

$$U_r(r, 0, t) = -\frac{\sqrt{c_2^2 - c_R^2}}{\pi c_2^3 c_R^3 R'} \int_{-\infty}^\infty \left[\omega^2 J_1(\omega r/c_R) \int_0^\infty \tilde{m}\tilde{f}^*(\omega/c_R, z, \omega) \exp(-z \frac{\omega}{c_R} \sqrt{1 - c_R^2/c_1^2}) dz \right] \exp(i\omega t) d\omega, \quad (19)$$

$$U_z(r, 0, t) = -\frac{(c_2^{-2} - c_R^{-2})^2 - c_R^{-4}}{2\pi R'} \times \int_{-\infty}^\infty \left[\omega^2 J_0(\omega r/c_R) \int_0^\infty \tilde{m}\tilde{f}^*(\omega/c_R, z, \omega) \exp(-z \frac{\omega}{c_R} \sqrt{1 - c_R^2/c_1^2}) dz \right] \exp(i\omega t) d\omega, \quad (20)$$

где

$$R' = \frac{1}{\omega^3} \frac{d}{d\lambda} \left[(\lambda^2 + \beta_2^2)^2 - 4\lambda^2\beta_1\beta_2 \right] \Big|_{\lambda=\omega/c_R} = \text{const}.$$

Полученные выражения являются решением поставленной задачи в виде квадратур. Частным случаем является ситуация когда коэффициент объемного расширения считается независимым от температуры. Пусть $\alpha_T(T) = \text{const}$, тогда $f(r, z, t) = \alpha_T T(r, z, t)$ и внутренний интеграл по переменной z легко вычисляется и

$$U_r(r, 0, t) = -\frac{\sqrt{c_2^2 - c_R^2} m\lambda_T R_0^2}{\pi c_2^3 c_R^3 R' 2\pi\lambda_q t_0^3} \int_{-\infty}^\infty [F_R(\omega) J_1(\omega r/c_R)] \exp(i\omega t) d\omega, \quad (21)$$

$$U_z(r, 0, t) = -\frac{(c_2^{-2} - c_R^{-2})^2 - c_R^{-4} m\lambda_T R_0^2}{2\pi R' 2\pi\lambda_q t_0^3} \int_{-\infty}^\infty [F_R(\omega) J_0(\omega r/c_R)] \exp(i\omega t) d\omega, \quad (22)$$

где

$$F_R(\omega) = \frac{\omega^2 \tilde{I}^*(\omega R_0, \omega)}{\omega^2/c_R^2 + i\omega/a + \omega\sqrt{c_R^{-2} - c_1^{-2}} \sqrt{\omega^2/c_R^2 + i\omega/a}} \quad (23)$$

– спектр рэлеевской волны. Функции Бесселя в выражениях (21) и (22) отвечают за распространение акустического импульса со скоростью c_R в сторону увеличения координаты r и уменьшение его амплитуды как $1/\sqrt{r}$. Выражение (23) связывает спектр ПАВ с распределением интенсивности в лазерном пучке.

Численное моделирование

Для вычисления параметров температурного распределения и акустического поля методами численного моделирования выберем распределение энергии в падающем излучении в виде функции Гаусса в пространстве и во времени:

$$I_r(r) = \exp(-r^2 / R_0^2) / \pi R_0^2, \quad I_t(t) = \exp(-t^2 / t_0^2) / \sqrt{\pi t_0^2}, \quad (24)$$

поскольку такое распределение является наиболее близким к реальному; $2 \cdot t_0$ – длительность лазерного импульса (по критерию e^{-1}); R_0 – радиус лазерного пятна (e^{-1}) на поверхности. Тогда

$$\tilde{T}^*(\lambda, \omega) = \frac{W}{2\pi} \exp\left(\frac{\lambda^2 R_0^2}{4} + \frac{\omega^2 t_0^2}{4}\right). \quad (25)$$

При численном интегрировании использовался метод Симпсона с переменным шагом интегрирования.

А. Распределение температуры. Выражение для трансформанты температуры следует из (4)

$$\tilde{T}^*(\lambda, z, \omega) = \frac{W}{2\pi\lambda_q \sqrt{\lambda^2 + i\omega/a}} \exp\left(\frac{\lambda^2 R_0^2}{4} + \frac{\omega^2 t_0^2}{4} - z\sqrt{\lambda^2 + i\omega/a}\right). \quad (26)$$

Распределение температуры в полупространстве определяется действительной частью выражения (5). Вычисления достаточно трудоемки, но в случае малой глубины проникновения ($\sqrt{at_0} \ll R_0$) r -профиль функции $T(r, z, t)$ при малых t хорошо описывается функцией Гаусса [5]. Это позволяет из уравнения теплопроводности (2) получить приближенную формулу:

$$T(\eta, \chi, \tau) = T_0 \cdot T_\eta(\eta) \cdot T_{\chi\tau}(\chi, \tau), \quad (27)$$

где

$$T_0 = \frac{2W\sqrt{at_0}}{\pi^2 R_0^2 \lambda_q t_0}, \quad (28)$$

$$T_\eta(\eta) = \exp\left(-\eta^2 / \sqrt{1 + 4\tau / \rho_r}\right) / \sqrt{1 + 4\tau / \rho_r}, \quad (29)$$

$$T_{\chi\tau}(\chi, \tau) = \int_0^\infty \exp\left(-\frac{\xi^4}{4} - \frac{\sqrt{2}}{2} \xi \chi\right) \cos\left(\frac{\sqrt{2}}{2} \xi \chi - \xi^2 \tau + \frac{\pi}{4}\right) d\xi \quad (30)$$

и введены безразмерные параметры:

$$\xi = \omega t_0; \quad \tau = \frac{t}{t_0}; \quad \eta = \frac{r}{R_0}; \quad \chi = \frac{z}{\sqrt{at_0}}; \quad \rho_r = \frac{R_0^2}{at_0}. \quad (31)$$

Анализ выражений (25)–(27) показывает, что T_0 можно использовать для оценки максимальной температуры в полупространстве по порядку величины.

$$T(0, 0, 0) = T_0 \frac{\sqrt[4]{4}}{4} \Gamma\left(\frac{1}{4}\right) \cos\left(\frac{\pi}{4}\right) \approx 0.906 \cdot T_0, \quad T_{\max} \approx T(0, 0, 0.541) \approx 1.076 \cdot T_0. \quad (32)$$

На рис. 1а, б представлены распределения температуры по глубине полупространства и по времени. Расчеты проведены для железа ($a = 1.78 \cdot 10^{-5}$ м/с², $\lambda_q = 70$ Вт/м·°С) при $W = 1$ мДж: $2t_0 = 30$ нс; $R_0 = 1$ мм. В этом случае $T_0 = 99.7$ К, $T_{\max} = 107.3$ К. Условно полагается, что начальная температура образца равна нулю, т.е. вычисляется разность температур. При тех же условиях возбуждения в ферромагнитом сплаве инвар (Fe + 36 % Ni, $a = 3 \cdot 10^{-6}$ м/с², $\lambda_q = 12.2$ Вт/м·°С) максимальный скачок температуры, по нашим расчетам, составляет 254.4 К. Если считать, что термоупругий механизм возбуждения акустических волн реализуется в упомянутых средах вплоть до температуры плавления $T_m \approx 1400$ – 1500 °С, то максимальная плотность теплового потока, при которой еще реализуется термоупругое возбуждение, определяется формулой (28):

$$Q_{\max} = \frac{W}{\pi R_0^2 t_0} = \frac{T_m \pi \lambda_q}{2\sqrt{at_0}}. \quad (33)$$

Для железа это 32.6, для инвара – 13.5 МВт/см².

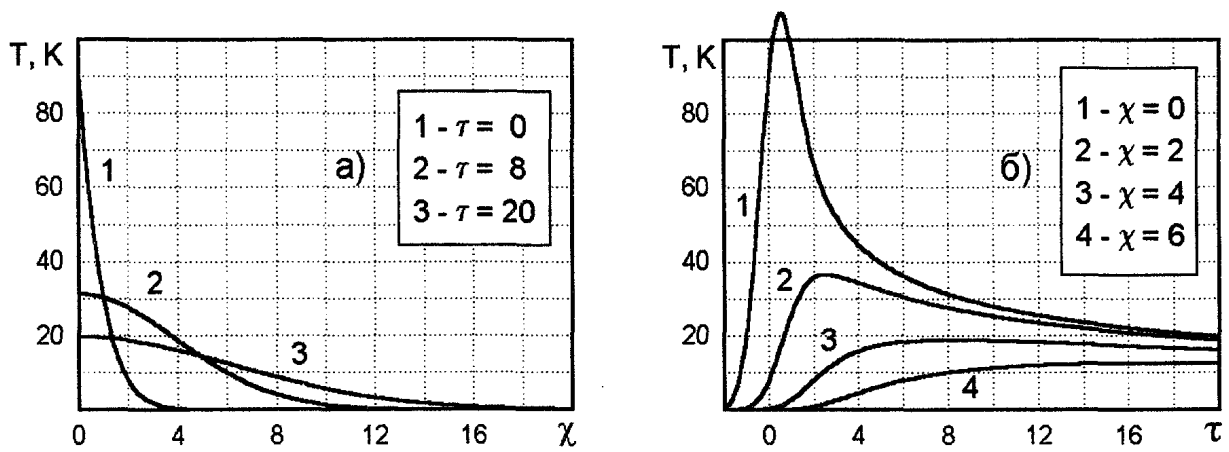


Рис. 1. а) распределение температуры по глубине χ для различных τ
 б) временное распределение температуры на различных глубинах χ

В. Компоненты вектора смещений в поверхностной волне. Предварительный численный анализ параметров акустического поля можно провести на основе относительно простых соотношений (21)–(23). На рис. 2а, б представлены акустические импульсы для железа ($\rho = 7870$ кг/м³, $c_1 = 5630$ м/с, $c_2 = 3230$ м/с, $c_R = 2749$ м/с, $m = 2.07$), полученные на основе (21) и (22) с параметрами лазерного излучения $W = 1$ мДж, $2t_0 = 30$ нс, $R_0 = 1$ мм. Их можно наблюдать в эксперименте на широкополосном (толстом) пьезопреобразователе. Часто в наблюдениях используют ЭМА датчики, которые фиксируют скорость движения поверхности (скорость смещений). Скорости смещений в акустическом импульсе представлены на рис. 3а, б. Максимуму в спектре рэлеевской волны соответствует частота 0.81 МГц (длина волны 3.95 мм). Полосы частот в 3 МГц должно быть достаточно для наблюдения таких импульсов.

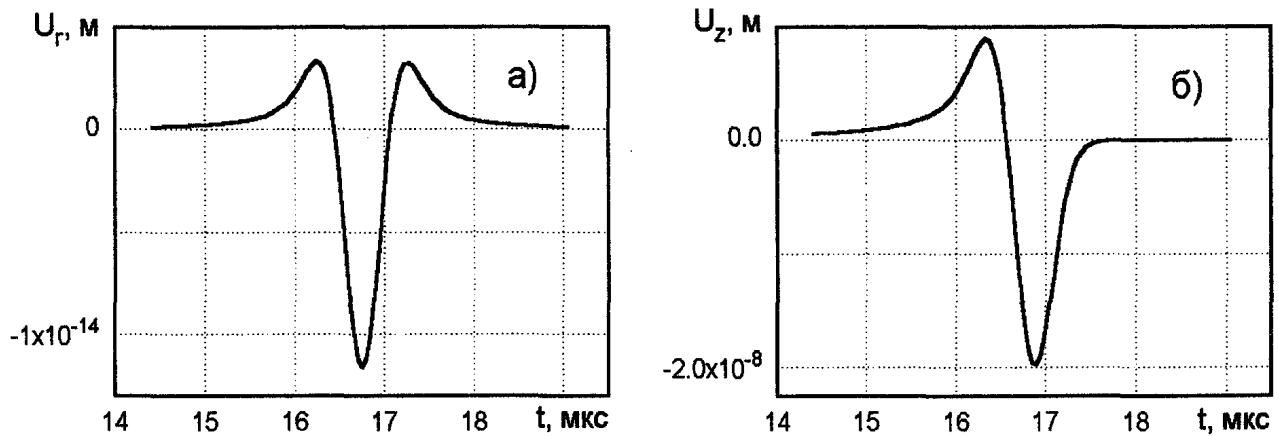


Рис. 2. Проекция вектора смещения в поверхностном акустическом импульсе: а) U_r ; б) U_z

Выводы

Получены выражения описывающие смещения в импульсе ПАВ при лазерном возбуждении с учетом зависимости коэффициента объемного расширения от температуры (19), (20). Решение из-за математических трудностей не удастся представить в аналитической форме, поэтому дальнейшее исследование проводится методами численного интегрирования. Сделана оценка для границы термоупругого механизма возбуждения ПАВ импульсным оптическим излучением (33), совпадающая с точностью до $\sqrt{\pi}$ с [5]. По нашим расчетам, для железа она составляет 32.6, для ферромагнитного сплава инвар – 13.5 МВт/см². Экспериментально определяемая граница для инвара 30–40 МВт/см². Длительность наблюдаемых импульсов примерно в два раза меньше. Несответствие можно объяснить неточностью определения радиуса лазерного пучка и энергии им-

пульса. Форма рассчитанных импульсов (рис. 3б) прекрасно согласуется с наблюдаемой в эксперименте на инваре.

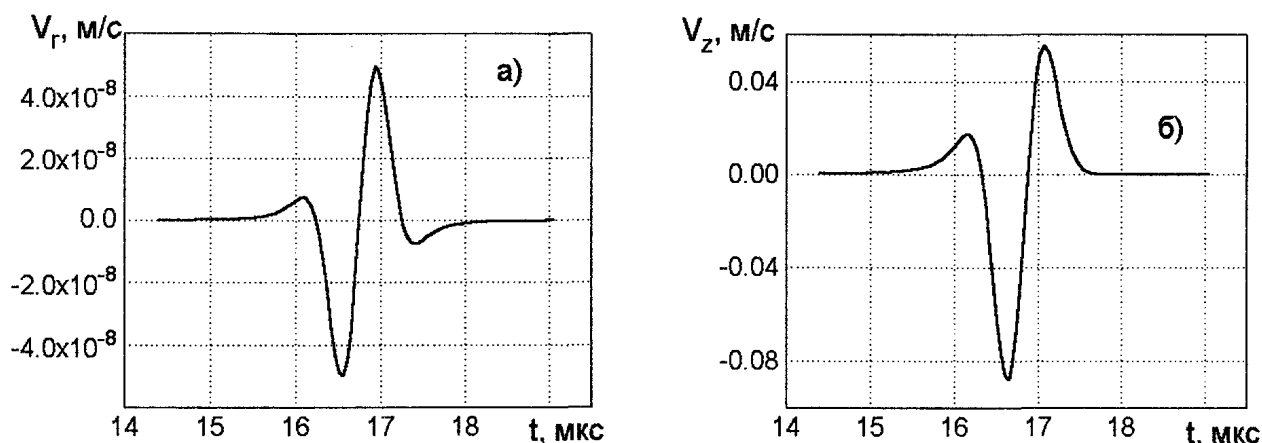


Рис. 3. Компоненты вектора скорости смещений в акустическом импульсе: а) V_x ; б) V_z

Благодарности

Авторы выражают глубокую благодарность Шульгинову А.А. за помощь и полезные замечания, а также Петрову Ю.В. за сотрудничество и предоставленные экспериментальные данные.

Работа выполнена при поддержке Минобразования РФ (Тематический план, № 802).

Литература

1. Лямшев Л.М. Лазерное термооптическое возбуждение звука. – М.: Наука, 1989. – 219 с.
2. Лариков Л.Н., Усов Ю.В. Тепловые свойства железоникелевых сплавов // *Металлофизика*. – 1977. – № 68. – С. 28–35.
3. Тихонов А.Н., Самарский А.А. Уравнения математической физики. – М.: Наука, 1972. – 735 с.
4. Гуревич С.Ю., Петров Ю.В., Прокопьев К.В., Шульгинов А.А. Исследование влияния магнитного фазового перехода на спектр акустических импульсов, возбуждаемых лазерным импульсом в ферромагнетике // *Акустический журнал*. – 1999. – Т. 45, № 4. – С. 494–498.
5. Апполонов В.В., Прохоров А.М., Хомич В.Ю., Четкин С.А. Термоупругое воздействие импульсно-периодического излучения на поверхность твердого тела // *Квантовая электроника*, 9, № 2 (1982).

СТРУКТУРА ЭЛЕКТРОМАГНИТНОГО ПОЛЯ, НАВЕДЕННОГО НАД ПОВЕРХНОСТЬЮ ФЕРРОМАГНИТНОГО ПОЛУПРОСТРАНСТВА, ПРИ ПАДЕНИИ СФЕРИЧЕСКОГО ФРОНТА АКУСТИЧЕСКОЙ ВОЛНЫ

К.В. Прокопьев, А.А. Шульгинов, С.Ю. Гуревич

В работе решена задача обратного электромагнито-акустического преобразования для ферромагнитной среды находящейся при температуре Кюри. Расчет выполнен для сферического фронта падающей акустической волны. Вычислены параметры электромагнитных волн, возникающих над поверхностью ферромагнитного полупространства.

При взаимодействии переменного электромагнитного поля с поверхностью проводящей среды в последней возникает акустическое поле. Если упругая среда является ферромагнитной, то акустическое поле обладает рядом особенностей. При температурах, соответствующих магнитному фазовому переходу (точка Кюри), обратное электромагнито-акустическое (ЭМА) преобразование происходит за счет объемной магнитострикции в следствии отсутствия анизотропии магнитоупругих и магнитных свойств среды. Задача обратного ЭМА преобразования успешно решена для плоского фронта падающей акустической волны [2]. В развитие подходов, сформулированных в этой работе, решим аналогичную задачу для падающей акустической волны, имеющей сферический фронт.

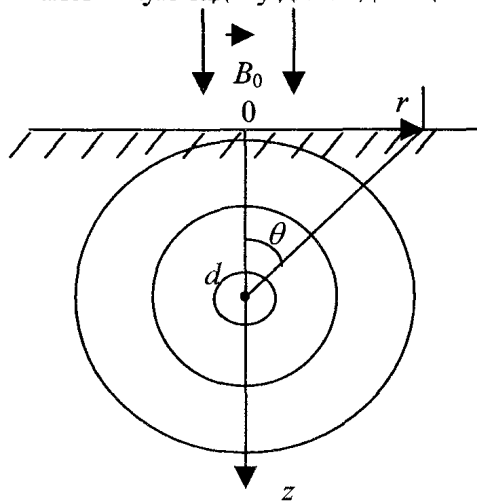


Рис. 1. Система координат задачи

Пусть на плоскую границу ферромагнитной среды, заполняющую полупространство $Z > 0$, из точки d падает сферическая волна частотой ω и амплитудой потенциала W (рис. 1). Среда магнитополяризована внешним магнитным полем B_0 . В результате действия упругой волны на элементы ферромагнитной среды возникает изменение намагниченности среды, которая создает переменное магнитное поле, последнее над поверхностью полупространства может быть зарегистрировано с помощью электромагнито-акустических преобразователей.

Смещение частиц твердого тела представим через скалярный и векторный потенциал. Уравнения запишем для цилиндрической системы координат

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} + \frac{\partial^2 \Phi}{\partial z^2} - \frac{1}{C_l^2} \frac{\partial^2 \Phi}{\partial t^2} = W \frac{\delta(r)}{2\pi r} \delta(z-d) e^{-i\omega t}; \quad (1)$$

$$\frac{\partial^2 \Psi}{\partial r^2} + \frac{1}{r} \frac{\partial \Psi}{\partial r} - \frac{\Psi}{r^2} + \frac{\partial^2 \Psi}{\partial z^2} - \frac{1}{C_t^2} \frac{\partial^2 \Psi}{\partial t^2} = 0. \quad (2)$$

Граничные условия, записанные для границы металл-воздух

$$\sigma_{zz}|_{z=0} = 0, \quad \sigma_{rz}|_{z=0} = 0$$

преобразуем через потенциалы

$$(1 - 2\vartheta) \left(\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} \right) + \frac{\partial^2 \Phi}{\partial z^2} + 2\vartheta \left(\frac{\partial^2 \Psi}{\partial r \partial z} + \frac{1}{r} \frac{\partial \Psi}{\partial z} \right) = 0; \quad (3)$$

$$2 \frac{\partial^2 \Phi}{\partial r \partial z} - \frac{\partial^2 \Psi}{\partial z^2} + \frac{\partial^2 \Psi}{\partial r^2} + \frac{1}{r} \frac{\partial \Psi}{\partial r} - \frac{\Psi}{r^2} = 0, \quad (4)$$

где $\vartheta = \frac{C_t^2}{C_l^2}$, C_t , C_l – скорости поперечной и продольной волны в среде.

Решение системы (1), (2) будем искать в виде

$$\Phi(z, r, t) = \Phi^*(z, r)e^{-i\omega t}, \Psi(z, r, t) = \Psi^*(z, r)e^{-i\omega t}.$$

С учетом этого система (1), (2) примет вид

$$\frac{\partial^2 \Phi^*}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi^*}{\partial r} + \frac{\partial^2 \Phi^*}{\partial z^2} + k_l^2 \Phi^* = W \frac{\delta(r)}{2\pi r} \delta(z-d), \quad (5)$$

$$\frac{\partial^2 \Psi^*}{\partial r^2} + \frac{1}{r} \frac{\partial \Psi^*}{\partial r} - \frac{\Psi^*}{r^2} + \frac{\partial^2 \Psi^*}{\partial z^2} + k_t^2 \Psi^* = 0, \quad (6)$$

$$k_l = \frac{\omega}{C_l}, \quad k_t = \frac{\omega}{C_t}.$$

Решение системы (5), (6) будем искать методом интегральных преобразований Бесселя, где

$$\tilde{\Phi}^*(z, k) = \frac{1}{2\pi} \int_0^\infty \Phi^*(z, k) J_0(kr) r dr; \quad \tilde{\Psi}^*(z, k) = \frac{1}{2\pi} \int_0^\infty \Psi^*(z, k) J_1(kr) r dr -$$

прямое, а

$$\Phi^*(z, r) = \int_0^\infty \tilde{\Phi}^*(z, k) J_0(kr) k dk; \quad \Psi^*(z, r) = \int_0^\infty \tilde{\Psi}^*(z, k) J_1(kr) k dk -$$

обратное преобразование. Учитывая, что функции Бесселя $J_0(x)$ и $J_1(x)$ удовлетворяют уравнению Бесселя

$$\frac{d^2 J_0}{dx^2} + \frac{1}{x} \frac{dJ_0}{dx} + J_0 = 0, \quad \frac{d^2 J_1}{dx^2} + \frac{1}{x} \frac{dJ_1}{dx} + J_1 - \frac{J_1}{x^2} = 0,$$

получим

$$\frac{d^2 \tilde{\Phi}^*}{dz^2} - (k^2 - k_l^2) \tilde{\Phi}^* = \frac{W}{4\pi^2} \delta(z-d); \quad (7)$$

$$\frac{d^2 \tilde{\Psi}^*}{dz^2} - (k^2 - k_t^2) \tilde{\Psi}^* = 0. \quad (8)$$

Запишем общее решение уравнений (7), (8)

$$\tilde{\Phi}^*(z, k) = C_1 e^{-z\sqrt{k^2 - k_l^2}} + C_1' e^{-z\sqrt{k^2 - k_l^2}} + \tilde{\Phi}_1^*(z, k); \quad (9)$$

$$\tilde{\Psi}^*(z, k) = C_2 e^{-z\sqrt{k^2 - k_t^2}}, \quad (10)$$

где $\tilde{\Phi}_1^*(z, k)$ – частное решение неоднородного уравнения (7). Второе слагаемое в уравнении (9) дает расходящееся решение, тогда $C_1' = 0$. Учитывая соотношение

$$\delta(z-d) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\gamma(z-d)} d\gamma,$$

и введя прямое и обратное преобразование Фурье

$$\tilde{\Phi}_1^* = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi(\gamma, k) e^{i\gamma k} d\gamma, \quad \Phi(\gamma, k) = \int_{-\infty}^{+\infty} \tilde{\Phi}_1^*(z, k) e^{-i\gamma k} dz,$$

уравнение (7) преобразуется к виду

$$\Phi = -\frac{W}{4\pi^2} \frac{e^{-i\gamma d}}{\gamma^2 + k^2 - k_l^2}. \quad (11)$$

Сделаем прямое преобразование Бесселя уравнения (11)

$$\tilde{\Phi}_1^*(z, k) = -\frac{W}{8\pi^3} \int_{-\infty}^{+\infty} \frac{e^{-i\gamma(d-z)}}{\gamma^2 + k^2 - k_l^2} d\gamma = -\frac{W}{8\pi^3} \int_{-\infty}^{+\infty} \frac{e^{-i\gamma(d-z)}}{(\gamma - \sqrt{k_l^2 - k^2})(\gamma + \sqrt{k_l^2 - k^2})} d\gamma. \quad (12)$$

Как видно из рис. 1, нас интересует решение при $z < d$. Следовательно, $(d-z) > 0$. Введем обозначения

$$\gamma_1 = i\sqrt{k^2 - k_l^2}, \quad \gamma_2 = -i\sqrt{k^2 - k_l^2}.$$

Для нахождения интеграла (12) дополним область интегрирования до замкнутого контура.

Тогда

$$\int_{-\infty}^{+\infty} \frac{e^{-i\gamma(d-z)}}{(\gamma-\gamma_1)(\gamma-\gamma_2)} d\gamma = \oint \frac{e^{-i\gamma(d-z)}}{(\gamma-\gamma_1)(\gamma-\gamma_2)} d\gamma = 2\pi i \sum \text{res} \frac{e^{-i\gamma(d-z)}}{(\gamma-\gamma_1)(\gamma-\gamma_2)} = 2\pi i \frac{e^{-i\gamma(d-z)}}{\gamma_2-\gamma_1}.$$

Найдем частное решение

$$\tilde{\Phi}_1^*(z, k) = -\frac{W}{8\pi^2 \sqrt{k^2 - k_l^2}} e^{-(d-z)\sqrt{k^2 - k_l^2}}, \quad z < d.$$

Общее решение уравнений (9), (10):

$$\tilde{\Phi}^*(z, k) = C_1^{(k)} e^{-z\sqrt{k^2 - k_l^2}} - \frac{W}{8\pi^2 \sqrt{k^2 - k_l^2}} e^{-(d-z)\sqrt{k^2 - k_l^2}}; \quad (13)$$

$$\tilde{\Psi}^*(z, k) = C_2^{(k)} e^{-z\sqrt{k^2 - k_l^2}}. \quad (14)$$

Первое слагаемое в уравнении (13) – отраженная l -волна, второе – падающая l -волна, уравнение (14) – отраженная S -волна. Коэффициенты $C_1(k)$ и $C_2(k)$ находятся из граничных условий (3), (4). В результате получим:

$$C_1(k) = \frac{W}{8\pi^2} e^{-d\sqrt{k^2 - k_l^2}} \frac{\Delta_1^*}{\Delta} = \frac{W}{8\pi^2} e^{-d\sqrt{k^2 - k_l^2}} C_1^*(k),$$

$$C_2(k) = \frac{W}{8\pi^2} e^{-d\sqrt{k^2 - k_l^2}} \frac{\Delta_2^*}{\Delta} = \frac{W}{8\pi^2} e^{-d\sqrt{k^2 - k_l^2}} C_2^*(k),$$

$$\Delta = (2\vartheta k^2 - k_l^2)(2k^2 - k_l^2) - 4k^2 \vartheta \sqrt{k^2 - k_l^2} \sqrt{k^2 - k_l^2};$$

$$\Delta_1^* = \frac{(2\vartheta k^2 - k_l^2)(2k^2 - k_l^2) + 4k^2 \vartheta \sqrt{k^2 - k_l^2} \sqrt{k^2 - k_l^2}}{\sqrt{k^2 - k_l^2}};$$

$$\Delta_2^* = 4k(2\vartheta k^2 - k_l^2)$$

Общее решение для амплитуд потенциалов продольной и поперечной волн запишем следующим образом:

$$\Phi^*(z, r) = \frac{W}{8\pi^2} \int_0^\infty C_1^*(k) e^{-(z+d)\sqrt{k^2 - k_l^2}} J_0(kr) k dk - \frac{A}{8\pi^2} \int_0^\infty \frac{k}{\sqrt{k^2 - k_l^2}} e^{-(z+d)\sqrt{k^2 - k_l^2}} J_0(kr) k dk; \quad (15)$$

$$\Psi^*(z, r) = \frac{W}{8\pi^2} \int_0^\infty C_2^*(k) e^{-z\sqrt{k^2 - k_l^2} - d\sqrt{k^2 - k_l^2}} J_1(kr) k dk. \quad (16)$$

В акустическом поле на грани элементарного объема, выделенного в среде, действуют нормальные и касательные напряжения, однако последние вызывают лишь изменение формы объема, величина объема при этом не меняется. Следовательно, касательные напряжения и соответствующие им касательные составляющие смещений на поверхности среды не вызовут изменение объема и, следовательно, не приведут к изменению намагниченности за счет объемной магнито-стрикции. По этой причине далее рассматриваются только нормальные к поверхности среды составляющие смещений. Найдем z -компоненту смещения U_z , которая связана с потенциалами следующим образом

$$U_z(z, r, t) = \frac{\partial \Phi}{\partial z} + \frac{\partial \Psi}{\partial r} + \frac{\Psi}{r},$$

для трансформант Бесселя получим

$$\tilde{U}_z^*(z, r) = \left[\frac{d\tilde{\Phi}^*}{dz} + k\tilde{\Psi}^* \right]. \quad (17)$$

При наложении на ферромагнетик периодически меняющихся во времени внешних напряжений под действием упругой волны в нем происходит обратимые и необратимые смещения границ доменов и процесс смещения вектора намагниченности. В области парапроцесса в районе

температуры Кюри намагниченность насыщения меняется [1], что приводит к генерации электромагнитной волны.

Пусть ферромагнитное полупространство находится в постоянном магнитном поле, направленном вдоль оси z . Падающая акустическая волна вызывает смещение границ доменов, что приводит к генерации электромагнитной волны.

Запишем уравнения Максвелла

$$\operatorname{rot} \vec{B} = \mu\mu_0(\vec{j}_0 - \vec{j}_{cm}) = \mu\mu_0(\vec{j}_0 - \varepsilon\varepsilon_0 \frac{\partial \vec{E}}{\partial t}); \quad (18)$$

$$\operatorname{div} \vec{B} = 0; \quad (19)$$

$$\operatorname{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}; \quad (20)$$

$$\operatorname{div} \vec{E} = 0; \quad (21)$$

$$\vec{j}_0 = \vec{j} + \sigma \vec{E}, \quad (22)$$

где \vec{j}_0 – полная плотность тока внутри ферромагнетика; \vec{j} – плотность тока, вызванного смещением границ доменов; $\sigma \vec{E}$ – плотность тока, индуцированного электромагнитным полем. Введем вектор потенциал A ,

$$\vec{B} = \operatorname{rot} \vec{A}, \quad \vec{E} = -\frac{\partial \vec{A}}{\partial t}.$$

Тогда уравнение (18) можно представить в виде

$$\operatorname{rot} \operatorname{rot} \vec{A} = \mu\mu_0 \vec{j} + \mu\mu_0 \sigma \vec{E} + \mu\mu_0 \varepsilon\varepsilon_0 \frac{\partial \vec{E}}{\partial t}$$

или

$$\operatorname{rot} \operatorname{rot} \vec{A} = \mu\mu_0 \vec{j} - \mu\mu_0 \sigma \frac{\partial \vec{A}}{\partial t} - \frac{1}{v^2} \frac{\partial \vec{A}}{\partial t}, \quad (23)$$

где

$$v^2 = \frac{1}{\mu\mu_0 \varepsilon\varepsilon_0} = \frac{c^2}{n^2}.$$

Решение уравнения (23) ищем в виде

$$\operatorname{rot} \operatorname{rot} \vec{A}^* = \mu\mu_0 \vec{j}^* - \mu\mu_0 \sigma i\omega \vec{A}^* + k_0^2 \vec{A}^*, \quad (24)$$

где

$$\vec{A} = \vec{A}^* e^{-i\omega t}; \vec{j} = \vec{j}^* e^{-i\omega t}; k_0 = \frac{\omega}{v}.$$

В цилиндрических координатах уравнение (24), записанное для ферромагнетика и воздуха, будет выглядеть следующим образом

$$\frac{\partial^2 A_{\phi}^*}{\partial z^2} + \frac{\partial^2 A_{\phi}^*}{\partial r^2} + \frac{1}{r} \frac{\partial A_{\phi}^*}{\partial r} + \frac{A_{\phi}^*}{r^2} + k_{\phi}^2 A_{\phi}^* = F^*; \quad (25)$$

$$\frac{\partial^2 A_s^*}{\partial z^2} + \frac{\partial^2 A_s^*}{\partial r^2} + \frac{1}{r} \frac{\partial A_s^*}{\partial r} - \frac{A_s^*}{r^2} + k_s^2 A_s^* = 0, \quad (26)$$

где

$$F^* = -\mu\mu_0 j = -\mu\mu_0 \operatorname{rot} \vec{M} \Big|_{\vec{e}_r}; \quad k_{\phi}^2 = (i\omega \sigma \mu\mu_0 + k_0^2), \quad k_s = \frac{\omega}{C},$$

A_{ϕ}^* , A_s^* – амплитуды вектор-потенциала в ферромагнетике и в воздухе.

Из условий равенства тангенциальных компонент электрического и магнитного поля граничные условия запишутся в виде

$$A_s^* \Big|_{z=0} = A_{\phi}^* \Big|_{z=0}; \quad (27)$$

$$\left. \frac{dA_g^*}{dz} \right|_{z=0} = \frac{1}{\mu} \left. \frac{dA_\phi^*}{dz} \right|_{z=0}. \quad (28)$$

Используя интегральные преобразование Бесселя получим

$$\frac{d^2 \tilde{A}_\phi^*}{dz^2} - (\beta^2 - k_\phi^2) \tilde{A}_\phi^* = \tilde{F}^*(z, \beta); \quad (29)$$

$$\frac{d^2 \tilde{A}_g^*}{dz^2} - (\beta^2 - k_g^2) \tilde{A}_g^* = 0. \quad (30)$$

Как показано в работе [2], связь между упругими и магнитными константами можно представить в виде

$$F^* = -\mu\mu_0 \operatorname{rot} \vec{M} \Big|_{\vec{e}_r} = b \frac{\partial^2 U_z}{\partial z^2},$$

где

$$b = \mu(\hat{\lambda} + \frac{2}{3}\hat{\mu}) \frac{\partial \lambda_v}{\partial H_{0z}},$$

$\hat{\lambda}, \hat{\mu}$ – постоянные Ламэ; λ_v – коэффициент объемной магнитострикции. Для трансформанты Бесселя получим

$$\begin{aligned} \tilde{F}^*(z, k) = & \tilde{N}_1 \exp(-(z+d)\sqrt{k^2 - k_l^2}) + \tilde{N}_2 \exp(-(d-z)\sqrt{k^2 - k_l^2}) + \\ & + \tilde{N}_3 \exp(-z\sqrt{k^2 - k_l^2} - d\sqrt{k^2 - k_l^2}), \end{aligned} \quad (31)$$

где

$$\tilde{N}_1(k) = \frac{Wb}{8\pi^2} C_1^*(k)(k^2 - k_l^2)^{\frac{3}{2}} - \text{амплитуда отраженной продольной волны};$$

$$\tilde{N}_2(k) = \frac{Wb}{8\pi^2} (k^2 - k_l^2) - \text{амплитуда падающей продольной волны};$$

$$\tilde{N}_3(k) = \frac{Wb}{8\pi^2} C_2^*(k)k(k^2 - k_l^2) - \text{амплитуда отраженной поперечной волны}.$$

Запишем общее решение уравнений (25), (26)

$$\tilde{A}_g^*(z, \beta) = D_1 \exp(z\sqrt{\beta^2 - k_g^2}); \quad (32)$$

$$\tilde{A}_\phi^*(z, \beta) = D_2 \exp(-z\sqrt{\beta^2 - k_\phi^2}) + \tilde{G}(z, \beta), \quad (33)$$

где \tilde{G} – частное решение уравнения (25). Для нахождения коэффициентов в уравнении (32), (33) используем граничные условия (27), (28). Частное решение определяем подстановкой в (29). В результате получим выражение для трансформанты вектор потенциала в воздухе

$$\tilde{A}_g^*(z, \beta) = E(\beta) \exp(z\sqrt{\beta^2 - k_g^2} - d\sqrt{\beta^2 - k_l^2}), \quad (34)$$

$$E(\beta) = \frac{i}{2\pi} \frac{S_1(\beta) \frac{\sqrt{\beta^2 - k_\phi^2}}{\mu} + S_2(\beta)}{(k_\phi^2 - \beta^2) \left(\frac{\sqrt{\beta^2 - k_\phi^2}}{\mu} + \sqrt{\beta^2 - k_g^2} \right)},$$

где

$$S_1(\beta) = \tilde{N}_1(\beta) + \tilde{N}_2(\beta) + \tilde{N}_3(\beta),$$

$$S_2(\beta) = \sqrt{\beta^2 - k_l^2} (\tilde{N}_2(\beta) - \tilde{N}_1(\beta)) - \sqrt{\beta^2 - k_l^2} \tilde{N}_3(\beta).$$

Применив к выражению (34) обратное преобразование Бесселя, все интегралы берутся методом перевала, окончательно получим выражение для напряженности электромагнитного поля в воздухе на поверхности полупространства

$$E^*(z=0, r) = \frac{1}{d} \omega k_l \cos \theta E(k_l \sin \theta). \quad (35)$$

Результаты расчетов приведены на рис. 2, 3. В качестве ферромагнетика выбран железоникелевый сплав 32 НКД. Из анализа графиков видно, что максимум напряженности электромагнитного поля наблюдается в диапазоне частот от 2 МГц, при узком угле наблюдения 4–10 град. С ростом частоты эффективность преобразования усиливается, но при этом становится существенным поглощение в среде (в расчетах не учитывалось), что будет ограничивать верхний диапазон эффективно преобразуемых частот.

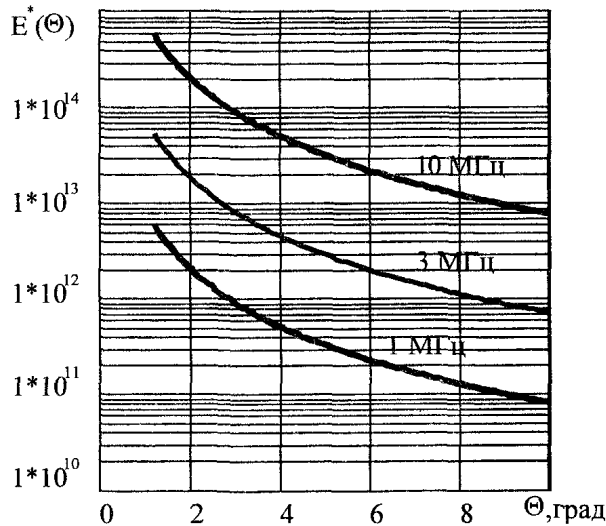


Рис. 2. Зависимость амплитуды E от Θ для различных частот падающей сферической волны ($d = 5$ см)

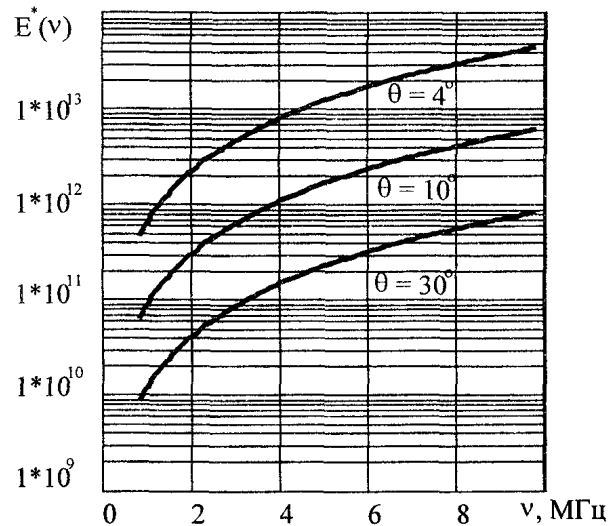


Рис. 3. Зависимость амплитуды E от ν для различных углов наблюдения ($d = 5$ см)

Полученные результаты могут служить основой при конструировании ЭМА преобразователей, где главной задачей является подбор размеров и формы самого преобразователя.

Работа выполнена при поддержке РФФИ (грант № 01-02-96469) и Минобразования РФ (грант № Е00-3.4-346).

Литература

1. Вонсовский С.В. Магнетизм. – М.: Наука, 1971. – 1032 с.
2. Гуревич С.Ю., Писарев М.С. Акусто-электромагнитное преобразование в ферромагнитных металлах при магнитном фазовом переходе // *Физические методы испытаний материалов и веществ: Тематич. сб. научн. тр.* – Челябинск, ЧГТУ, 1995. – С. 3–15.

ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ ПРОДОЛЬНЫХ УЛЬТРАЗВУКОВЫХ ВОЛН, ВОЗБУЖДАЕМЫХ ТЕРМОАКУСТИЧЕСКИМ ИЗЛУЧАТЕЛЕМ

Ю.В. Петров, А.В. Шушарин, В.Ф. Божко

Представлены результаты экспериментальных исследований по изучению импульсов продольных волн, возбуждаемых в металле коротким лазерным импульсом. Исследовано влияние интенсивности излучения лазера, размера пятна облучения, угла излучения на амплитуду, форму и длительность импульса. Результаты обсуждены, исходя из модели термоакустического излучателя в виде нагретого лазерным импульсом тонкого диска металла, в котором вследствие действия термоупругого и гидродинамического механизмов генерации, возникает импульс радиальных и осевых механических напряжений.

Использование ОКГ для возбуждения ультразвука в металлах в сочетании с оптимальной регистрацией коротких ультразвуковых импульсов дает возможность создать контролируемую аппаратуру с повышенной разрешающей способностью и уменьшенной «мертвой» зоной. С этой целью в настоящей работе проведены экспериментальные исследования продольных ультразвуковых волн, возбуждаемых термооптическим излучателем.

Диаграмма направленности термооптического излучателя

Эффективная работа аппаратуры для лазерной генерации и бесконтактной регистрации ультразвука с целью контроля материалов и изделий достигаются в результате оптимизации их конструкции, основанной, как правило, на исследовании характеристик направленности термооптических излучателей. Исследования проводились на стенде, описанном в [1]. В качестве образца использовался полуцилиндр из сплава Д-16 радиусом 40 мм. Излучение лазера фокусировалось в центр грани в пятно нужного диаметра. Прием осуществлялся либо широкополосным электромагнитоакустическим преобразователем (ЭМАП) с высокочастотным индуктором в виде прямого проводника длиной 15 мм, расположенного параллельно образующей полуцилиндра, либо поверхностно возбуждаемым толстым пьезопреобразователем (ПВП) [2]. Результаты эксперимента приведены на рис. 1а.

При малой интенсивности лазерного излучения, когда гидродинамический механизм генерации ультразвука еще не достиг своего максимума, амплитуда акустического импульса с ростом угла наблюдения меняется слабо, незначительно возрастает и уходит в ноль при угле наблюдения 90° . При больших значениях интенсивности лазерного излучения в диапазоне действия гидродинамического и испарительного механизмов генерации, амплитуда с увеличением угла наблюдения сначала несколько уменьшается. Но затем она увеличивается, достигает максимума в районе 30° и потом монотонно спадает. В диапазоне $60-70^\circ$ спад амплитуды замедляется, она остается почти без изменений, после этого уменьшается до нуля при 90° .

Полученные диаграммы направленности можно качественно объяснить, если принять за модель термооптического излучателя нагретый лазерным импульсом диск на поверхности металла. Толщина диска меняется в зависимости от глубины распространения тепла в металл и составляет несколько микрометров. Основание диска равно сечению лазерного луча и его диаметр измеряется несколькими миллиметрами. Дно диска можно принять за поршневой излучатель импульсов сжатия и растяжения. Боковая поверхность диска нагрева является источником радиальных напряжений сжатия. Акустический импульс, излучаемый боковой поверхностью, можно представить как сумму импульсов от ее ближней и дальней частей, причем импульс от дальней части запаздывает на акустическую разность хода и имеет противоположную полярность. В результате в точку приема от боковой поверхности диска приходит также импульс сжатия и растяжения. Таким образом, импульсы от дна и от боковой поверхности диска нагрева имеют одинаковую полярность и длительность, но форма у них будет несколько разная.

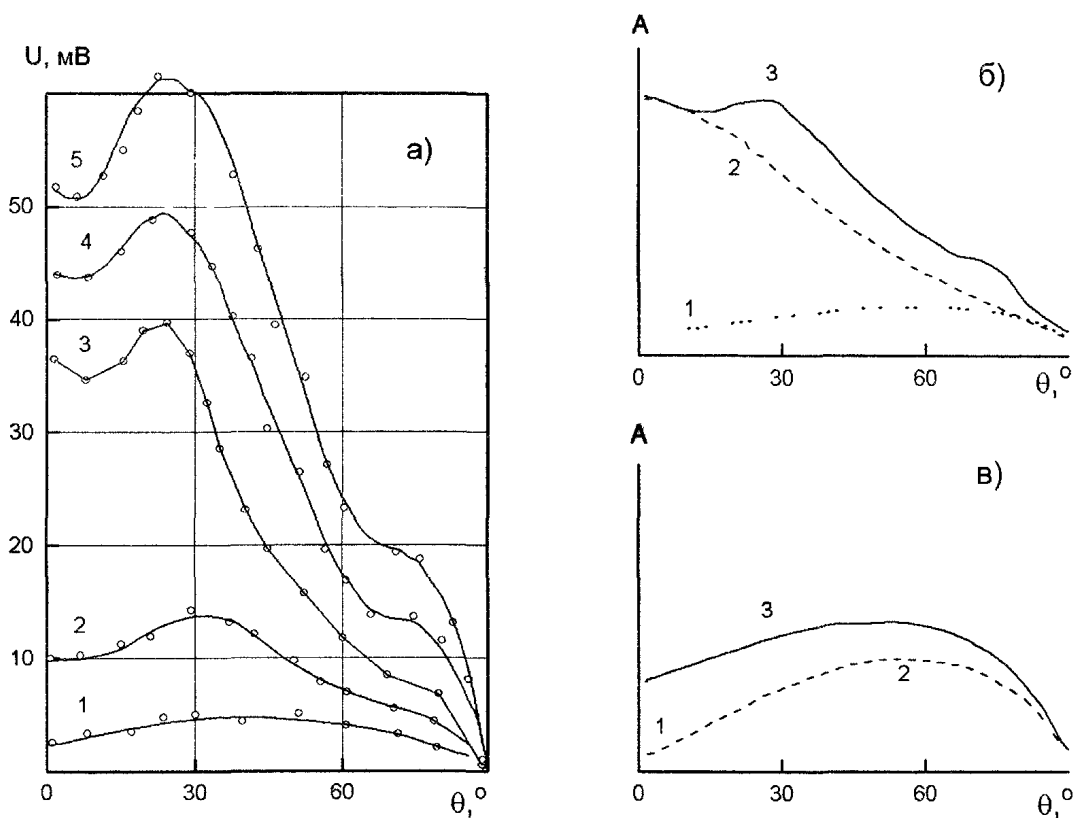


Рис. 1. а) экспериментальные диаграммы направленности термооптического излучателя диаметром 2 мм при разных интенсивностях лазерного излучения: 1 – 63, 2 – 120, 3 – 230, 4 – 300, 5 – 500 МВт/см²; б), в) качественное сложение диаграмм направленности от дна (2) и боковой поверхности (1) диска нагрева при интенсивности лазерного излучения соответственно больше и меньше 120 МВт/см²

Итак, диаграмму направленности термооптического излучателя можно представить как суперпозицию диаграмм направленности дна и боковой поверхности. При термоупругом механизме генерации ультразвука главную роль играет излучение от боковой поверхности диска нагрева. На оси термооптического излучателя амплитуда излучения небольшая, так как обусловлена излучением только от дна. С увеличением угла наблюдения амплитуда излучения от дна падает, но зато увеличивается амплитуда излучения от боковой поверхности диска. Результат сложения этих двух излучений и будет определять результирующую диаграмму направленности (рис. 1в). Она неплохо согласуется с экспериментальной диаграммой направленности при малой интенсивности лазерного излучения. Чем меньше интенсивность лазерного излучения, тем меньше влияние излучения от дна и тем ближе максимум диаграммы направленности будет располагаться к максимуму излучения от боковой поверхности (угол 65°).

При реализации гидродинамического и испарительного механизмов генерации ультразвука, определяющим является излучение от дна диска, поэтому результат сложения акустических импульсов от дна со сравнительно большой амплитудой и от боковой поверхности с малой амплитудой будет определять результирующую диаграмму направленности (рис. 1б). Она хорошо согласуется с экспериментальной диаграммой направленности для больших интенсивностей лазерного излучения. Медленный спад амплитуды в области 60–70° обусловлен максимумом излучения от боковой поверхности диска.

Были проведены исследования по влиянию размеров лазерного пятна на диаграмму направленности при постоянной интенсивности лазерного излучения. С этой целью лазерный луч перекрывался диафрагмой с отверстиями разного диаметра. На рис. 2а приведены результаты эксперимента. Видно, что чем больше диаметр пятна, тем больше амплитуда акустических импульсов. Качественно это можно объяснить также с помощью модели термооптического излучателя. В направлении оси диска нагрева излучает только дно, и амплитуда этого излучения пропорциональна площади дна диска (или квадрату диаметра пятна (рис. 2в)). В направлении угла наблюдения 70° излучением от дна можно пренебречь по сравнению с излучением от боковой поверхности. Тогда амплитуда результирующего излучения должна быть пропорциональна площади бо-

ковой поверхности или диаметру пятна (рис. 2б). Таким образом, проведенный эксперимент подтверждает наличие двух излучающих ультразвук источников.

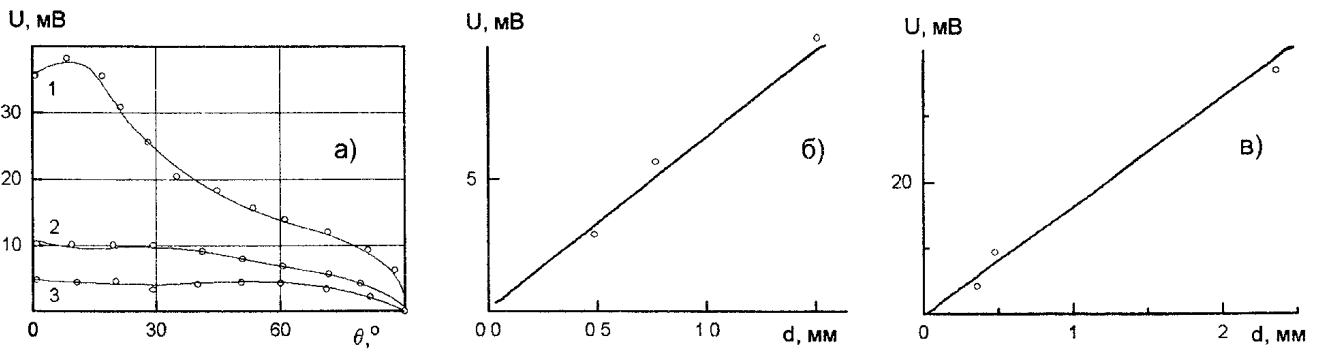


Рис. 2. а) диаграммы направленности термооптического излучателя диаметром 1 – 1.5 мм, 2 – 0.8 мм, 3 – 0.5 мм при интенсивности лазерного излучения 300 МВт/см²; б), в) зависимости амплитуды акустического импульса от диаметра и его квадрата светового пятна для $\theta = 70^\circ$ и $\theta = 0^\circ$

Длительность акустических импульсов, излучаемых термооптическим источником

Результаты эксперимента по выяснению зависимости длительности импульсов, излучаемых термооптическим источником ультразвука, от угла наблюдения приведены на рис. 3. Видно, что акустический импульс состоит из отрицательного полупериода длительностью примерно 50 нс и положительного полупериода, втрое большей длительности и в два раза меньше по амплитуде. Это хорошо согласуется с моделью термооптического излучателя, если считать, что акустические импульсы, излучаемые дном и боковой поверхностью диска нагрева являются импульсами сжатия–растяжения, незначительно отличающимися по форме. Накладываясь, они и дают результирующий импульс сжатия–растяжения. Длительность такого импульса равна сумме длительностей импульса напряжений τ_0 в диске нагрева и времени задержки импульса от дальней части диска нагрева относительно импульса от его ближней части, т.е.

$$\tau = \tau_0 + \frac{d \sin \theta}{V_l}$$

где d – диаметр светового пятна; θ – угол наблюдения; V_l – скорость продольных волн в образце.

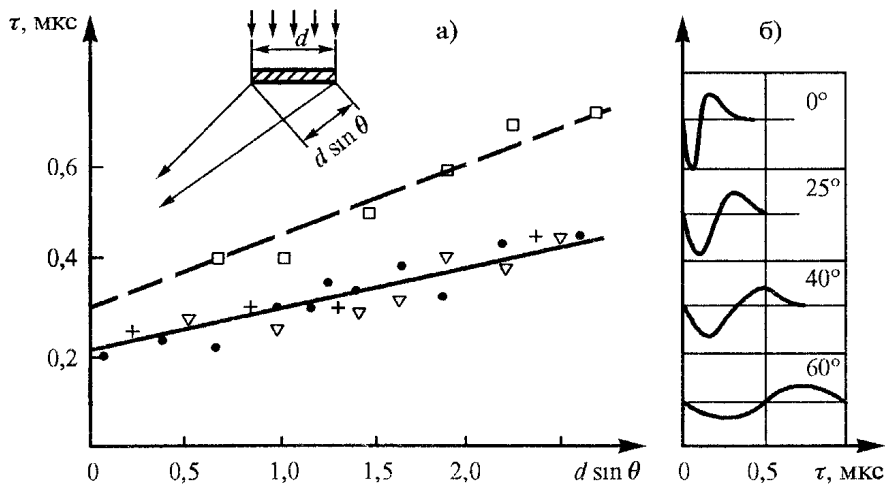


Рис. 3. Зависимость длительности акустического импульса: а) от акустической разности хода (интенсивность лазерного излучения 500 МВт/см², диаметр светового пятна 2 мм); б) от угла наблюдения (интенсивность лазерного излучения 250 МВт/см², диаметр светового пятна 2 мм);

Это значит, что длительность результирующего импульса с увеличением угла наблюдения должна линейно расти. Приведенные экспериментальные данные по зависимости длительности импульса от угла наблюдения (рис. 3б) и от акустической разности хода крайних лучей (рис. 3а)

подтверждают это. Несмотря на относительно большой разброс экспериментальных точек, угловой коэффициент по величине близок к обратной скорости продольных волн дюралюминия – 6,26 мм/мкс. График, построенный по данным, полученным с помощью ЭМАП с высокочастотным индуктором в виде проводника длиной $l = 15$ мм, располагается примерно на 150 нс выше графика для ПВТП. Это обусловлено тем, что акустические импульсы достигают края индуктора при акустическом пути $R = 40$ мм на время

$$\frac{l^2}{4RV_l} = 130 \text{ нс}$$

позже, чем до центра индуктора. Если учесть, что при диаметре ПВТП равном 8 мм также существует разность времени прихода акустических импульсов в центр и на край, равной 30 нс, то время существования импульса напряжений в диске нагрева будет

$$\tau_0 = 210 - 30 = 180 \text{ нс.}$$

Это примерно в три раза превышает длительность лазерного импульса.

Влияние размера светового пятна на амплитуду акустических импульсов от термооптического излучателя

Ультразвук возбуждался в дюралюминиевом бруске толщиной 26 мм. Размер светового пятна изменялся перемещением образца, относительно собирающей линзы с фокусным расстоянием 180 мм. Так как термооптический излучатель обладает осевой симметрией, то прием возбуждаемых им продольных волн удобно осуществлять с помощью ЭМАП с высокочастотным индуктором в виде проводящего витка. Размер витка обусловлен параметрами магнитной системы ЭМАП и равен 8 мм. В этом случае все элементы индуктора работают синфазно и обеспечивают оптимальный прием продольных волн, излученных под углом 12° . При таком малом угле излучение ультразвука от диаметрально противоположных элементов боковой поверхности диска нагрева практически компенсируются. Таким образом, регистрируемые акустические импульсы • обусловлены импульсами напряжений сжатие–растяжение у дна и повторяют их форму.

Результаты эксперимента, приведенные на рис. 4а можно объяснить, исходя из принятой модели термооптического излучателя. Основной вклад в результирующее излучение вносит дно диска нагрева. По мере удаления образца от линзы амплитуда акустического импульса должна уменьшаться, так как уменьшается диаметр светового пятна, а значит и площадь излучающей поверхности дна. Однако, при постоянной энергии лазерного излучения, его интенсивность (энергия на единицу площади) увеличивается. Таким образом, влияние интенсивности лазерного излучения оказывается более сильным и в результате амплитуда акустического импульса с уменьшением диаметра пятна увеличивается. Особенно сильный рост наблюдается при энергии лазерного излучения 0,26 Дж и диаметре пятна менее 2 мм. Поверхностная плотность тепловой мощности при этом достигает 100 МВт/см^2 . Учитывая результаты исследования, изложенные в предыдущем разделе, можно считать, что основной вклад в излучение акустических волн термооптическим источником, вносит гидродинамический механизм генерации напряжений.

Достигнув максимума при диаметре пятна 1,8 мм и плотности потока излучения 300 МВт/см^2 , амплитуда импульса несколько уменьшается вплоть до пятна минимального диаметра. Происходит это потому, что при росте плотности потока лазерного излучения более 300 МВт/см^2 , амплитуда акустического импульса возрастает, но более медленней, так как начинает проявляться влияние на излучение уменьшение площади поверхности пятна. При меньших энергиях лазерного импульса максимум амплитуды сдвигается в область меньших диаметров, вплоть до минимального.

Дальнейшее увеличение расстояния между линзой и образцом приводит к расфокусировке светового пятна и увеличения его диаметра. При этом амплитуда акустического импульса в зависимости от энергии лазерного излучения, либо сразу уменьшается, либо сначала увеличивается, достигая максимума, а затем уменьшается. Чем меньше энергия лазерного импульса, тем ближе максимум амплитуды к пятну наименьшего диаметра.

Из приведенных графиков видно, что при больших энергиях лазерного импульса при одном и том же диаметре светового пятна, амплитуда акустического импульса после фокуса линзы больше, чем до фокуса. Связано это, очевидно, с тем, что луч лазера несколько расходится, и его центральная (внутренняя) часть имеет большую, по сравнению с внешней, интенсивность. После

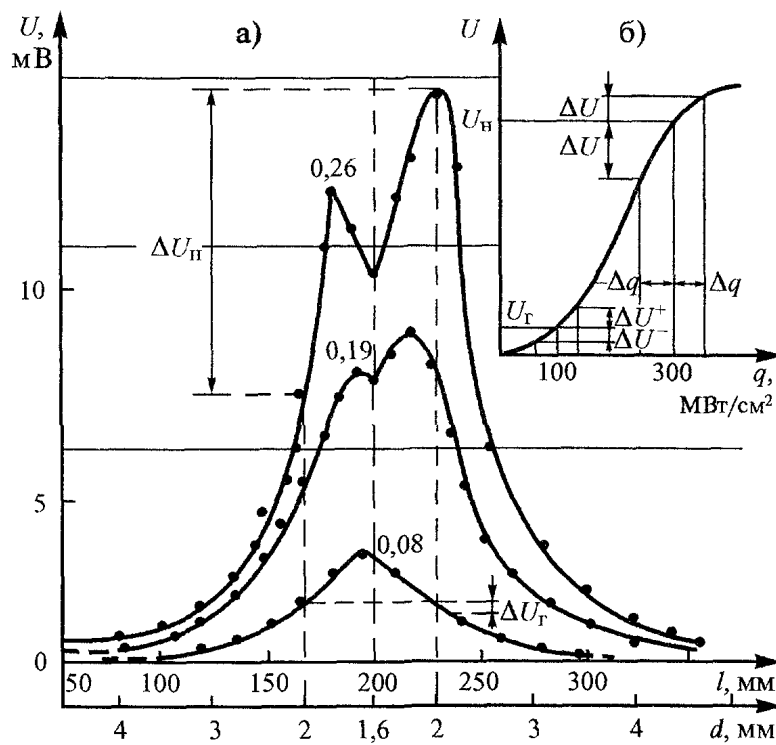


Рис. 4. Зависимость амплитуды акустических импульсов от диаметра пятна облучения или от расстояния между линзой и образцом

прохождения линзы внешняя часть луча, преломившись сильнее, сфокусируется несколько дальше за фокусом линзы, чем внутренняя, и окажется в центре луча. В результате интенсивность излучения по сечению луча выравнивается. Это хорошо подтверждается визуальными наблюдениями. Выравнивание интенсивности приводит к тому, что радиальные напряжения у боковой поверхности диска нагрева увеличиваются и, как следствие этого, увеличивается и амплитуда акустического импульса. Сказанное выше можно проиллюстрировать рис. 4б, где качественно показано изменение амплитуды акустического импульса при неравномерном и равномерном распределении интенсивности излучения.

Пусть средняя плотность потока лазерного излучения близка к плотности потока насыщения $\langle q \rangle_{\text{нас}} = 300 \text{ МВт/см}$ и амплитуда акустического импульса от светового пятна диаметром 2 мм после фокуса равна $U_{\text{нас}}$. От пятна такого же диаметра, но до фокуса средняя плотность потока такая же, но в его центре она больше на величину Δq^+ , а по его поверхностным краям меньше на величину Δq^- . В результате, это приводит к небольшому росту амплитуды акустического импульса от центра диска нагрева на величину ΔU^+ и более сильному спаду амплитуды от его периферийной части ΔU^- , т.е. $|\Delta U^+| < |\Delta U^-|$. Таким образом, амплитуда акустического импульса от всего излучателя до фокуса меньше, чем после фокуса.

При небольших энергиях лазерного излучения (до 0,08 Дж) амплитуда акустических импульсов с уменьшением размера светового пятна, растет более медленно и достигает максимума при его минимальном диаметре. Но для пятен после фокуса линзы амплитуда несколько меньше, чем для пятен такого же диаметра после фокуса. Связано это с тем, что при поверхностной плотности потока излучения около 100 МВт/см^2 происходит резкое увеличение амплитуды акустических импульсов U_r (рис. 4б). Поэтому увеличение излучения от центральной поверхности диска нагрева до фокуса линзы превышает спад излучения от его периферийной части $|\Delta U^+| > |\Delta U^-|$.

Влияние интенсивности лазерного излучения на амплитуду акустических волн

Интенсивность лазерного излучения регулировалась с помощью оптического аттенюатора. Результаты эксперимента приведены на рис. 5. Полученные зависимости можно качественно объяснить на основе принятой модели термооптического излучателя в виде диска нагрева. В точку приема приходят акустические импульсы под углом не более 12° . Так как в этом случае основной вклад в акустическое излучение вносит дно диска, то форма регистрируемого импульса колебательной скорости упругого смещения повторяет форму осевых напряжений у дна диска. В диапазоне действия термоупругого механизма генерации ультразвука амплитуда акустических импульсов растет прямо пропорционально интенсивности. Угловой коэффициент прямо пропорционален площади дна или квадрату диаметра светового пятна. При интенсивности лазерного излучения более 50 МВт/см^2 в световом пятне начинают наблюдаться яркие вспышки и амплитуда растет более быстро. Это значит, что в отдельных точках поверхности металла начинается процесс его плавления и вступает в действие более эффективный гидродинамический механизм генерации осевых напряжений. При интенсивности лазерного излучения до 300 МВт/см^2 ампли-

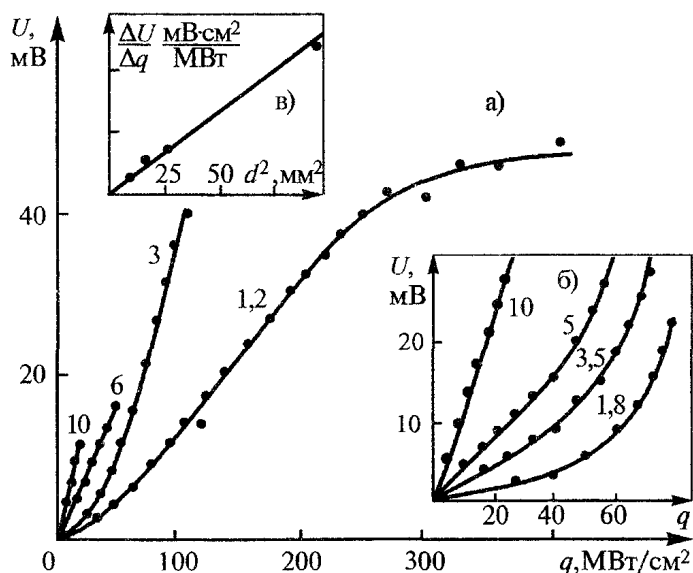


Рис. 5. Зависимости амплитуды акустических волн от интенсивности лазерного излучения, построенные в разных масштабах (а, б); зависимость амплитуды от квадрата диаметра пятна облучения (в)

туда акустического импульса быстро растет по линейному закону. При значениях интенсивности излучения лазера более 300 MW/cm² рост амплитуды замедляется. Это может быть связано с потерями энергии на испарение металла и экранированием светового пятна облаком образующейся плазмы.

Выводы

Проведено экспериментальное исследование диаграмм направленности термооптического излучателя продольных акустических волн при разных механизмах лазерной генерации упругих напряжений в металле. Найдены зависимости возбуждаемых акустических волн от угла наблюдения. Выяснено влияние размеров светового пятна на их амплитуду. Результаты экспериментов качественно объяснены с помощью принятой модели термооптического излучателя акустических волн в виде диска нагрева на поверхности металла.

Полученные результаты исследований могут успешно использоваться как в дальнейших теоретических исследованиях термоакустического преобразования в металлах с помощью лазеров, так и в прикладной лазерной ультразвуковой дефекто- и спектроскопии при конструировании и изготовлении средств неразрушающего контроля с улучшенными техническими характеристиками.

Литература

1. Волегов Ю.В., Петров Ю.В., Шушарин А.В. О механизмах возбуждения механических волн импульсами оптического квантового генератора // *Физические методы испытания материалов и веществ. Тематический сборник научных трудов.* – Челябинск. ЧГТУ, 1992. – С. 62–82.
2. Королев М.В., Карпельсон А.Е. Широкополосные ультразвуковые преобразователи. – М.: Машиностроение, 1982. – 157 с.

ОСОБЕННОСТИ УСТАНОВЛЕНИЯ КОЛЕБАНИЙ В КРУТИЛЬНОМ ВИСКОЗИМЕТРЕ

А.Е. Коренченко, В.П. Бескачко

Численными методами рассматривается модельная задача о течениях вязкой жидкости, заключенной внутри колеблющегося круглого цилиндра (вискозиметра). Для решения используется метод конечных разностей. Получены картины течений в вискозиметре, поля скоростей и давления. Исследуются частотные характеристики колебаний вискозиметра.

В связи с интерпретацией опытных данных, обнаруживающих аномалии в поведении политерм вязкости металлических расплавов, в работах [1, 2] были предприняты попытки учесть влияние на результаты прямых измерений факторов, отсутствующих в стандартной вискозиметрической теории [3, 4], но обычно присутствующих в экспериментах. В настоящей работе мы намерены продолжить анализ теоретических основ метода крутильных колебаний, обратив внимание на те предположения, на которых базируются все существующие ныне расчетные схемы. Именно, всегда считается, что:

а) колебания совершаются столь долго, что все переходные процессы, связанные с выбором начальных условий (выбором способа возбуждения колебаний) уже завершились, и система находится в режиме установившихся затухающих колебаний;

б) единственной существенной компонентой скорости жидкости является азимутальная. Первое предположение избавляет нас от рассмотрения эволюции системы во времени, второе (вместе с соображениями симметрии) – от нелинейностей в уравнениях Навье–Стокса, а оба они вместе позволяют найти решение вискозиметрической задачи в аналитическом виде, который и используется на практике во всех схемах расчета вязкости.

Выполнение первого условия в экспериментах обеспечить тем легче, чем меньше декремент затухания системы, то есть чем меньше отношение I'/I , где I' и I моменты инерции «замороженной» жидкости и подвесной системы соответственно. Справедливость второго условия зависит от того, насколько мала угловая амплитуда Φ колебаний подвеса. Ясно, что амплитуда азимутальной компоненты скорости жидкости в цилиндре также имеет порядок Φ , а амплитуды двух оставшихся компонент имеют порядок Φ^2 , как показано в [5]. Таким образом, каждое из высказанных предположений справедливо лишь асимптотически (при $t \rightarrow \infty$) – в пределе исчезающе малых колебаний.

Стремление к последнему пределу на практике быстро ограничивается падением отношения «сигнал/шум» при измерении необходимых для расчета вязкости параметров колебаний – периода и декремента затухания. Поэтому для корректной интерпретации опытных данных необходимы количественные оценки длительности переходного процесса и роли «вторичных» течений, обусловленных радиальной и осевой компонентами скорости жидкости. Такие оценки позволили бы ответить на следующие вопросы, касающиеся методики выполнения экспериментов.

1. Спустя какое время τ_1 после старта подвесной системы, осуществляемого тем или иным способом, устанавливается режим затухающих колебаний, а данные прямых измерений уже не содержат методических погрешностей, связанных с наблюдением переходного процесса?

2. Начиная с какого момента времени τ_2 амплитуда колебаний становится достаточно малой для того, чтобы влиянием вторичных течений на наблюдаемые параметры можно было пренебречь, и как выбирать начальную амплитуду колебаний, чтобы τ_2 было по возможности меньше?

3. Каким способом можно обнаружить присутствие указанных двух погрешностей по опытным данным, полученным с помощью той или иной схемы регистрации колебаний?

Насколько нам известно, попыток систематического исследования этих вопросов еще не предпринималось. В отсутствие ответов на них остается только надеяться, что «конкретно в наших экспериментах» условия, требуемые для применения стандартных вискозиметрических ме-

тодик, выполняются. Целью настоящей работы является создание математической модели, описывающей совместное движение подвесной системы вискозиметра и содержащейся в нем жидкости, на основе точных уравнений движения, и не использующей приближения, указанные выше. В силу последнего обстоятельства такая модель является численной, поскольку надежд на аналитическое решение нет.

Математическая формулировка задачи. Пусть прямой круговой цилиндр с внутренним радиусом R и моментом инерции I подвешен вдоль своей геометрической оси на упругой нити, имеющей крутильную жесткость κ (рис. 1). Цилиндр частично или до конца заполнен исследуемой жидкостью, которая полагается ньютоновской и обладает плотностью ρ и кинематической вязкостью ν . Высота столба жидкости $H \leq H_c$, где H_c – высота внутренней части боковой поверхности цилиндра, так что момент инерции «замороженной» жидкости есть $I' = MR^2/2 = \rho\pi R^2 H/2$. Рассматриваются два способа возбуждения крутильных колебаний:

1) при $t < 0$ цилиндр вместе с жидкостью покоится в положении равновесия, а в момент $t = 0$ ему сообщается угловая скорость ω_0 ;

2) при $t < 0$ цилиндр с жидкостью удерживается в состоянии покоя в положении, повернутом на угол ϕ_0 относительно положения равновесия.

При $t > 0$ цилиндр предоставляется самому себе, совершая крутильные колебания под действием моментов упругих ($M_e = -\kappa\phi$) и вязких (M_f) сил, приложенных к нему со стороны нити и заполняющей его жидкости. Уравнение его движения имеет вид

$$I \frac{d\omega}{dt} = -\kappa\phi + M_f. \tag{1}$$

В соответствии со сказанным выше начальными условиями для уравнения (1) являются:

$$\omega(0) = \omega_0, \quad \phi(0) = 0, \tag{2}$$

или
$$\omega(0) = 0, \quad \phi(0) = \phi_0. \tag{3}$$

Момент вязких сил относительно оси вращения определяется полем скорости жидкости на ее границе с твердым телом:

$$M_f = M_1 + M_2, \tag{4}$$

где моменты M_1 и M_2 обусловлены трением жидкости о боковую поверхность цилиндра и его торцы соответственно:

$$M_1 = 2\pi\eta R^2 \int_0^H \left(\frac{\partial V_\phi}{\partial r} - \frac{V_\phi}{r} \right) \Big|_{r=R} dz, \tag{5}$$

$$M_2 = 2\pi\eta \int_0^R \left(\frac{\partial V_\phi}{\partial z} \right) \Big|_{z=0} r^2 dr - 2\pi\eta \int_0^R \left(\frac{\partial V_\phi}{\partial z} \right) \Big|_{z=H} r^2 dr, \tag{6}$$

Здесь $\eta = \nu\rho$ – динамическая вязкость жидкости; $V_\phi = V_\phi(r, \phi, z, t)$ – азимутальная компонента поля ее скорости $\mathbf{V} = (V_r, V_\phi, V_z)$ (мы ввели цилиндрическую систему координат, ось которой совпадает с осью цилиндра, а начало отсчета располагается на его дне). В случае, если жидкость не смачивает верхнюю крышку цилиндра, последнее слагаемое в (6) отсутствует. Движение цилиндра возбуждает течение заполняющей его жидкости, подчиняющееся уравнениям:

$$\frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V}\nabla)\mathbf{V} = -\frac{1}{\rho}\nabla P + \nu\Delta\mathbf{V} + \rho\mathbf{g}, \tag{7}$$

$$(\nabla, \mathbf{V}) = 0. \tag{8}$$

Уравнения (7) есть уравнения баланса импульса (уравнения Навье–Стокса), а (8) – условие несжимаемости жидкости. В этих уравнениях P – поле давления; \mathbf{g} – ускорение свободного падения. Граничными условиями для (7)–(8) служат:

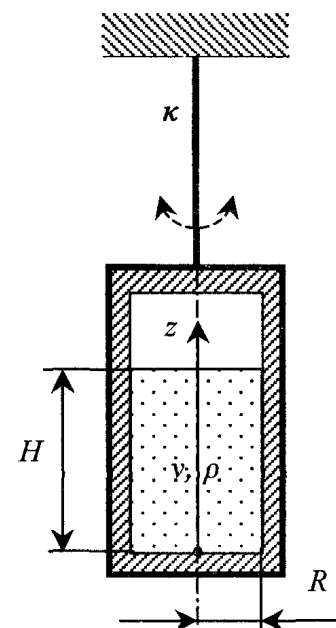


Рис. 1. Схема крутильного вискозиметра

$$V_\varphi(R, z, t) = \omega R, \quad (9)$$

$$V_\varphi(r, 0, t) = \omega r, \quad V_\varphi(r, H_c, t) = \omega r, \quad (10)$$

$$V_r(r, 0, t) = V_z(r, 0, t) = V_r(r, H_c, t) = V_z(r, H_c, t) = V_r(R, z, t) = V_z(R, z, t) = 0. \quad (11)$$

Они означают условия прилипания жидкости ко всем твердым границам. Последнее равенство в (10) имеет место, если жидкость целиком заполняет цилиндр ($H = H_c$). В противном случае ее верхняя граница свободна. В каждом из рассматриваемых случаев (2), (3) возбуждения колебаний жидкость в начальный момент времени покоится

$$\mathbf{V}(\mathbf{r}, 0) = 0. \quad (12)$$

Соотношениями (1)–(3), (7)–(12) полностью определена сопряженная краевая задача о движении твердого цилиндра с содержащейся в нем жидкостью. В приведенной постановке она является весьма громоздкой даже и для численного решения. Поэтому на данном этапе исследования рассмотрим следующие возможности ее упрощения. Во-первых, мы ограничимся изучением только осесимметричных режимов течения, когда ни одна из интересующих нас полевых переменных не зависит от азимутального угла φ . Во-вторых, поскольку нас интересует случай малых колебаний, мы будем считать их малыми настолько, что можно пренебречь деформацией свободной поверхности жидкости и считать ее плоской в любой момент времени. Заодно мы считаем, что радиус цилиндра достаточно велик для того, чтобы можно было пренебречь деформацией границы также и вследствие капиллярных эффектов. Первое предположение позволяет опустить в уравнениях (7)–(8) все слагаемые, содержащие производные по φ , получив в результате

$$\begin{aligned} \frac{\partial V_r}{\partial t} + V_r \frac{\partial V_r}{\partial r} + V_z \frac{\partial V_r}{\partial z} - \frac{V_\varphi^2}{r} &= -\frac{1}{\rho} \frac{\partial P}{\partial r} + \nu \left[\frac{\partial^2 V_r}{\partial r^2} + \frac{1}{r} \frac{\partial V_r}{\partial r} - \frac{V_r}{r^2} + \frac{\partial^2 V_r}{\partial z^2} \right], \\ \frac{\partial V_\varphi}{\partial t} + V_r \frac{\partial V_\varphi}{\partial r} + V_z \frac{\partial V_\varphi}{\partial z} + \frac{V_r V_\varphi}{r} &= \nu \left[\frac{\partial^2 V_\varphi}{\partial r^2} + \frac{1}{r} \frac{\partial V_\varphi}{\partial r} - \frac{V_\varphi}{r^2} + \frac{\partial^2 V_\varphi}{\partial z^2} \right], \\ \frac{\partial V_z}{\partial t} + V_r \frac{\partial V_z}{\partial r} + V_z \frac{\partial V_z}{\partial z} &= -\frac{1}{\rho} \frac{\partial P}{\partial z} + \nu \left[\frac{\partial^2 V_z}{\partial r^2} + \frac{1}{r} \frac{\partial V_z}{\partial r} + \frac{\partial^2 V_z}{\partial z^2} \right] + \rho g, \\ \frac{1}{r} \frac{\partial(rV_r)}{\partial r} + \frac{\partial V_z}{\partial z} &= 0. \end{aligned}$$

Второе предположение позволяет записать условие отсутствия сдвиговых напряжений на границе в виде

$$\left. \frac{\partial V_\varphi}{\partial z} \right|_{z=H} = 0, \quad V_z|_{z=H} = 0.$$

Если ввести теперь безразмерные переменные и параметры так, что все расстояния отнесены к внутреннему радиусу R цилиндра, скорость – к ν/R , давление – к $\rho \nu^2/R^2$, время – к R^2/ν , движение цилиндра и заполняющей его жидкости описывается следующей системой уравнений:

$$\begin{aligned} \frac{\partial u}{\partial \tau} + v \frac{\partial u}{\partial y} + u \frac{\partial u}{\partial x} &= -\frac{\partial p}{\partial x} + \frac{\partial^2 u}{\partial y^2} + \frac{1}{y} \frac{\partial u}{\partial y} + \frac{\partial^2 u}{\partial x^2}, \\ \frac{\partial v}{\partial \tau} + v \frac{\partial v}{\partial y} + u \frac{\partial v}{\partial x} - \frac{w^2}{y} &= -\frac{\partial p}{\partial y} + \frac{\partial^2 v}{\partial y^2} + \frac{1}{y} \frac{\partial v}{\partial y} - \frac{v}{y^2} + \frac{\partial^2 v}{\partial x^2}, \\ \frac{\partial w}{\partial \tau} + v \frac{\partial w}{\partial y} + u \frac{\partial w}{\partial x} + \frac{vw}{r} &= \frac{\partial^2 w}{\partial y^2} + \frac{1}{y} \frac{\partial w}{\partial y} - \frac{w}{y^2} + \frac{\partial^2 w}{\partial x^2}, \\ \frac{1}{y} \frac{\partial(yv)}{\partial y} + \frac{\partial u}{\partial x} &= 0, \\ \frac{df}{d\tau} &= -G\varphi + m_f. \end{aligned} \quad (13)$$

Введенные здесь безразмерные величины означают: (x, y) , $x = \frac{z}{R}$, $y = \frac{r}{R}$ – компоненты радиус-вектора; (u, v, w) – компоненты вектора скорости; p – возмущение давления относительно $\rho g x$; f – угловая скорость цилиндра; $G = \frac{R^4 \kappa}{\nu^2 I}$ – коэффициент упругости,

$$m_j = \frac{2\pi\rho R^5}{I} \left(\int_0^{H/R} \left(\frac{\partial w}{\partial y} - \frac{w}{y} \right) \Big|_{y=1} dx + \int_0^1 \left(\frac{\partial w}{\partial x} \right) \Big|_{x=0} y^2 dy - \int_0^1 \left(\frac{\partial w}{\partial x} \right) \Big|_{x=H/R} y^2 dy \right)$$

– момент вязких сил. Граничные условия запишутся в виде

$$w(1, x, \tau) = f,$$

$$w(y, 0, \tau) = f \cdot y, \quad w(y, H_c/R, \tau) = f \cdot y, \tag{14}$$

$$v(y, 0, \tau) = u(y, 0, \tau) = v(y, H_c/R, \tau) = u(y, H_c/R, \tau) = v(1, x, \tau) = u(1, x, \tau) = 0.$$

Кроме того, на оси сосуда выполняются условия $\frac{\partial u}{\partial y} = \frac{\partial p}{\partial y} = 0$, $v = 0$, обусловленные осевой симметрией течения. Задача заключается в отыскании закона движения цилиндра $f = f(\tau)$ и заполняющей его жидкости.

Численное решение. Численное решение системы (13) с граничными условиями (14) находилось методом сеток. Использовались равномерные пространственные сетки с максимальной размерностью 64×64 в осевом и радиальном направлениях. Дискретизация гидродинамических уравнений производилась по схеме центральных разностей с точностью $(\Delta x)^2$ по пространственным переменным. Линеаризация получаемых в результате нелинейных разностных уравнений проводилась методом Ньютона, а решение линеаризованных систем уравнений – методом исключения Гаусса. Консервативные свойства решения контролировались на каждом временном шаге, соответствующие погрешности оказались находящимися в пределах схемной точности. Расчеты проводились для цилиндра с внутренним радиусом 1–3 см и высотой заполнения его жидкостью 1–3 см, так что диапазон изменения параметра задачи H/R был от $1/3$ до 1. Физические параметры жидкости отвечали воде.

Результаты и обсуждение. В настоящей работе приводятся результаты численных экспериментов для случая, когда жидкость целиком заполняет цилиндр. На рис. 2 показаны линии тока «вторичных» течений, обусловленных радиальной и осевой компонентами поля скорости, для различных фаз колебаний цилиндра. Рис. 2а соответствует максимальной, а рис. 2б – нулевой угловой скорости цилиндра. В первом случае образуется вихрь, в котором жидкость поднимается вверх около боковой поверхности цилиндра и опускается вниз вблизи его оси. Во втором случае около дна цилиндра возникает второй вихрь с противоположным направлением циркуляции. Подобные вихри существуют в любой фазе движения цилиндра. Вблизи верхнего торца сосуда, где скорости велики, формируется зона низкого давления, куда выталкиваются линии тока (на рисунке видна область сгущения линий тока). Основной вихрь не всегда занимает всю область

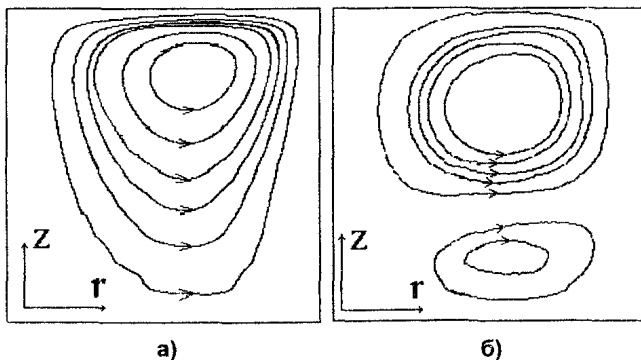


Рис. 2. Линии тока вторичных течений:
а) угловая скорость цилиндра максимальна;
б) угловая скорость цилиндра равна нулю

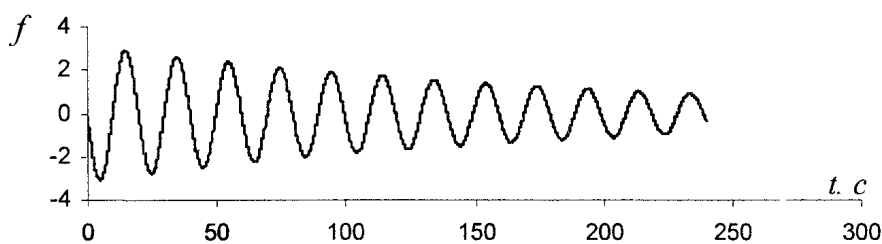
внутри цилиндра, в нижних областях (вблизи дна цилиндра) в моменты, когда скорость цилиндра близка к нулю, возникают дополнительные рециркуляционные зоны (рис. 2б). На рис. 3а и 3б показаны законы изменения угловой скорости цилиндра f , а также осевой $u(\tau)$ и радиальной $v(\tau)$ скоростей жидкости вблизи боковой стенки для случая, когда цилиндр стартует из состояния покоя, будучи предварительно повернут на некоторый начальный угол, а затухание велико. Особенностью движения жидкости в цилиндре является то, что частота колебаний осевой и радиальной скоростей в два раза выше частоты основных колебаний

(рис. 3а и 3б). Этот факт легко объяснить присутствием в уравнении Навье–Стокса для радиальной скорости слагаемого, описывающего действие центробежной силы $\frac{w^2}{y}$. Ясно, что этот член

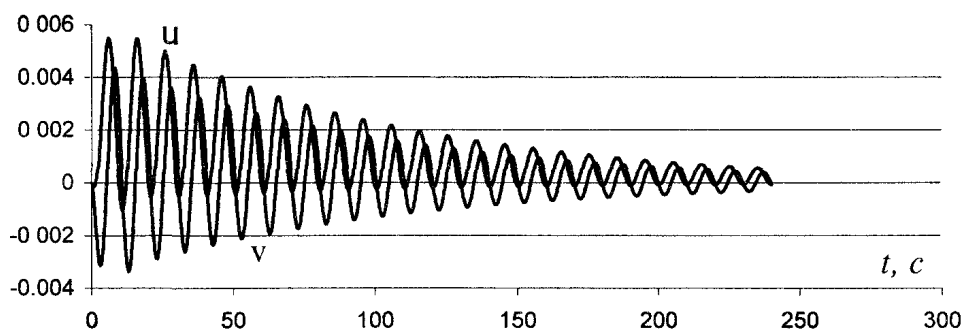
имеет смысл источника для уравнения радиальной скорости и имеет двойную, по сравнению с основной, частоту колебаний. В свою очередь, колебания осевой и радиальной скоростей вызывают появление высших гармоник в спектре азимутальной скорости, благодаря членам вида

$-w \cdot \frac{\partial u}{\partial x}$ и $-w \cdot \frac{\partial v}{\partial x}$. Визуально трудно заметить существование какого-либо переходного процесса в колебаниях цилиндра (рис. 3б). Однако анализ спектра этого движения указывает, что наряду с основной частотой присутствуют и высшие гармоники, обусловленные затуханием колебаний,

влиянием начальных условий, а также нелинейным взаимодействием между различными колебаниями. На рис. 4 изображены зависимости отношения амплитуд второй и первой (основной) гармоник колебаний для двух случаев возбуждения колебаний. В процессе установления колебаний

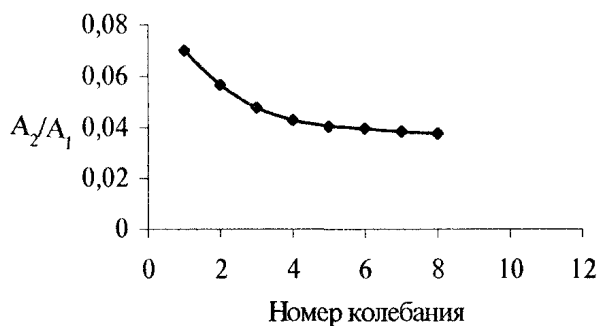


а)

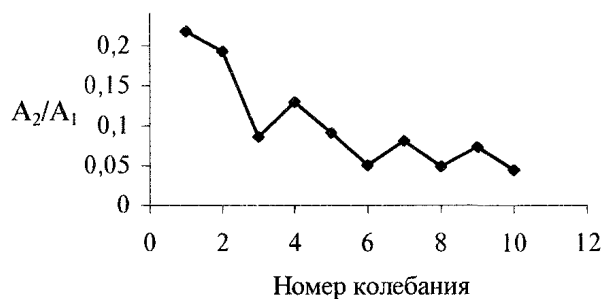


б)

Рис. 3. а) закон $f(t)$ движения цилиндра при старте из состояния $f(0) = 0$, $\phi(0) = \phi_0$; б) законы $u(t)$ и $v(t)$ для скоростей жидкости во внутренней точке при начальных условиях, соответствующих рис. 3а



а)



б)

Рис. 4. Относительное убывание амплитуды гармоник для двух случаев возбуждения колебаний

отношение амплитуд второй и первой, а также нулевой и первой гармоник изменяется, приближаясь к некоторому постоянному значению (рис. 4). Отметим, что установление колебаний в случае возбуждения их толчком (рис. 4б), имеет колебательный характер и длится дольше, по сравнению со случаем, показанным на рис. 4а. Приблизительно с десятого колебания в обоих случаях можно считать, что колебания установились, и проводить измерения затухания.

Работа выполнена при поддержке РФФИ, грант 01-01-96424 Урал.

Литература

1. Бескачко В.П., Вяткин Г.П., Писарев Н.М., Щека А.И. Влияние поверхностных пленок на результаты измерения вязкости по методу Швидковского. I. Теория // *Расплавы*. – 1990. – № 6. – С. 3–8; II. Численные эксперименты. – Там же. – С. 9–16.
2. Бескачко В.П., Вяткин Г.П., Писарев Н.М., Хисматулин М.Б. Теория крутильного вискозиметра, помещенного в осевое магнитное поле // *Магнитная гидродинамика*. – 1992. – № 2. – С. 65–70.
3. Швидковский Е.Г. Некоторые вопросы вязкости расплавленных металлов. – М.: ГИТТЛ, 1955. – 206 с.
4. Kestin J., Newell G.F. Theory of oscillating type viscometers: the oscillating cap. Part I. // *ZAMP*. – 1957. – V. VIII. – P. 433–449.
5. Grouvel J.M., Kestin J. Working equations for the oscillating-cup viscometer // *Appl. Sci. Res.* – 1978. – V. 34. – P. 427–443.

ТЕЧЕНИЕ, ВОЗБУЖДАЕМОЕ ДИСКОМ, ВРАЩАЮЩИМСЯ НА ПОВЕРХНОСТИ ЖИДКОСТИ

В.П. Бескачко, А.Е. Коренченко

Рассматривается задача о равномерном вращении толстого диска, погруженного в жидкость на некоторую глубину, получены картины возникающих в жидкости течений. Изучено затухающее движение диска после прекращения действия сил, поддерживающих его вращение. Расчеты показали, что при затухающем движении диска в ньютоновской жидкости не возникают возвратные движения, наблюдавшиеся в [1].

Непосредственной мотивацией для выполнения настоящей работы послужили результаты экспериментов, выполненных в [1] и указывающих на возможность присутствия упругих свойств в жидкостях, обычно считающихся ньютоновскими. В этих экспериментах поплавков в виде диска помещался на поверхность исследуемой жидкости и приводился во вращение вокруг собственной оси. По достижении некоторой (небольшой) угловой скорости диск предоставлялся самому себе, и наблюдали процесс затухания его вращения вследствие вязкого трения о жидкость. Обнаружилось, однако, что затухание носит осциллирующий характер, неожиданный для ньютоновской жидкости. Это дало повод авторам [1] предположить, что в исследованных ими жидкостях наряду с вязкими присутствуют также и упругие свойства. Не исключая подобной возможности, мы хотели бы обратить внимание также и на то, что характеристики течения, возбуждаемого вращающимся диском в жидкости, сложным образом зависят от:

- 1) «начальных условий» – точнее от истории раскрутки диска от состояния покоя до прекращения действия внешнего вращающего момента;
- 2) от формы и размеров твердых поверхностей, ограничивающих жидкость;
- 3) от механических параметров самого диска (его момента инерции, например).

Учитывая нестационарность, трехмерность и нелинейность задачи, даже в ньютоновском случае трудно предвидеть характер течения, например, априори исключить возможность возвратных течений. Последние могли бы быть причиной наблюдаемого движения диска и в отсутствие упругих свойств жидкости. В связи с этим в настоящей работе предпринята попытка численного исследования некоторых упрощенных вариантов задачи.

Математическая формулировка задачи. Пусть диск радиуса R_d (рис. 1) плавает на поверхности вязкой несжимаемой жидкости (с плотностью ρ и кинематической вязкостью ν), погружившись в нее на некоторую глубину. Жидкость содержится в круговом цилиндре радиуса R_c и заполняет его до высоты H_c . Геометрические оси цилиндра и диска совпадают. Рассматриваются следующие случаи относительного движения диска и цилиндра.

1. В момент $t = 0$ диск начинает вращение с угловой скоростью ω_d , цилиндр неподвижен. Изучается поле скорости жидкости до момента установления стационарного вращения.

2. Цилиндр с жидкостью, покоившийся до момента $t = 0$, начинает вращаться с угловой скоростью ω_c . Диск остается неподвижным все время движения. Исследуется стационарное распределение скоростей в сосуде.

3. До момента $t = 0$ вся система покоится. При $t = 0$ диск начинает вращаться с угловой скоростью ω_d . После достижения стационарного состояния приводящий диск в движение момент исчезает, и диск продолжает движение по инерции. Исследуется характер этого затухающего движения.

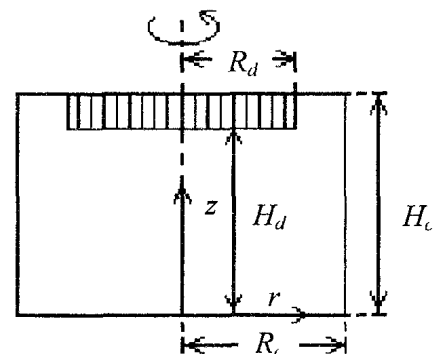


Рис. 1. Схема установки

В безразмерных переменных, когда все расстояния отнесены к радиусу R_c цилиндра, скорость – к ν/R_c (ν – кинематическая вязкость жидкости), возмущение давления сверх $\rho g z$ – к $\rho \nu^2/R_c^2$ (ρ – плотность жидкости), время – к R_c^2/ν , движение диска и окружающей его жидкости описывается следующей системой уравнений:

$$\frac{\partial \mathbf{u}}{\partial t} = \Delta \mathbf{u} - \mathbf{u} \cdot \nabla \mathbf{u} - \nabla p, \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2)$$

$$\frac{d\omega_d}{dt} = M_{fr}. \quad (3)$$

Здесь мы ввели цилиндрическую систему координат (r, θ, z) , начало отсчета которой поместили на дне сосуда (рис. 1), $\mathbf{u} = (u_z, u_r, u_\theta)$ и p – поля безразмерной скорости и безразмерного возмущения давления в жидкости, ω_d – угловая скорость диска, M_{fr} – безразмерный момент вязких сил трения, приложенных к диску со стороны жидкости, учитывающий трение о боковую поверхность и нижний торец диска. Выражение для момента учитывает направление внешней нормали к поверхности и записывается в виде [2]:

$$M_{fr} = M_1 + M_2 = T \left[- \int_{\text{по нижней поверхности диска}} r^2 \frac{\partial u_\theta}{\partial z} dr + \int_{\text{по боковой поверхности диска}} \left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right) dz \right], \quad T = \frac{2\pi\rho R^5}{I},$$

здесь I – момент инерции диска.

Уравнение (1) – уравнение баланса импульса в вязкой несжимаемой жидкости (уравнение Навье–Стокса), (2) – условие ее несжимаемости, (3) – уравнение движения диска по инерции. Граничные условия состоят в прилипании жидкости ко всем твердым границам:

$$z = 0, \quad 0 < r < R_c: \quad u_z = 0, \quad u_r = 0, \quad u_\theta = \omega_c \cdot r;$$

$$z = H_d, \quad 0 < r < R_d: \quad u_z = 0, \quad u_r = 0, \quad u_\theta = \omega_d r;$$

$$0 < z < H_c, \quad r = R_c: \quad u_z = 0, \quad u_r = 0, \quad u_\theta = \omega_c R_c;$$

$$H_d < z < H_c, \quad r = R_d: \quad u_z = 0, \quad u_r = 0, \quad u_\theta = \omega_d R_d;$$

и в отсутствии сдвиговых напряжений на свободной границе:

$$z = H_c, \quad R_d < r < R_c: \quad u_z = 0, \quad \frac{\partial u_r}{\partial z} = 0, \quad \frac{\partial u_\theta}{\partial z} = 0.$$

В условиях осевой симметрии при $r = 0$ будут иметь место условия $\frac{\partial f}{\partial r} = 0, u_r = 0$ для всех функций $f(r, \theta, z)$.

Численные методы решения. Численное решение проводилось методом конечных разностей. Использовались равномерные пространственные сетки с максимальной размерностью 45×45 в осевом и радиальном направлениях. Дискретизация гидродинамических уравнений производилась по схеме центральных разностей с точностью $(\Delta x)^2$ по пространственным переменным. Линеаризация получаемых в результате нелинейных разностных уравнений проводилась методом Ньютона, а решение линеаризованных систем алгебраических уравнений – методом исключения Гаусса [3, 4]. Консервативные свойства решения контролировались на каждом временном шаге, и соответствующие погрешности оказались находящимися в пределах схемной точности ($\sim 1\%$). Расчеты проводились для цилиндрического контейнера с внутренним радиусом 3 см, параметр задачи H_c/R_c принимал значения в пределах между 1 и 3. Радиус диска варьировался в пределах 1...3 см. Физические параметры жидкости отвечали воде.

Результаты и обсуждение. Через некоторое время после начала движения во всех рассмотренных случаях наступало стационарное состояние. Линии тока стационарных движений для случаев 1 и 2 изображены на рис. 2а, б. Твердотельное вращение в таких обстоятельствах не на-

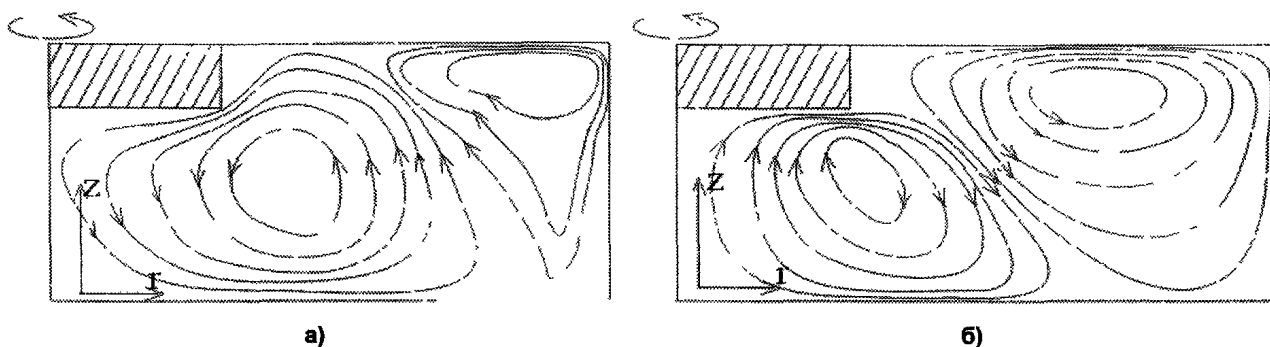


Рис. 2. Линии тока при вращательном движении: а) вращается контейнер; б) вращается диск

блюдалось даже при самых малых угловых скоростях. Движение жидкости в этих случаях имело сходные черты. Для обоих случаев характерно образование двух вихрей: первый возникает вблизи свободной поверхности, центр второго находится под диском. Однако имеется существенное различие в направлениях вращения жидкости, связанное с тем, что в случае вращения диска движение зарождается вблизи диска под его поверхностью, радиальные давления под диском обусловлены центробежной силой и направлены в сторону возрастания радиуса. Вихрь, возникающий вблизи свободной поверхности, имеет вторичное происхождение и появляется в результате распространения первого вихря внутри жидкости под действием центробежных и вязких сил. В случае вращения контейнера при покоящемся диске элементы жидкости, расположенные вблизи свободной поверхности, имеют меньшую, по сравнению со случаем твердотельного вращения, скорость и поэтому под действием центробежной силы приобретают некоторый импульс в направлении возрастания радиуса. Вихрь, сформированный вблизи свободной поверхности, является первичным, в результате его распространения в жидкости появляется вторичный вихрь, в котором жидкость затекает под поверхность диска. Если диск вращается в покоящемся контейнере и после достижения стационарного состояния вращение диска не поддерживается извне, то под действием сил трения со стороны жидкости диск постепенно останавливается. На рис. 3 изображена зависимость угловой скорости такого затухающего вращательного движения диска от времени. Уменьшение угловой скорости описывается экспоненциальной зависимостью, по коэффициенту затухания этой кривой можно судить о коэффициенте кинематической вязкости жидкости. В затухающем движении диска в жидкости, подчиненной ньютоновскому закону вязкости не наблюдалось возвратных движений, из чего можно сделать вывод, что адекватное описание эксперимента [1] возможно лишь с помощью моделей жидкости, включающих возможность релаксации напряжений.

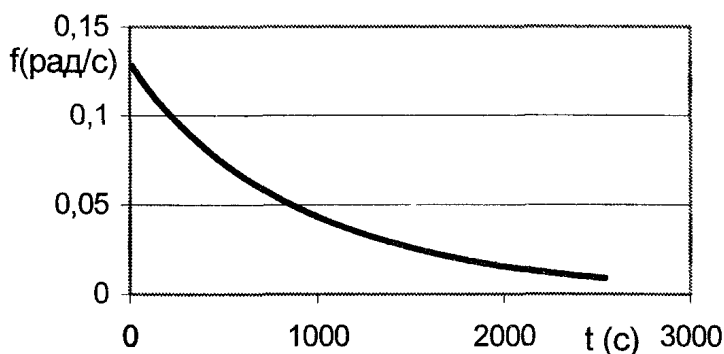


Рис. 3. Временная зависимость угловой скорости при затухающем движении диска

Работа выполнена при поддержке РФФИ, грант 01-01-96424 Урал.

Литература

1. Apakashev R.A., Pavlov V.V. Determination of the shear strength and modulus of water at low flow velocities // *J. Fluid Dynamics*. – 1997. – V. 32, № 1. – P. 1.
2. Ландау Л.Д., Лившиц Е.М. Механика сплошных сред. – М.: Наука, 1988. – 518 с.
3. Ши Д. Численные методы в задачах теплообмена. – М.: Мир, 1988. – 537 с.
4. T. Inamuro, A. Yamaguchi, F. Ogino. Fluid flow in a rotating cylindrical container with a rotating disk at the fluid surface // *Fluid Dynamics Research*. – 21 (1997). – P. 417.

ИССЛЕДОВАНИЕ ПОВЕРХНОСТНЫХ АКУСТИЧЕСКИХ ВОЛН, ВОЗБУЖДАЕМЫХ ЛАЗЕРНЫМ ИМПУЛЬСОМ В ФЕРРОМАГНЕТИКАХ ПРИ ТЕМПЕРАТУРЕ МАГНИТНОГО ФАЗОВОГО ПЕРЕХОДА

С.Ю. Гуревич, Ю.В. Петров, Е.В. Голубев, К.Б. Хабиров

Проведено экспериментальное исследование зависимости параметров акустического поля поверхностных волн при лазерном возбуждении от температуры в ферромагнитном сплаве 32НКД (инвар). Магнитный фазовый переход в ферромагнетике приводит к увеличению амплитуды импульса ПАВ и изменению в спектре. Также зафиксирован рост скорости распространения волн Рэлея с температурой.

I. Введение

Возбуждение ультразвука в ферромагнитном металле короткими лазерными импульсами сопровождается локальным повышением его температуры. Если в ферромагнетике происходит магнитный фазовый переход (точка Кюри), то это может существенно повлиять на некоторые физические параметры металла.

Как было показано в работах [1, 2], изменение свойств ферромагнетиков при магнитном фазовом переходе существенно влияет на амплитуду, спектр и длительность импульсов возбуждаемых объемных акустических волн. В данной работе исследуется влияние магнитного фазового перехода на параметры высокотемпературного преобразования импульсного лазерного излучения в акустическое поле поверхностных волн (ПАВ).

II. Эксперимент

Для решения поставленной задачи была использована экспериментальная установка, состоящая из NdYAG лазера ЛТИ-403 (1), оптического аттенюатора (2), измерителя оптической мощности (3), набора диафрагм с отверстиями от 0,5 до 4 мм (4), фокусирующей линзы (5), исследуемого образца из ферромагнитного материала (6), измерителя температуры (7), нагревателя (8), приемника ПАВ (9), широкополосных усилителя регистрируемых сигналов (10) и осциллографа (11), видеокамеры (12) (рис. 1).

Лазер, работающий в режиме активной модуляции добротности, излучает световой импульс длительностью 20–30 нс с частотой следования 12,5–100 Гц. Длина волны излучения составляет 1,06 мкм, максимальная энергия импульса 0,3 Дж. Исследуемый образец, моделирующий свойства ферромагнетиков при их нагревании, представляет собой прямоугольную пластину 270×60×10 мм, изготовленную из ковanej заготовки железоникелевого сплава 32НКД инварного состава. Температура Кюри, при которой происходит магнитный фазовый переход у такого сплава, относительно невелика и составляет 210 °С. В эксперименте использовался электронагреватель мощностью 300 Вт. Температура измерялась с помощью стандартной хромель-алюмелевой термопары и прибора Digi-Temper, model 3527 TDS-RUGA с цифровой индикацией.

Для исследования влияния магнитного фазового перехода на процесс возбуждения ПАВ, необходимо, чтобы температура ферромагнетика, обусловленная

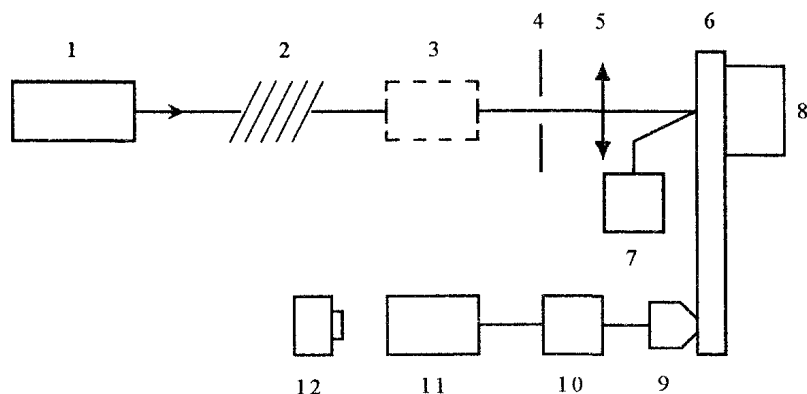


Рис. 1. Схема экспериментальной установки

энергией лазерного излучения, не превышала температуру Кюри. Это возможно только в том случае, если возбуждение ПАВ происходит за счет термоупругого эффекта. С целью нахождения пороговых значений энергии лазерного излучения, определяющих действие термоупругого механизма генерации ПАВ, был проведен ряд экспериментов.

Исследования проводились по методике, изложенной в работе [3]. В качестве приемника ПАВ использовался усовершенствованный ЭМАП поверхностных волн [4]. Применение НЖБ-магнитов и более совершенной схемы предусилителя, позволили значительно увеличить чувствительность такого преобразователя. Располагался ЭМАП на освещаемой поверхности образца на расстоянии 180 мм от места возбуждения ПАВ. Оптическим attenuатором и измерителем мощности задавались и измерялись значения средней мощности лазерного излучения в достаточно широком диапазоне. Размер лазерного пятна и равномерность его освещения осуществлялись с помощью диафрагм и линзы. Регистрируемые на экране осциллографа электрические импульсы, пропорциональные импульсам скорости упругих смещений в ПАВ, снимались на видеокамеру, вводились в компьютер и обрабатывались.

На рис. 2 приведены результаты исследований для лазерного пятна диаметром 2 мм. Абсцисса точки пересечения касательных к начальному и наиболее крутому участкам кривой зависимости амплитуды электрического импульса на выходе приемного тракта U от натурального логарифма средней поверхностной плотности мощности лазерного излучения $\langle q \rangle$, определяет конец действия термоупругого и начало вступления гидродинамического механизмов генерации ПАВ. Подобным образом можно установить и начало действия испарительного механизма.

При исследовании температурной зависимости параметров импульсов смещений в ПАВ лазерный луч через диафрагму проецируется на поверхность образца, с противоположной стороны которого находится плоский электронагреватель. Вблизи места возбуждения ПАВ закреплялась термopара. В качестве приемников ПАВ использовались клиновидный поверхностно-возбуждаемый толстый пьезопреобразователь (ПВТП) и ЭМАП поверхностных волн.

Выбор клиновидного ПВТП обусловлен тем, что при своей широкополосности он обеспечивает регистрацию ПАВ, возбуждаемых при минимально возможной мощности лазерного излучения – 0,014 Вт. Конструкция такого приемника аналогична конструкции прямого ПВТП, описанного в работе [4]. Недостатками этого приемника является необходимость переходной среды, относительно большая площадь контакта с поверхностью образца, а также наличие переотраженных волн в клине. Все это приводит к незначительным искажениям импульса смещения ПАВ. ЭМАП поверхностных волн, хотя и обладает меньшей, по сравнению с клиновидным ПВТП, чувствительностью, однако имеет ряд существенных преимуществ, основные из которых – отсутствие контактной среды и локальность зоны приема (диаметр высокочастотного индуктора в виде прямого проводника длиной 10 мм равен 0,5 мм). Импульсы скоростей смещений в ПАВ, в этом случае, регистрируются без искажений.

По мере нагревания образца от комнатной температуры до температуры несколько выше точки Кюри, осциллограммы регистрируемых импульсов смещений ПАВ снимались на видеокамеру, затем вводились в компьютер и обрабатывались. На рис. 3–5 приведены результаты эксперимента для случая приема ПАВ клиновидным ПВТП и лазерного пятна диаметром 1,5 мм.

III. Результаты и выводы

1. На основании экспериментальных исследований предложена методика определения границ механизмов термооптического возбуждения ПАВ в металлах. В частности, найдено, что действие термоупругого механизма в инваре ограничивается средней поверхностной плотностью тепловой мощности $\langle q \rangle = 10$ Мвт/см² (рис. 2).

2. Обнаружено существенное влияние магнитного фазового перехода на амплитуду смещений в ПАВ в случае их возбуждения за счет термоупругого механизма. При значении $\langle q \rangle = 1,2$ Мвт/см² амплитуда линейно растет с температурой примерно до 150°, после чего начинается ее резкий рост и в области температуры Кюри она достигает максимума (рис. 3 график 1). График 2 соответствует значению $\langle q \rangle = 50$ Мвт/см². В этом случае реализуется гидродинамический и испарительный механизмы возбуждения и амплитуда упругих смещений с ростом температуры ферромагнетика остается практически без изменений.

3. Влияние магнитного фазового перехода на процесс возбуждения ПАВ подтверждается и изменением формы и спектра импульсов ПАВ. Из рис. 4 видно, что в области температуры Кюри

максимум спектра смещается в область более низких частот. Кроме того, с увеличением температуры ферромагнетика происходит смещение максимума импульса примерно на 1,5 мкс к началу развертки, что соответствует увеличению скорости ПАВ. Аналогичные исследования были проведены и с ЭМАП поверхностных волн.

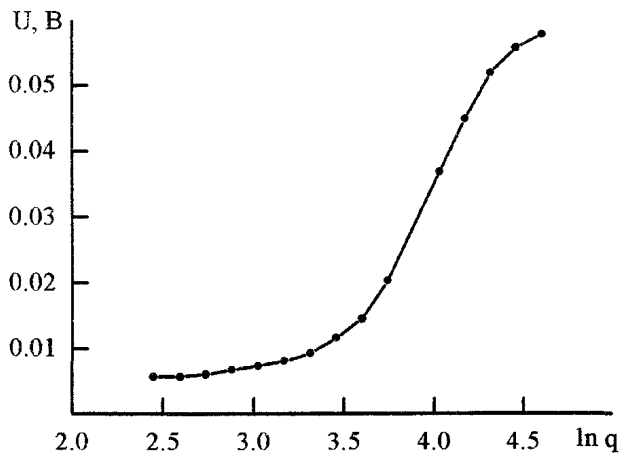


Рис. 2. Зависимость амплитуды импульса скоростей упругих смещений в ПАВ от логарифма средней интенсивности светового потока

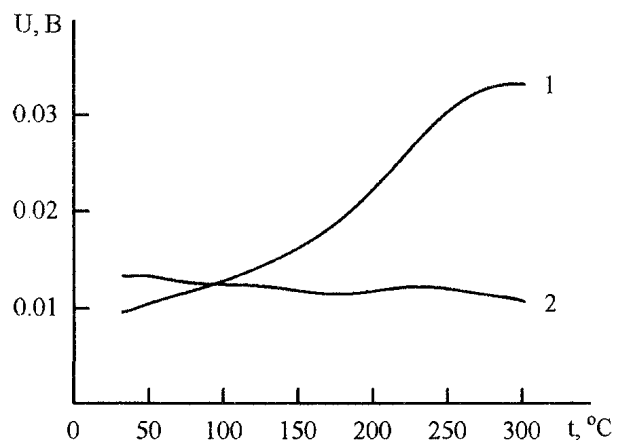


Рис. 3. Зависимость амплитуды импульса ПАВ от температуры: 1 – $\langle q \rangle = 1.2$; 2 – $\langle q \rangle = 50 \text{ МВт/см}^2$

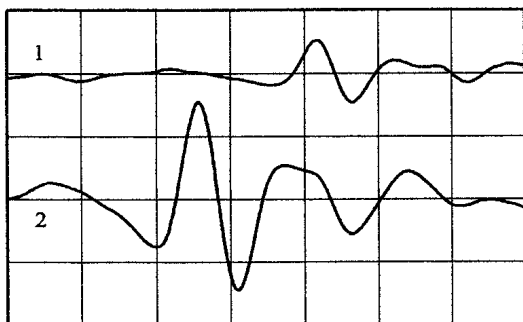


Рис. 4. Форма импульсов смещений в ПАВ: 1 – 27 °С; 2 – 275 °С. Масштаб по горизонтали 1 мкс/см, по вертикали 0,01 В/см

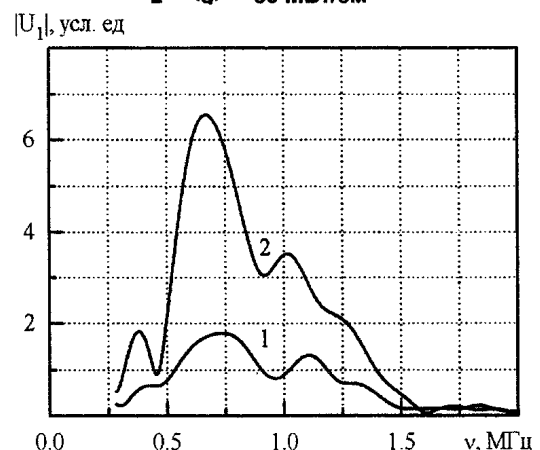


Рис. 5. Спектр импульсов смещений ПАВ: 1 – 27 °С; 2 – 275 °С

IV. Благодарности

Авторы выражают глубокую благодарность Шульгинову А.А. за помощь и полезные замечания.

Работа выполнена при поддержке Минобрания РФ (Тематический план, № 802) и Администрации города.

Литература

1. Исследование влияния магнитного фазового перехода на спектр акустических импульсов, возбуждаемых лазерным импульсом в ферромагнетике / С.Ю. Гуревич, Ю.В. Петров, К.В. Прокопьев, А.А. Шульгинов // Акустический журнал. – 1999. – № 4. – С. 497–501.
2. Исследование температурной зависимости спектров акустических импульсов, возбуждаемых лучом лазера в ферромагнитных металлах / С.Ю. Гуревич, Ю.В. Петров, А.А. Шульгинов, К.В. Прокопьев // Дефектоскопия. – 2000. – № 5. – С. 31–35.
3. О границах механизмов термоакустического преобразования при лазерной генерации ультразвука в металлах / С.Ю. Гуревич, Ю.В. Петров, А.А. Шульгинов, К.В. Прокопьев // Дефектоскопия. – 2001. – № 4. – С. 69–75.
4. Буденков Г.А., Петров Ю.В., Лукманов А.М. Авт. свид. № 406155//Бюлл. изобр. – 1973. – № 45.

ТЕРМОДИНАМИКА УГЛЕРОДА И КИСЛОРОДА В ЖЕЛЕЗЕ

О.И. Качурина, Б.И. Леонович, О.В. Платонова

Описаны термодинамические свойства растворов внедрения на основе железа. Раствор внедрения рассматривался как состоящий из двух подрешеток, одна из которых полностью заполнена основным элементом, в другой субрешетке часть мест занимают растворенные атомы, а вакансии рассматриваются как дополнительный «компонент» системы.

В черной металлургии при анализе процессов восстановления оксидов, науглероживания металла, образования карбидных и оксидных фаз основополагающими являются системы железо–углерод и железо–кислород. Для количественного описания этих процессов необходимо знать термодинамические характеристики как компонентов внедрения, так и растворителя. Однако системный анализ поведения этих элементов в различных гомогенных фазах в широком интервале концентраций и температур практически отсутствует. При расчетах преимущественно используется теория Вагнера для разбавленных растворов. Возможность обобщенного подхода к определению термодинамических характеристик появляется при аналитическом описании фазовых равновесий и построении соответствующей диаграммы состояния.

В настоящей работе описаны термодинамические свойства растворов внедрения на основе железа. При этом раствор внедрения рассматривается как состоящий из двух подрешеток [1]. Одна из них полностью заполнена основным элементом, в другой субрешетке часть мест занимают растворенные атомы, а вакансии (ν) рассматриваются как дополнительный «компонент» системы.

При расчетах фазовых диаграмм необходимо знать концентрационную зависимость термодинамических функций раствора. Для растворов замещения обычно используется полиномиальная модель Редлиха–Кистера, согласно которой зависимость избыточной энергии Гиббса от состава для двойных систем имеет вид [2]:

$$G_m^{\text{изб}} = x_1 x_2 \sum^h L(x_1 - x_2)^h \quad (1)$$

где $^h L$ – коэффициенты полинома; x_1, x_2 – мольные доли компонентов раствора.

Хиллерт с сотрудниками [1, 3] применили это соотношение к металлическим растворам внедрения. В рамках этой подрешеточной модели мольная энергия Гиббса «трехкомпонентного» раствора выражается уравнением

$$G_m = y_i y_\nu G_{i\nu}^0 + y_i y_j G_{ij}^0 + RT \left[y_i \ln y_i + z (y_j \ln y_j + y_\nu \ln y_\nu) \right] + y_i y_j y_\nu ({}^0 L + (y_j - y_\nu) {}^1 L) + G_m^{\text{мар}}. \quad (2)$$

Здесь y_i, y_j – концентрации компонентов, пересчитанные с учетом вакансий: $y_i = x_i / (1 - x_j) = 1$, $y_j = (1/z)[x_j / (1 - x_j)]$, $y_\nu + y_j = 1$; x_i и x_j – мольные доли основного компонента и элемента внедрения соответственно; z – число позиций для внедрения элемента j в расчете на один атом растворителя i (для жидких расплавов и ГЦК-структур $z = 1$, для структуры с ОЦК-решеткой $z = 3$); G_{ij}^0 и $G_{i\nu}^0$ – стандартные значения энергии Гиббса гипотетического бинарного сплава; $G_m^{\text{мар}}$ – магнитная составляющая энергии Гиббса, значение которой рассчитывали по формуле [4]

$$G_m^{\text{мар}} = RT \ln(\beta + 1) f(\tau), \quad (3)$$

где $\tau = T/T_c$; $T_c = 1043$ К – критическая температура магнитного превращения железа; $\beta = 2,22$;

$$f(\tau) = \begin{cases} 1 - 0,90530\tau^{-1} - 0,153\tau^3 - 6,8 \cdot 10^{-3}\tau^9 - 1,53 \cdot 10^{-3}\tau^{15} & (\tau < 1); \\ -6,417 \cdot 10^{-2}\tau^{-5} - 2,037 \cdot 10^{-3}\tau^{-15} - 4,278 \cdot 10^{-4}\tau^{-25} & (\tau > 1). \end{cases} \quad (4)$$

Подставив в выражение (2) значения y_i , получим:

$$G_m = y_v G_{i,v}^0 + y_j G_{i,j}^0 + zRT(y_j \ln y_j + y_v \ln y_v) + y_j y_v [{}^0L + (y_j - y_v) {}^1L] + G_m^{\text{mar}}. \quad (5)$$

В соответствии с этой функцией уравнения парциальных молярных энергий Гиббса (химических потенциалов) компонентов бинарного металлического раствора на основе железа будут иметь вид

$$\bar{G}_j = G_{\text{Fe}_j}^0 + RT \ln \frac{y_j}{1-y_j} + (1-2y_j) {}^0L_{jv}^{\text{Fe}} + (6y_j - 6y_j^2 - 1) {}^1L_{jv}^{\text{Fe}}, \quad (6)$$

$$\bar{G}_{\text{Fe}} = G_{\text{Fe}_v}^0 + zRT \ln(1-y_j) + y_j^2 [{}^0L_{jv}^{\text{Fe}} + (4y_j - 3) {}^1L_{jv}^{\text{Fe}}] + G_{\text{Fe}}^{\text{mar}}.$$

Температурная зависимость параметров модели 0L и 1L , а также изменение энергии Гиббса при фазовом переходе элемента внедрения в железе относительно выбранного стандартного состояния приведены в табл. 1.

Таблица 1

Температурная зависимость параметров гомогенных фаз системы

Фаза*	0L , Дж/моль	1L , Дж/моль	$G_j^0 - G_j^{0\text{ст}}$, Дж/моль
Система железо–углерод			
Ж	$-12223 - 17,669T$	0	$40245 - 1,648T$
α, δ	0	0	$108299 - 39,603T$
γ	$-21079 - 11,555T$	0	$67194 - 7,628T$
Система железо–кислород			
Ж ₂	$433531,8 - 723,032T$	$-348172,5 + 358,1885T$	$-460697,5 + 522,6667T$
W	$-827292,3 + 52,745T$	$351027 - 71,943T$	$226174,5 + 97,103T$
Ж ₁	$478592,3 - 201,009T$	0	$-595751,7 + 226,087T$
α, δ	0	0	$-185708,4 + 95,575T$
γ	0	0	$-152640,5 + 69,976T$

*Ж и Ж₁ – жидкий металлический расплав; α, δ, γ – твердый раствор на основе модификаций железа; Ж₂ – раствор жидких оксидов железа; W – вюстит FeO_{1+x}

При расчете координат линий двухфазного равновесия решали систему уравнений

$$\bar{G}_{i(j)}^{(f)} = \bar{G}_{i(j)}^{(f')} + \Delta_r G_{i(j)}^0(f' \rightarrow f), \quad (7)$$

где $\Delta_r G_{i(j)}^0(f' \rightarrow f) = G_{i(j)}^{0(f)} - G_{i(j)}^{0(f')}$ – изменение стандартной энергии Гиббса i -го (или j -го) компонента при фазовом переходе.

Температурные зависимости фазовых переходов железа определяли по формулам [4]:

$$G_{\text{Fe}}^{0(\text{Ж})} - G_{\text{Fe}}^{0(\gamma)} = -11274 + 163,878T - 22,03T \ln T + 0,0041755T^2, \quad \text{Дж/моль}, \quad (8)$$

$$G_{\text{Fe}}^{0(\alpha)} - G_{\text{Fe}}^{0(\gamma)} = 1462,4 - 8,282T + 1,15T \ln T - 0,00064T^2 + G_{\text{Fe}}^{\text{mar}}, \quad \text{Дж/моль}, \quad (9)$$

где $G_{\text{Fe}}^{\text{mar}}$ рассчитывается по уравнению (3).

Система железо–углерод. Изучению растворимости углерода в жидком железе (Ж), аустените (γ) и феррите (δ, α) посвящены многочисленные экспериментальные и теоретические исследования. При этом авторами предлагаются различные полиномиальные и интерполяционные уравнения для расчета численных значений активности углерода [5, 6]. Термодинамический расчет диаграммы состояния Fe–C с использованием настоящей модели впервые проведен Чипманом [7] и Агреном [8].

Парциальные энергии Гиббса компонентов бинарного раствора системы Fe–C в соответствии с уравнениями (6) будут иметь вид

$$\bar{G}_C = \bar{G}_{\text{Fe}_C} - \bar{G}_{\text{Fe}_v} = RT \ln \frac{y_C}{1-y_C} + (1-2y_C) {}^0L_{Cv}^{\text{Fe}} + (6y_C - 6y_C^2 - 1) {}^1L_{Cv}^{\text{Fe}} + (G_{Cv}^0 - G_C^{\text{orp}}), \quad (10)$$

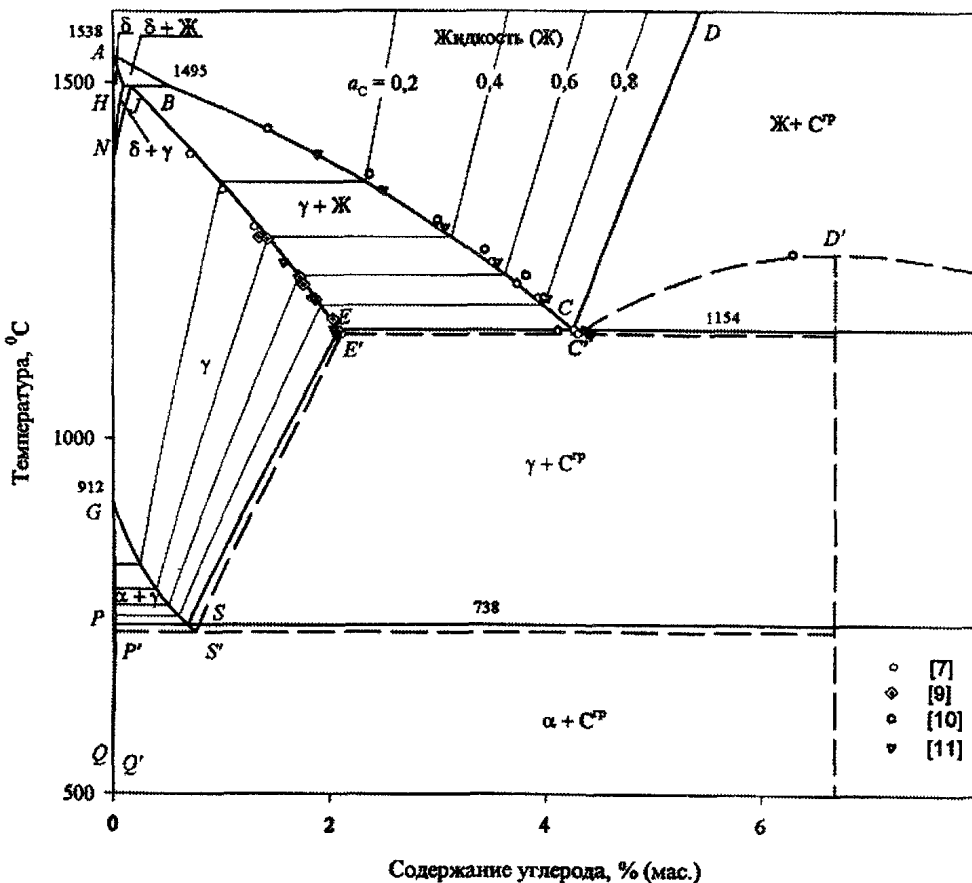
$$\bar{G}_{\text{Fe}} \equiv \bar{G}_{\text{Fe}_v} = G_{\text{Fe}_v}^0 + zRT \ln(1-y_C) + y_C^2 {}^0L_{Cv}^{\text{Fe}} + y_C^2 (4y_C - 3) {}^1L_{Cv}^{\text{Fe}} + G_{\text{Fe}}^{\text{mar}}.$$

При расчетах координат линий двухфазных равновесий системы железо–углерод использованы параметры ${}^0L_{Cv}^{\text{Fe}}$ и ${}^1L_{Cv}^{\text{Fe}}$, а также энергии Гиббса фазовых переходов углерода относительно стан-

дартного состояния графита ($G_{Cv}^{0(f)} - G_C^{0гр}$) (f – индекс фазы) (табл. 1), приведенные в работе [8]. Результаты расчета, полученные при использовании этих зависимостей, анализировались по экспериментальным данным температурной зависимости концентрации углерода на линии двухфазного равновесия аустенит (HE) + жидкий расплав (BC) [7, 9–11], предельной растворимости углерода в аустените (ES) [12] и в жидком расплаве (CD) [7, 13], а также по данным о концентрационной зависимости активности углерода в гомогенных фазах при различных температурах [6, 14].

При расчете метастабильных равновесий с участием цементита принята следующая температурная зависимость изменения энергии Гиббса при образовании Fe_3C [8]:

$$\Delta_f G_T^0(Fe_3C) = G_{Fe_3C}^0 - 3G_{Fe}^{0\gamma} - G_C^{0гр} = 39828 - 193,296T + 22,3452T \ln T, \text{ Дж/моль.} \quad (11)$$



Результаты расчета представлены на диаграмме состояний системы железо–углерод (рис. 1), а значения координат точек невариантных превращений и тип этих превращений приведены в табл. 2.

Рис. 1. Диаграмма состояния системы Fe–C

Характеристика невариантных превращений в системе железо–углерод

Таблица 2

Реакция*	Концентрация углерода в равновесных фазах, % мас.			T, °C	Тип реакции
Система железо–углерод					
$Fe(ж) \leftrightarrow \delta-Fe$	0	0	–	1538	Кристаллизация
$\delta-Fe \leftrightarrow \gamma-Fe$	0	0	–	1394,5	Полиморфное превращение
$\gamma-Fe \leftrightarrow \alpha-Fe$	0	0	–	911,76	
$\delta + Ж \leftrightarrow \gamma$	0,09	0,52	0,18	1495	Перитектическая
$Ж \leftrightarrow \gamma + C^{тп}$	4,26	2,08	100	1154	Эвтектическая
$\gamma \leftrightarrow \alpha + C^{тп}$	0,69	0,02	100	738	Эвтектоидная
Система железо–цементит					
$Fe_3C(тв) \leftrightarrow Fe_3C(ж)$	7,67	7,67	–	1262	Плавление
$Ж \leftrightarrow \gamma-Fe_3C$	4,29	2,10	7,67	1148	Эвтектическая
$\gamma \leftrightarrow \alpha-Fe_3C$	0,77	0,02	7,67	727	Эвтектоидная

* Значения концентрации углерода приведены для последовательности фаз в уравнении для химической реакции.

Система железо–кислород. Система Fe–O остается наиболее важной для металлургии. Однако, несмотря на многочисленные исследования физико-химических свойств веществ этой системы, её диаграмма состояния полностью не изучена [18–21]. Недостаточной и противоречивой является информация о системе при температурах ниже 700 °С и выше ~1200 °С, при которых экспериментальные исследования свойств её затруднены. Усложняет анализ системы и то обстоятельство, что конденсированные вещества являются соединениями переменного состава; их свойства зависят не только от внешних параметров, но и от химического состава. В этой связи попытки теоретического описания свойств системы представляют значительный интерес.

Парциальные молярные энергии Гиббса кислорода и железа в твердых и жидких растворах выражаются формулами аналогичными системе Fe–C. Первое из уравнений соотношения (6) и найденные значения ${}^0L_{Ov}^{Fe}$, ${}^1L_{Ov}^{Fe}$ и $(G_{Ov}^0 - \frac{1}{2}G_{O_2}^0)$ (табл. 1) дают возможность получить температурную зависимость равновесного давления кислорода (атм.) над его растворами в железе:

$$\lg p_{O_2}^{Ж_1} = 2 \lg c_O + c_O \left(-\frac{1000}{T} + 0,42 \right) - \frac{12240}{T} - 1,38; \quad [18] \quad (12)$$

$$\lg p_{O_2}^{\alpha,\delta} = 2 \lg c_O - \frac{19401,5}{T} + 5,0308; \quad (13)$$

$$\lg p_{O_2}^{\gamma} = 2 \lg c_O - \frac{15946,8}{T} + 3,31062, \quad (14)$$

где c_O – концентрация кислорода в растворе, выраженная в мол. %.

При растворении газообразного кислорода парциальная молярная энергия Гиббса (химический потенциал) кислорода в вюстите (W) и расплаве ($Ж_2$) с комплектной решеткой по кислороду может быть выражена уравнением

$$\bar{G}_O = \frac{1}{2} RT \ln p_{O_2} = RT \ln(1 - y_{Fe}) + y_{Fe}^2 [{}^0L_{Fev}^O + (4y_{Fe} - 3) {}^1L_{Fev}^O] + \left(G_{Ov}^0 - \frac{1}{2} G_{O_2}^0 \right), \quad (15)$$

где $G_{Ov}^0 - \frac{1}{2} G_{O_2}^0$ – изменение энергии Гиббса при растворении кислорода в оксидной фазе; $G_{O_2}^0$ – энергия Гиббса чистого газообразного кислорода при давлении 1 атм.

Если состав вюстита выразить формулой FeO_{1+x} , то уравнение (15) будет иметь вид:

$$\bar{G}_O = RT \ln \frac{x}{1+x} + \left(\frac{1}{1+x} \right)^2 {}^0L_{Fev}^O + \frac{1-3x}{(1+x)^3} {}^1L_{Fev}^O + \left(G_{Ov}^0 - \frac{1}{2} G_{O_2}^0 \right). \quad (16)$$

Полученные термодинамические соотношения позволили рассчитать диаграмму состояния системы железо – кислород в области существования вюстита и растворов кислорода в жидком и твердом железе. Результаты расчета представлены на рис. 2. Координаты точек неинвариантных превращений и их тип приведены в табл. 3.

Таблица 3

Характеристика неинвариантных превращений в системе железо–кислород

Реакция	Концентрация кислорода в равновесных фазах, % мол.			T, °C	Тип реакции
$\gamma + W \leftrightarrow \alpha$	0,00069	51,27	0,0027	912	Перитектоидная
$\delta \leftrightarrow \gamma + Ж_2$	0,027	0,018	50,87	1392	Метатектическая
$Ж_1 \leftrightarrow \delta + Ж_2$	0,58	0,032	50,49	1528	Монотектическая
$Ж_2 \leftrightarrow \gamma + W$	50,94	0,017	51,22	1371	Эвтектическая
$W \leftrightarrow \alpha + Fe_3O_4$	51,61	~0	57,14	558	Эвтектоидная
$Ж_2 + Fe_3O_4 \leftrightarrow W$	54,14	~57,14	54,36	1424	Перитектическая
$Ж_2 \leftrightarrow Fe_3O_4$	57,143	57,143	–	1597	Кристаллизация

При определении границ сосуществования гомогенного вюстита, жидких оксидов и растворов кислорода в железе в различных полиморфных и агрегатных состояниях использовали следующие уравнения для давления p_{O_2} (атм), отвечающего равновесию двух конденсированных фаз, указанных в надстрочном индексе:

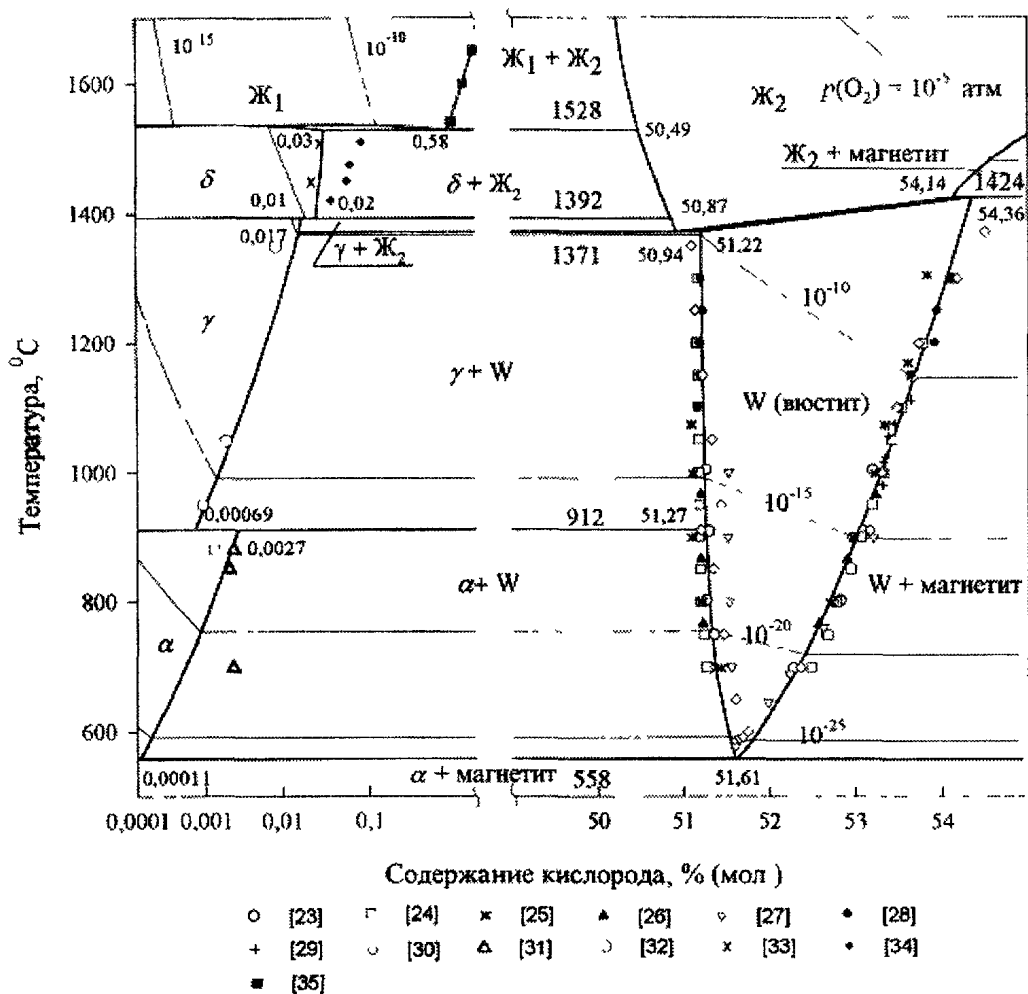


Рис. 2. Диаграмма состояния системы Fe–C

$$\lg p_{O_2}^{\alpha/W} = -\frac{27258,5}{T} + 6,53828; \quad (17)$$

$$\lg p_{O_2}^{\gamma/W} = -\frac{27637,333}{T} + 6,85789; \quad (18)$$

$$\lg p_{O_2}^{\gamma/Ж_2} = -\frac{23189,9}{T} + 4,15265; \quad (19)$$

$$\lg p_{O_2}^{\delta/Ж_2} = -\frac{23120,6}{T} + 4,11104; \quad (20)$$

$$\lg p_{O_2}^{Ж_1/Ж_2} = -\frac{24410,3}{T} + 4,82716; \quad (21)$$

$$\lg p_{O_2}^{W/Fe_3O_4} = -\frac{28485,77}{T} + 9,293487 \lg T - 19,24324; \quad (22)$$

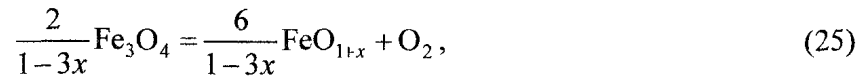
$$\lg p_{O_2}^{Fe_3O_4/Fe_2O_3} = -\frac{21158}{T} + 8,257 \lg T - 14,4926. \quad (23)$$

В приведенных соотношениях и далее давление кислорода (и других газов) выражено в атмосферах, а температура в Кельвин. Уравнения (17)–(22) были получены обработкой результатов расчетов соответствующих равновесий, а зависимость (23) заимствована из работы [22]. Результаты расчета хорошо согласуются с опытными данными Даркена и Гурри [21] и оценками этих линий, приведенными в работе [19]

Система железо–углерод–кислород. Процессы взаимодействия оксидов металлов с углеродом неоднократно были предметом экспериментального и теоретического исследования и доста-

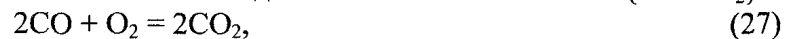
точно подробно изложены в литературе [16, 17]. При этом надежно определены температуры начала процессов восстановления оксидов и составы равновесной газовой фазы. Однако количественный расчет состава образующегося при восстановлении металла представляет определенные трудности, поскольку связан с выбором термодинамической модели, адекватно описывающей поведение двух элементов внедрения (углерода и кислорода) в металлической подрешетке.

В бинарной системе железо–кислород существуют оксиды: Fe_2O_3 – гематит, Fe_3O_4 – магнетит и FeO_{1+x} – вюстит, диссоциацию которых можно описать следующими уравнениями химических реакций:



Температурная зависимость равновесного давления кислорода, отвечающая приведенным реакциям, определяется уравнениями (17)–(23). Комбинирование равновесий реакций (24)–(26) позволяет определить характер фазовых превращений в исследуемой системе в зависимости от равновесного давления кислорода и температуры. При заданной температуре имеется только одно значение кислородного потенциала, обеспечивающее равновесное сосуществование трех конденсированных фаз.

Значение кислородного потенциала может быть также задано составом газовой смеси $\{\text{CO}-\text{CO}_2\}$



$$\lg K_{p(27)} = \frac{29791}{T} + 1,69 \cdot 10^{-4}T - \frac{32400}{T^2} - 9,495 \text{ [15]}.$$

Комбинирование констант равновесия реакций (24)–(26) и (27) позволяет выявить термодинамические закономерности восстановления оксидов железа углеродсодержащими газовыми смесями. В присутствии углерода соотношение $p_{\text{CO}}/p_{\text{CO}_2}$ в газовой фазе (и кислородного потенциала p_{O_2}) задается реакцией



$$\lg K_{p(28)} = -\frac{9369,54}{T} - 0,502766 \lg T - 1,47 \cdot 10^{-4}T + \frac{42207,11}{T^2} + 11,20291 \text{ [15]}.$$

Если в системе присутствует свободный углерод, то кривая Будуара отвечает равновесию графита с магнетитом, вюститом либо насыщенными железом углеродистыми растворами в модификации α , γ и жидким расплавом (Ж_1). В случае гомогенной металлической фазы при вычислении константы равновесия $K_{p(28)}$ необходимо учитывать активность растворенного в металле углерода. Внедрение углерода в железную матрицу, содержащую растворенный кислород, существенно меняет характер границ двухфазных областей в металле.

На рис. 3 представлены значения $\lg p_{\text{O}_2}$, соответствующие сосуществованию двух конденсированных фаз, в зависимости от температуры. Кривые на этом рисунке определяют области устойчивости двух оксидов или металлических фаз, металла и оксидов либо насыщенных углеродом растворов на основе железа. Штриховые линии обозначают границы фазовых областей металла в различных агрегатных состояниях, соответствующих системе железо–кислород.

В области гомогенности вюстита FeO_{1+x} и жидких оксидов тонким линиям отвечают изоконцентрационные составы. Значения x в нестехиометрическом вюстите (W) и жидких оксидах (Ж_2) определяли по зависимости давления кислорода от температуры и состава равновесной фазы, представленной уравнением (16).

Состояние металлической фазы в зависимости от состава углеродсодержащей газовой фазы и температуры представлено на рис. 4. Координаты линий двухфазного равновесия в металле ($\alpha + \gamma$, $\gamma + \delta$, $\delta + \text{Ж}_1$, $\gamma + \text{Ж}_1$) и металла с оксидами рассчитывали по значениям K_p реакции (28) и значениям парциальных молярных энергий Гиббса компонентов трехкомпонентного раствора железо–углерод–кислород

$$\begin{aligned} \bar{G}_{Fe} &= G_{Fe}^0 + zRT \ln(1 - y_C - y_O) + y_C^2 \cdot {}^0L_{Cv}^{Fe} + y_O^2 \cdot {}^0L_{Ov}^{Fe} + G_{Fe}^{mag}, \\ \bar{G}_C &= RT \ln \frac{y_C}{1 - y_C - y_O} + (1 - 2y_C - y_O) {}^0L_{Cv}^{Fe} + (G_{Cv}^0 - G_C^{0rp}), \\ \bar{G}_O &= RT \ln \frac{y_O}{1 - y_O - y_C} + (1 - 2y_O - y_C) {}^0L_{Ov}^{Fe} + (G_{Ov}^0 - \frac{1}{2}G_{O_2}^0). \end{aligned} \quad (29)$$

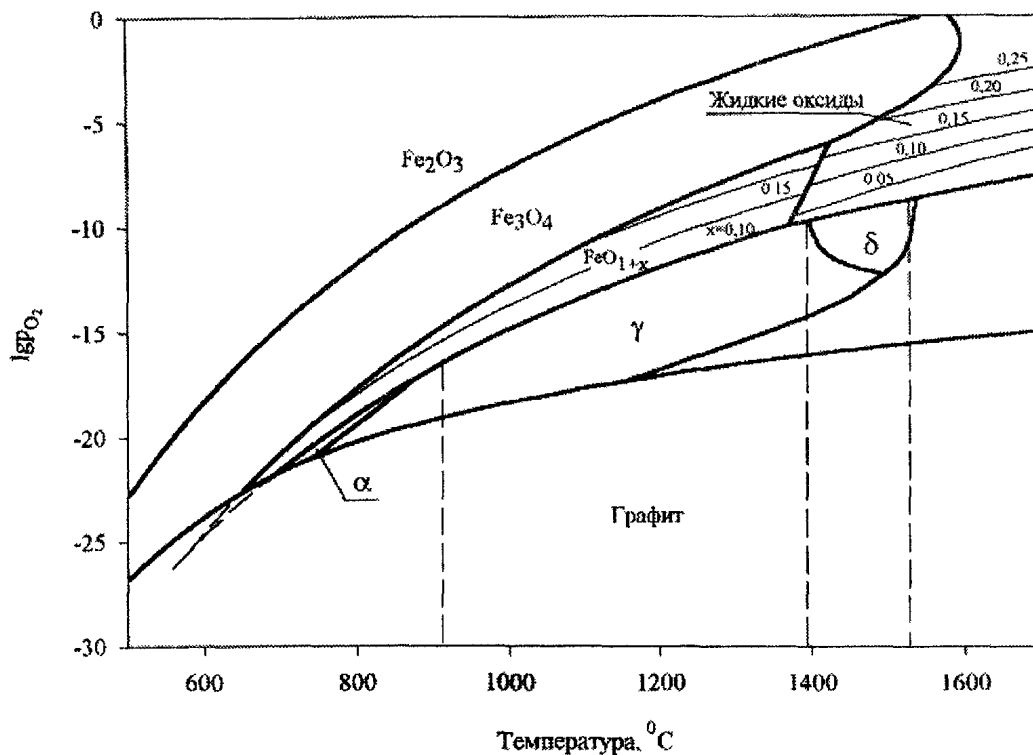


Рис. 3. Влияние углерода на фазовые равновесия в системе Fe-C

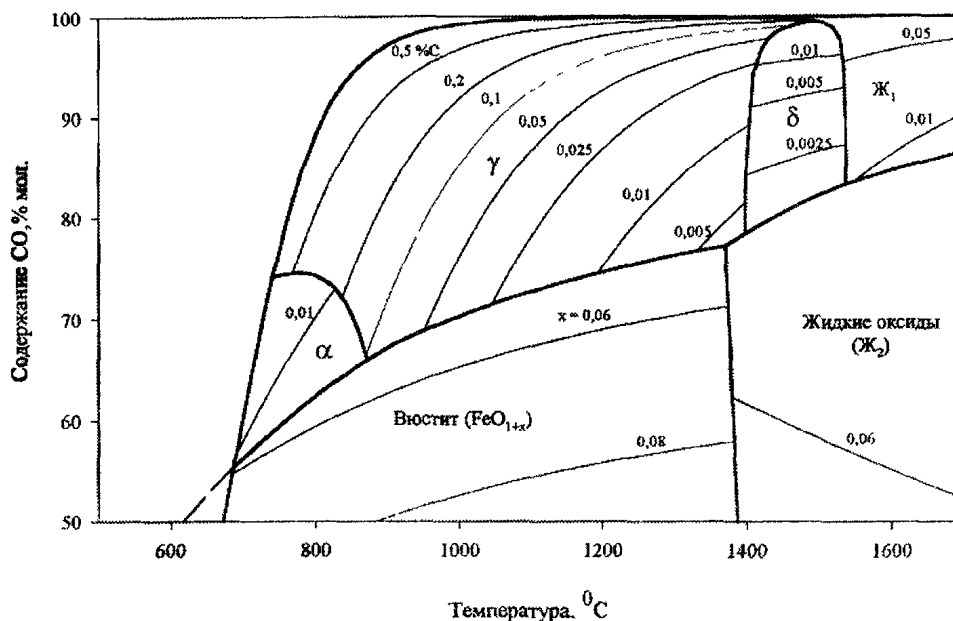


Рис. 4. Науглероживание железа при восстановлении оксидов

Рассчитанные по уравнениям (29) значения активности углерода позволяют определить из соотношения для $K_{p(28)}$ при заданной температуре (или концентрации CO в газовой фазе) равновесное содержание CO (мол. %) (либо температуру двухфазного равновесия). Таким образом, задаваясь температурой и составом газовой фазы, можно предсказать значение равновесной концентрации углерода в металлической фазе.

Более полное представление о характере фазовых равновесий можно получить из диаграммы, построенной в координатах температура–содержание углерода в металле (рис. 5). В этом случае диаграмма представляет собой политермическую поверхность растворимости углерода и кислорода в железе. Тонким линиям отвечают изобары кислорода, равновесного с конденсированными фазами. Правая часть диаграммы отвечает состоянию насыщенных железоуглеродистых сплавов. При низких концентрациях углерода металлические фазы будут находиться в равновесии с жидкими и твердыми оксидами. В табл. 4 приведены результаты расчета содержания углерода в металле и $\lg p_{O_2}$ в газовой фазе для невариантных равновесий в системе Fe–C–O.

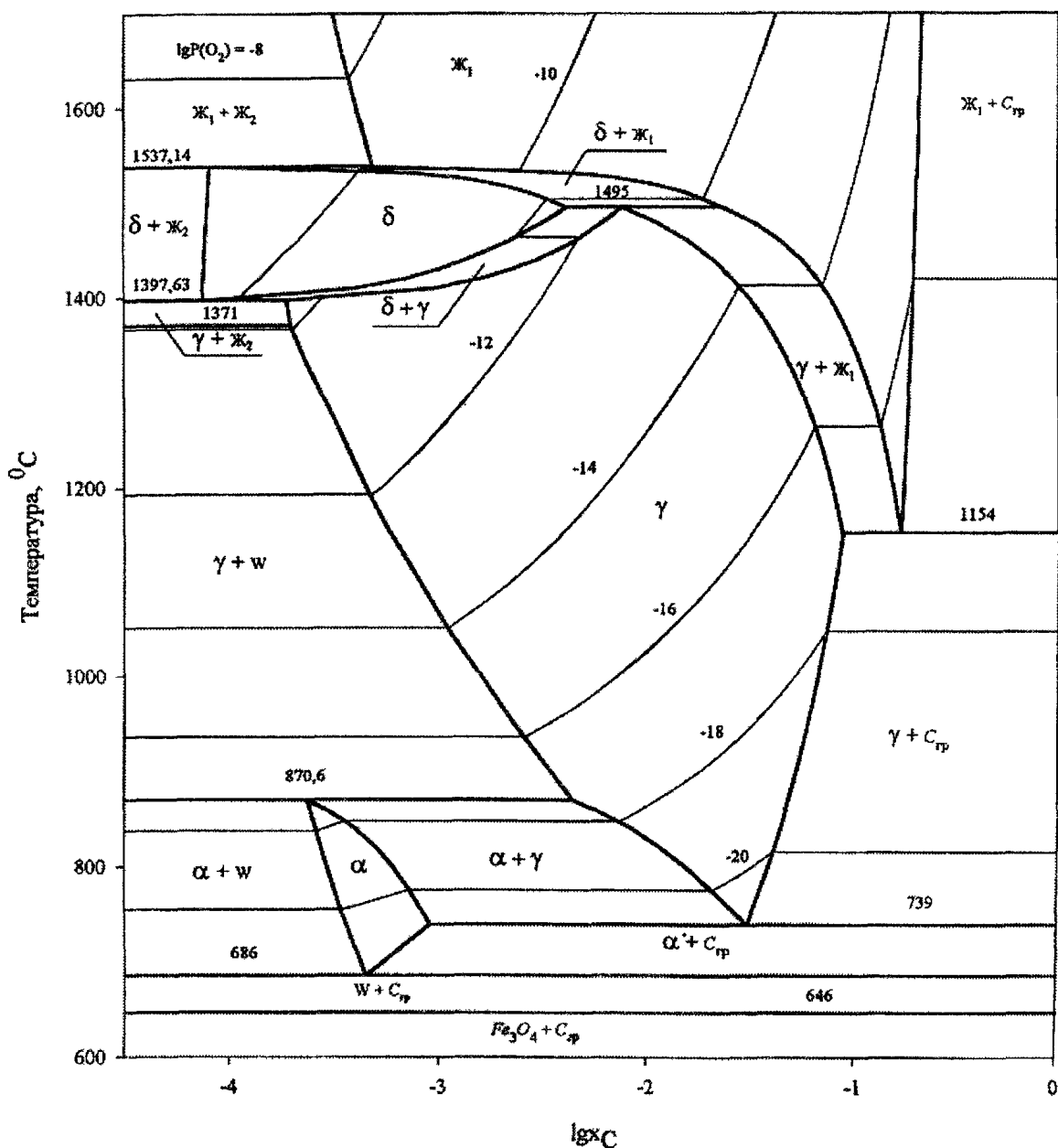


Рис. 5. Фазовые равновесия в системе Fe–C–O

Нонвариантные равновесия в системе Fe–C–O

$t, ^\circ\text{C}$	Реакция	$\lg p_{\text{O}_2}, (\text{атм})$	Содержание углерода, % мас.			Тип превращения
1537,14	$\text{Ж}_1 + \text{Ж}_2 \leftrightarrow \delta$	-8,66	0,0103	–	0,0017	Синтектическое
1495	$\delta + \text{Ж}_1 \leftrightarrow \gamma$	-12,25	0,0898	0,518	0,1669	Перитектическое
1397,63	$\delta \leftrightarrow \text{Ж}_2 + \gamma$	-9,73	0,0016	–	0,0040	Метатектическое
1371	$\text{Ж}_2 \leftrightarrow \text{W} + \gamma$	-9,95	–	–	0,0042	Эвтектическое
1154	$\text{Ж}_1 \leftrightarrow \gamma + \text{C}_{\text{гр}}$	-17,33	4,256	2,085	100	Эвтектическое
870,6	$\gamma + \text{W} \leftrightarrow \alpha$	-17,30	0,09635	–	0,00505	Перитектоидное
739	$\gamma \leftrightarrow \alpha + \text{C}_{\text{гр}}$	-21,00	0,678	0,0196	100	Эвтектоидное
686	$\alpha \leftrightarrow \text{W} + \text{C}_{\text{гр}}$	-21,90	0,00997	–	100	Эвтектоидное

Выводы

С использованием термодинамической модели растворов внедрения рассчитаны и построены диаграммы состояния железо – углерод, железо – кислород и железо – углерод – кислород. Определены значения концентраций и температур, отвечающих нонвариантным равновесиям в этих системах. Установлены концентрационные и температурные пределы равновесного сосуществования оксидов железа с металлическими фазами.

Работа выполнена при финансовой поддержке РФФИ по проекту р2001 ур чел-02-17.

Литература

- Hillert M., Staffanson L.I. The regular solution model for stoichiometric phases and ionic melts // Acta Chemica Scandinavica. – 1970. – V. 24, № 10. – P. 3618–3626.
- Морачевский А.Г. Термодинамика расплавленных металлических и солевых систем. – М.: Металлургия, 1987. – 240 с.
- Hillert M., Jansson B., Sundman B., Agren J. A two-sublattice model for molten Solutions with different tendency for ionization // Met. Trans. – 1985. – V. 16A. – P. 261–266.
- Hertzman S., Sundman B. A thermodynamic analysis of the Fe–Cr system // Met. Transaction. – 1987. – V. 18A. – P. 175–187.
- Ban-Ya S., Elliott J., Chipman J. Thermodynamics of austenitic Fe–C alloys // Metall. Trans. – 1970. – V. 1. – P. 1313–1320.
- Бурылев Б.П. Термодинамика металлических растворов внедрения. – Изд-во Ростовского университета, 1984. – 160 с.
- Chipman J. Thermodynamics and phase diagram of the Fe–C system // Metall. Trans. – 1972. – V. 3. – P. 56–63.
- Agren J. A Thermodynamic Analysis of the Fe–C and Fe–N Phase Diagrams // Metal. Trans. – 1979. – V. 10A. – P. 1847–1852.
- Benz M., Elliott J. Линия солидуса аустенита и пересмотр диаграммы железо–углерод // Trans. TMS-AIME. – 1961. – V. 221. – P. 323–331.
- Buckley R.A., Hume-Rothery W.J. Liquidus and Solidus relations in iron-rich iron-carbon alloys // Iron Steel Inst. – 1960. – V. 196, № 12. – P. 403–406.
- Chicco B., Thorpe W.R. Experimental determination of the austenite + liquid phase boundaries of the Fe–C system // Met. Trans. – 1982. – V. 13A, № 7. – P. 1293–1297.
- Шюлер П., Эркес Г.-В. Измерение активности углерода в сплавах Fe–X–C в интервале температур от 900 до 1100 °С при метано-водородной равновесной системе. Пер. № В-64287. – М.: Всесоюзный центр переводов научно-технической литературы и документации, 1981. – 23 с.
- Шанк Ф. Структуры двойных сплавов. Пер. с англ. – М., Металлургия, 1973. – 760 с.
- Gocken N. A. Comments on the activity of carbon in liquid iron // Steel research. – 1994. – V. 65, № 4. – P. 125–127.
- Краткий справочник физико-химических величин: Изд. 8-е, перераб. Под ред. А.А. Равделя и А.М. Пономаревой. – Л.: Химия, 1983. – 233 с.

16. Есин О.А., Гельд П.В. Физическая химия пирометаллургических процессов. Часть 1. Реакции между газообразными и твердыми фазами.
17. Леонидова М.Н., Шварцман Л.А., Шульц М. Физико-химические основы взаимодействия металлов с контролируруемыми атмосферами. – М.: Металлургия, 1980. – 264 с.
18. Wriedt H.A. The Fe–O (Iron-Oxygen) System // *Journal of Phase Equilibria*. – 1991. – V. 12, № 2. – P. 170–200.
19. Лыкасов А.А., Карел К., Мень А.Н., Варшавский М.Т., Михайлов Г.Г. Физико-химические свойства вюститита и его растворов. – Свердловск: УНЦ АН СССР, 1987. – 230 с.
20. Sundman B. An Assessment of the Fe–O System // *Journal of Phase Equilibria*. – 1991. – V. 12, № 1. – P. 127–140.
21. Darken L.S., Gurry R.W. The System Iron-Oxygen. I. The wustite field and related equilibria // *J. Amer. Chem. Soc.* – 1945. – V. 67. – P. 1398–1412; The system iron-oxygen. II. Equilibrium and thermodynamics of liquid oxide and other phases // *J. Amer. Chem. Soc.* – 1946. – V. 68. – P. 798–816.
22. Jacobsson E. Solid state em Studies of the Systems FeO–Fe₃O₄ and Fe₃O₄–Fe₂O₃ in the temperature range 1000–1600 K // *Scand. J. Metallurgy*. – 1985. – V. 14. – P. 252–256.
23. Ackermann R.J., Sandford R.W. A Thermodynamic Study of the Wustite Phase // USAEC, Argonne Nat. Lab., Rep. ANL-7250. – 1966.
24. Лыкасов А.А., Кузнецов Ю.С., Пилько Е.И., Шишков В.И., Кожеуров В.А. Термодинамика вюститита // *ЖФХ*. – 1969. – Т. 43, № 12. – С. 3124–3125.
25. Campserveux J., Boureau G., Picard C., Gerdanian P. Determination of the Boundaries of the Iron Protoxide Domain at High Temperature // *Rev. Int. Hautes Temp. Refract.* – 1969. – V. 6. – P. 165–174.
26. Rizzo F.E., Smith J.V. Coulometric Titration of Wustite // *J. Physi. Chem.* – 1968. – V. 72, № 2. – P. 485–488.
27. Asao H., Ono K., Yamaguchi A., Moriyama J. Thermodynamic Properties of Wustite (FeO_{1+y}) // *Mem. Fac. Eng., Kyoto Univ.* – 1970. – V. 33. – P. 66–70.
28. Takayama E., Kimizuka N. Thermodynamic Properties and Subphases of Wustite Field Determined by Means of Thermogravimetric Method the Temperature Range of 1100–1300 °C // *J. Electrochem. Soc.* – 1980. – V. 127(4). – P. 970–976.
29. Giddings R.A., Gordon R.S. Solid-State Coulometric Titration: Critical Analysis and Application to Wustite // *J. Electrochem. Soc.* – 1974. – V. 121, № 6. – P. 793–800.
30. Vallet P. The Fe–O (Iron-Oxygen) Phase Diagram in the Range of the Nonstoichiometric Monoxide and Magnetite at the Fe-Rich Limit: Reduction Diagrams // *Bull. Alloy Phase Diagrams*. – 1989. – V. 10, № 3. – P. 209–218.
31. Sifferlen R. On the Subject of the Insolubility of Oxygen in Zone-Refined α -Iron // *C.R. Acad. Sci. (Paris)*. – 1958, № 247. – P. 1608–1611.
32. Swisher J.H., Turkdogan E.T. Solubility, Permeability and Diffusivity of Oxygen in Solid Iron // *Trans. Metall. Soc. AIME*. – 1967, № 239. – P. 426–431.
33. Hepworth M.T., Smith R.P., Turkdogan E. T. Solubility, Permeability and Diffusivity of Oxygen in BCC Iron // *Trans. Metall. Soc. AIME*. – 1966, № 236. – P. 1278–1283.
34. Tankins E.S., Gokcen N.A. Solubility of Oxygen in Delta Iron // *Trans. ASM*. – 1961, № 53. – P. 843–852.
35. Левинский Ю.В. Диаграммы состояния металлов с газами: Справочник. – М., Металлургия, 1975. – 296 с.

РАВНОВЕСИЕ ТВЕРДОГО ШПИНЕЛЬНОГО РАСТВОРА $\text{Fe}_3\text{O}_4 - \text{CuFe}_2\text{O}_4$ С ДЕЛАФОССИТОМ И ГЕМАТИТОМ

А.Э. Катков, А.А. Лыкасов

Детально исследованы условия равновесия шпинельного раствора с сопряженными фазами в интервале температур 1073–323 К. Исследование выполнено методом измерения ЭДС гальванических ячеек с твердым кислород-проводящим электролитом.

Фазовые равновесия в системе Fe–Cu–O наиболее подробно исследованы при температуре 1273 К [1–8]. В работе [2] информация по термодинамическим свойствам веществ системы Fe–Cu–O, накопленная к тому времени, была обобщена в виде диаграммы $\lg p_{\text{O}_2}$ – состав. Впоследствии отдельные равновесия были исследованы более детально, и диаграмма приняла вид, изображенный на рис. 1. Очевидно, что в системе Fe–Cu–O кроме двойных соединений систем Fe–O и Cu–O образуются два тройных соединения: CuFeO_2 (делафоссит) и CuFe_2O_4 . Феррит меди CuFe_2O_4 и магнетит Fe_3O_4 при температуре 1273 К неограниченно растворяются друг в друге. Этот раствор с одной стороны сопряжен с Fe_2O_3 , а с другой, при более низких парциальных давлениях кислорода, – с медью, делафосситом или CuO . С оксидом CuO в равновесии находится раствор, близкий по составу к CuFe_2O_4 , а с металлической медью – раствор на основе Fe_3O_4 . Растворимость меди в вюстите не учитывается. Согласно нашим исследованиям [9] растворимость меди в вюстите при 1273 К достигает 0,77 ат.% и это оказывает влияние на структуру диаграммы в области устойчивости вюститной фазы.

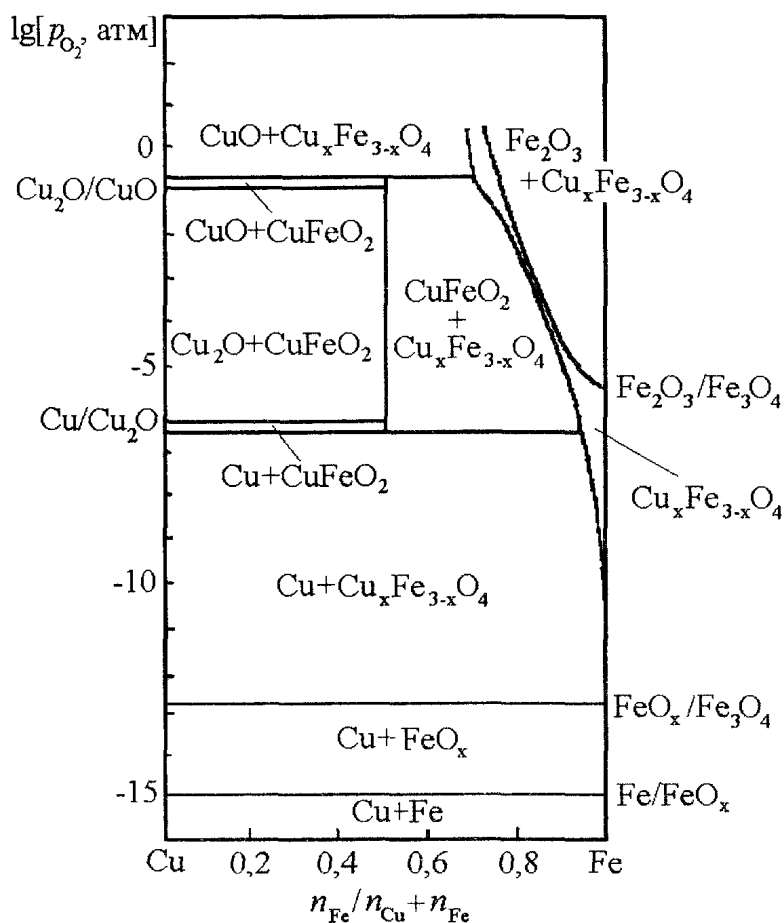


Рис. 1. Диаграмма состояния системы Fe–Cu–O при 1273 К

Фазовые равновесия в системе Fe–Cu–O при температурах ниже 1273 К исследовались в работах [4–8]. Полученные данные разноречивы. По данным работы [5] неограниченная растворимость CuFe_2O_4 и Fe_3O_4 сохраняется даже при 900 К, а согласно исследованию [7] уже при 1240 К реализуется равновесие делафоссита с Fe_2O_3 , то есть при этой температуре CuFe_2O_4 ограниченно растворяется в Fe_3O_4 . Изотермическое сечение фазовой диаграммы системы Fe–Cu–O при 1240 К, построенное авторами работы [7], приведено на рис. 2.

В настоящей работе более детально исследованы условия равновесия шпинельного раствора с сопряженными фазами в интервале температур 1073–1323 К. Исследование выполнено методом измерения ЭДС гальванических ячеек с твердым кислород-проводящим электролитом. Конструкции ячеек и процедура эксперимента подробно описаны в работах [10–11].

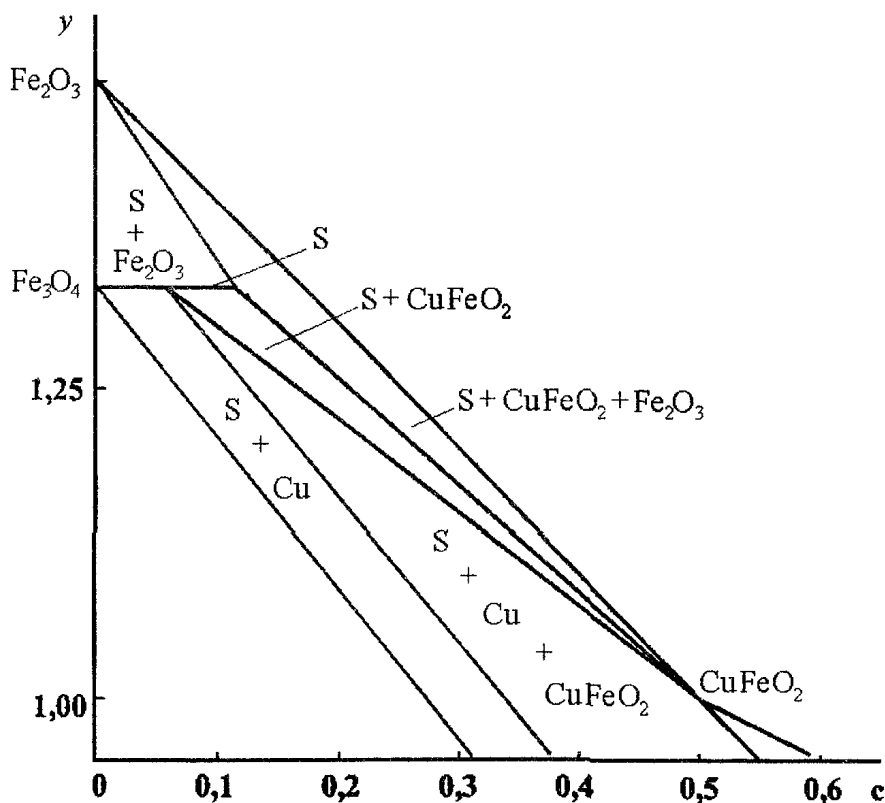


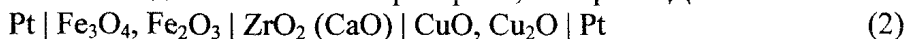
Рис. 2. Фазовая диаграмма системы Fe-Cu-O при 1200 К: S - шпинельный раствор

Для определения условий двухфазных равновесий твердого шпинельного раствора $Fe_3O_4 - CuFe_2O_4$ с дельтафосфитом и гематитом были измерены ЭДС гальванических элементов



в интервале температур 1090...1280 К. Здесь CuO, Cu_2O - эталонное вещество, $Cu_c Fe_{1-c} O_y$ - исследуемый образец. Величина c изменялась от 0,09 до 0,275 при $y = 1,30$ и от 0 до 0,125 при $y = 1,35...1,4$.

С целью предотвращения изменения состава электродов во время опытов равновесную смесь CuO, Cu_2O и исследуемый электрод $Cu_c Fe_{1-c} O_y$ замуровывали в тигли из электролита. Рабочие поверхности тиглей шлифовали друг к другу. Возможность использования смеси CuO, Cu_2O в качестве электрода сравнения в подобных ячейках проверяли, измерив ЭДС элемента



в интервале температур 1123...1323 К. Результаты представлены на рис. 3. Рассчитанные по ним значения $lg(P_{O_2}, атм)$ описываются уравнением

$$lg P_{O_2} = +9,1 - 13\,000/T \tag{3}$$

и хорошо согласуются с литературными данными, рис. 4.

Зависимости ЭДС элемента (1) от температуры представлены на рис. 5-7. Они хорошо воспроизводились при нагревании и охлаждении элемента, а также при повторном проведении опыта. Ошибка в определении величины ЭДС составляла ± 3 мВ. Из рис. 5-7 видно, что зависимости ЭДС от температуры имеют излом. При температурах ниже температуры излома ЭДС элемента (1) не зависит от валового

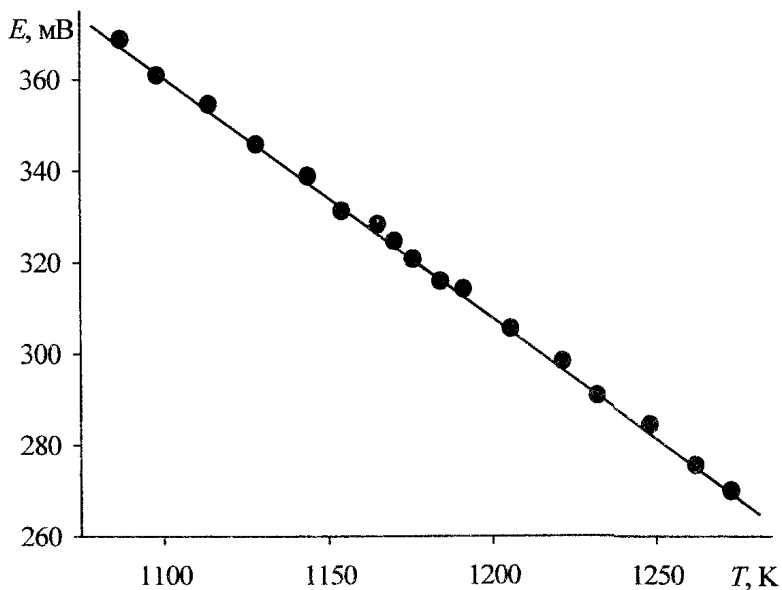


Рис. 3. Зависимость ЭДС элемента (2) от температуры

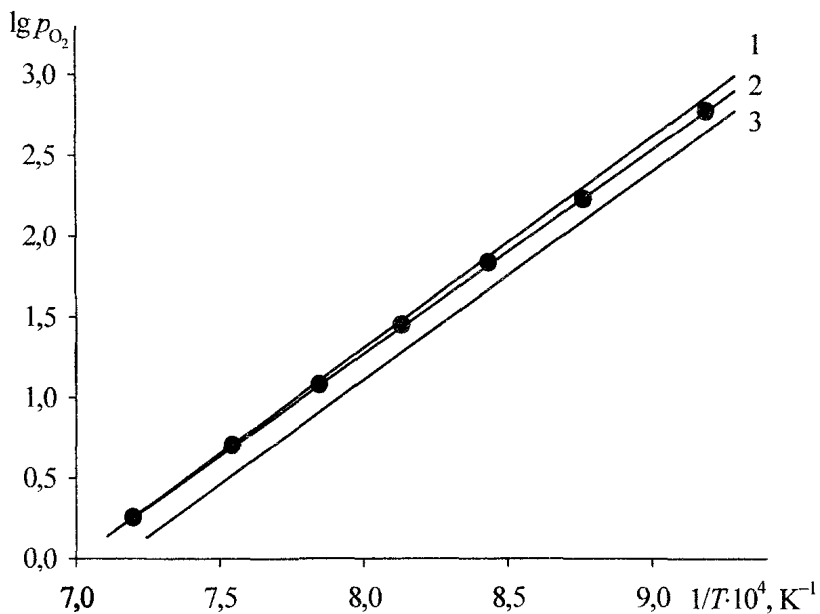


Рис. 4. Температурная зависимость $\lg P_{O_2}$ для равновесной смеси $CuO-Cu_2O$: 1 – [12]; 2 – наши данные; 3 – [13]

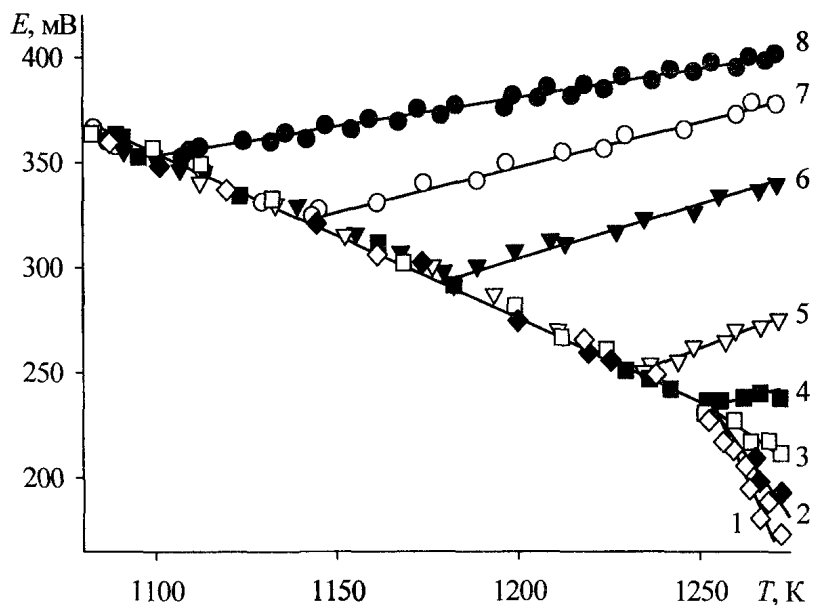


Рис. 5. Зависимость ЭДС элемента (1) от температуры: $y = 1,30$; c : 1 – 0,198; 2 – 0,191; 3 – 0,185; 4 – 0,178; 5 – 0,15; 6 – 0,12; 7 – 0,1; 8 – 0,09

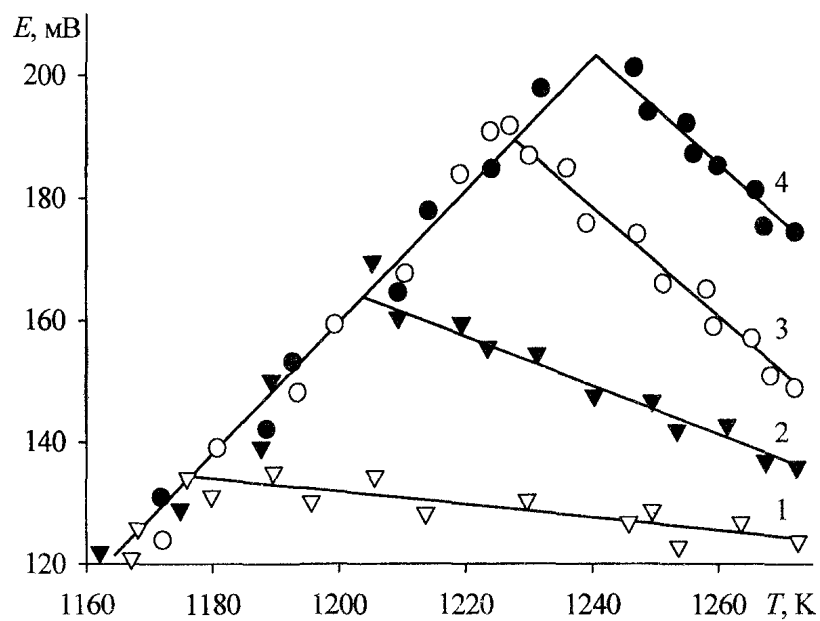


Рис. 6. Зависимость ЭДС элемента (1) от температуры: $y = 1,30$; c : 1 – 0,275; 2 – 0,26; 3 – 0,25; 4 – 0,225

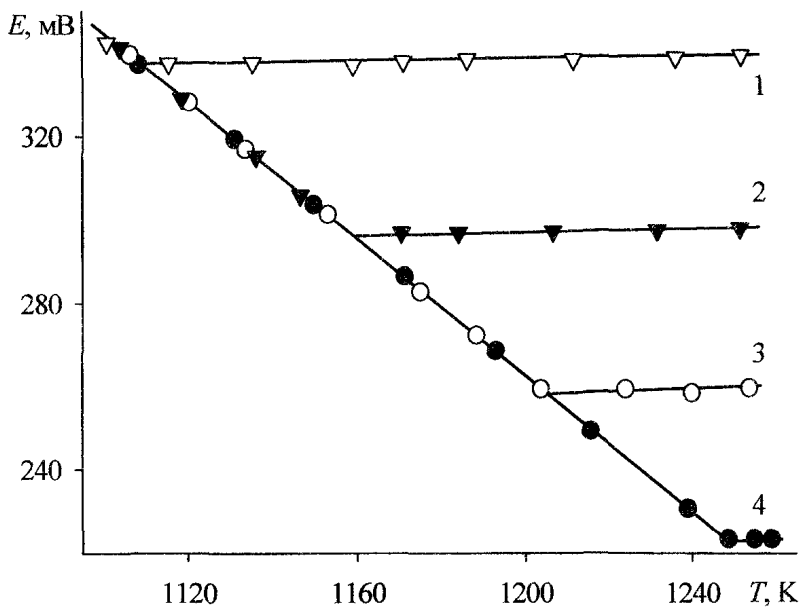


Рис. 7. Зависимость ЭДС элемента (1) от температуры: $y = 1,375$; c : 1 – 0,025; 2 – 0,050; 3 – 0,075; 4 – 0,10

состава образца, то есть в этих условиях образцы должны быть трехфазными. Это было подтверждено рентгенофазовым анализом образцов, закаленных с 1173 К. Образцы состояли из смеси шпинельной фазы, делафоссита и Fe_2O_3 . Для трехфазных образцов, фигуративные точки которых лежат левее линии соединяющей на фазовой диаграмме системы Fe–Cu–O делафоссит и Fe_2O_3 (см. рис. 2), ЭДС элемента (1) отличается от ЭДС этого элемента с образцами, фигуративные точки которых правее этой линии. Из этого следует, что на фазовой диаграмме системы Fe–Cu–O при температурах ниже ~1250 К по разные стороны указанной линии имеются две разные трехфазные области. Они возникают в результате диспропорционирования шпинельного

раствора с образованием Fe_3O_4 и делафоссита $CuFeO_2$. В результате в системе Fe–Cu–O появляются два раствора шпинельного типа: один раствор на основе Fe_2O_3 , другой – на основе $CuFeO_2$. Давления кислорода в системе, при которых реализуются эти трехфазные равновесия, описываются уравнениями:

$$S_1 + F + D: \lg(P_{O_2}, \text{атм}) = (-29\,068/T + 18,46) \pm 0,04; \quad (4)$$

$$S_1 + F + D: \lg(P_{O_2}, \text{атм}) = (18\,060/T - 19,22) \pm 0,04, \quad (5)$$

Согласно уравнениям (4) и (5) температура начала распада шпинельного раствора $T_{кр} = 1251$ К. При этой температуре система моновариантна. Ее фазовая область на диаграмме состояния системы вырождена в линию, соединяющую шпинель и продукты «распада». По координатам изломов на зависимостях ЭДС от температуры (рис. 5–7) была построена кривая диспропорционирования твердого шпинельного раствора. Она приведена на рис. 8.

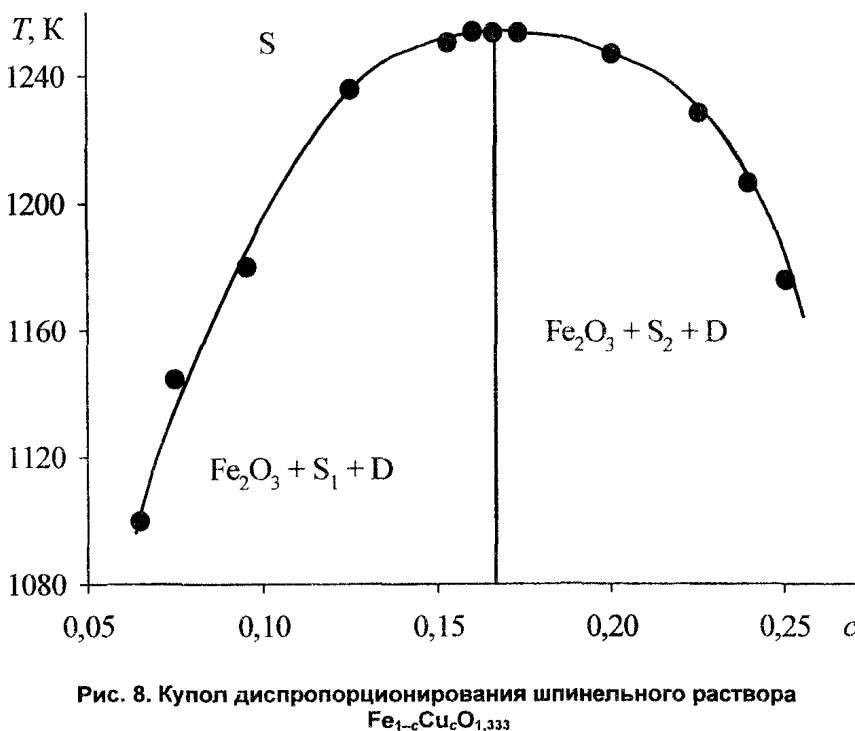


Рис. 8. Купол диспропорционирования шпинельного раствора $Fe_{1-c}Cu_cO_{1,333}$

Выше критической температуры 1251 К образуется непрерывный ряд твердых растворов от Fe_3O_4 , до $CuFe_2O_4$. При более низких температурах в результате диспропорционирования шпинели с образованием Fe_2O_3 и делафоссита на диаграмме зависимости T от c появляются две области. В каждой из областей линия купола отвечает составу шпинели, находящейся в равновесии одновременно с Fe_2O_3 и делафосситом. При температурах выше температуры излома на зависимостях ЭДС элемента (1) от температуры (рис. 5, 6) образцы содержат две фазы; в этих условиях в равновесии

находится делафоссит (D) и шпинельный раствор (S). Состав шпинельного раствора может быть установлен по положению конод в двухфазной области S + D фазовой диаграммы системы. Так как делафоссит является соединением практически постоянного состава CuFeO_2 , коноды в области равновесия твердый шпинельный раствор – делафоссит представляют собой прямые, соединяющие фигуративную точку делафоссита и фигуративные точки исследуемых образцов. Экстраполяция этих прямых до нижней границы области гомогенности шпинельного раствора ($y = 1,333$) дает составы шпинели, находящейся в равновесии с делафосситом.

По данным ЭДС были рассчитаны значения равновесного давления кислорода над шпинелью, находящейся в равновесии с делафосситом при температуре 1273K, рис. 9 (область III). Шпинельный раствор, состав которого соответствует области I рис. 9, находится в равновесии с вюститной фазой, а области II – в равновесии с практически чистой медью. Условия равновесия шпинельного раствора с вюститом и медью обсуждаются в работе [9]. Как видно из рис. 9, равновесное давление кислорода существенно увеличивается с повышением содержания меди в твердом шпинельном растворе.

Условия равновесия шпинельного раствора с Fe_2O_3 были определены по данным измерения ЭДС элемента (1), приведенным на рис. 7. При этом исходили из предположения, что шпинельный

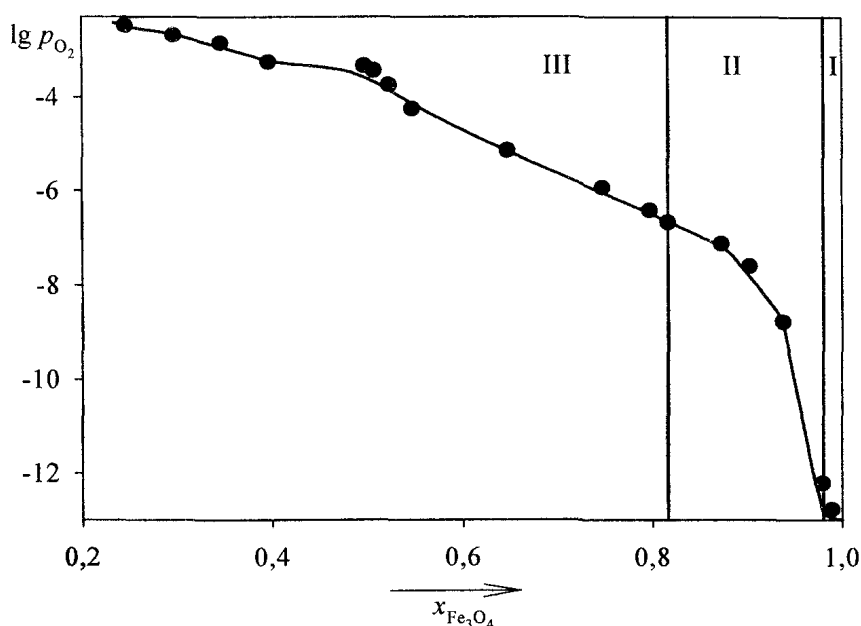


Рис. 9. Зависимость $\lg(P_{\text{O}_2}, \text{атм})$ от состава шпинельного раствора при температуре 1273 К

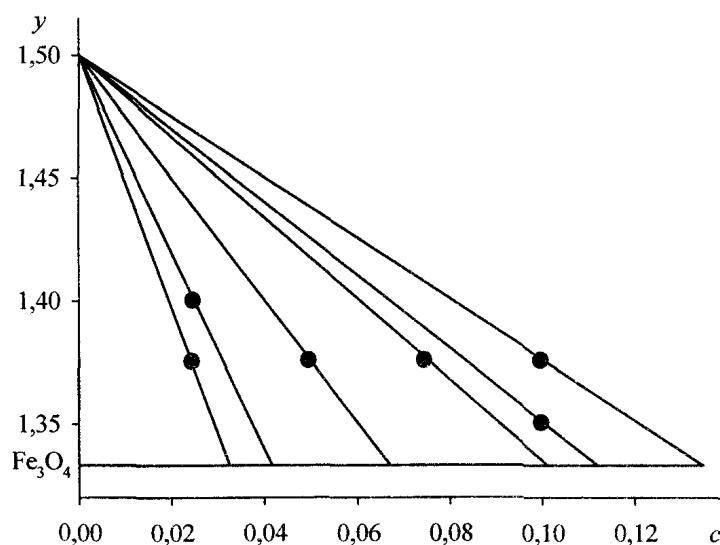


Рис. 10. Составы шпинельного раствора, находящиеся в равновесии с гематитом при 1273 К

раствор находится в равновесии с чистым Fe_2O_3 . Таким образом, для того, чтобы установить состав шпинели в двухфазной области, содержащей шпинельный раствор и Fe_2O_3 , достаточно определить положение фигуративной точки образца на фазовой диаграмме системы Fe–Cu–O. Соединив эту точку с фигуративной точкой Fe_2O_3 ($c = 0, y = 1,5$), получим коноду, пересечение которой с границей шпинельного раствора определит его состав. Результаты таких оценок состава шпинельного раствора в исследуемых образцах при температуре 1273 К приведены на рис. 10.

Используя графики зависимости $E = E(T)$ для образцов, содержащих шпинельный раствор и Fe_2O_3 , рис. 7, и уравнение Нернста, были рассчитаны давления кислорода, равновесные этим образцам. Это позволило установить зависимость P_{O_2} от состава шпинельного раствора, находящегося в равновесии с Fe_2O_3 . Она приведена на рис. 11. Как и для шпинельного раствора, сопряженного с низкокислородными фазами, при равновесии с Fe_2O_3 давление кислорода существенно повышается с увеличением содержания меди в растворе.

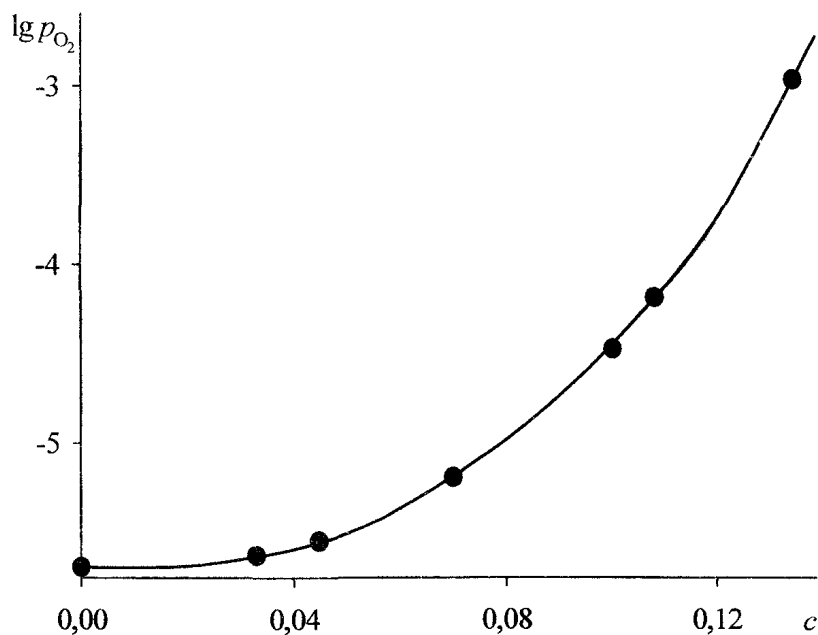


Рис. 11. Зависимость $\lg(P_{O_2}, \text{атм})$ от состава шпинельного раствора, находящегося в равновесии с Fe_2O_3 при температуре 1273К

Выводы

Методом измерения ЭДС гальванических ячеек с твердым электролитом исследованы условия равновесия шпинельного раствора системы Fe–Cu–O с делафосситом и гематитом. Показано, что Fe_3O_4 и $CuFe_2O_4$ неограниченно растворяются при температуре выше 1250 К. При более низких температурах раствор диспропорционирует с образованием делафоссита и Fe_2O_3 . При этом образуются два раствора: один на основе Fe_3O_4 , другой – на основе $CuFe_2O_4$. В интервале температур 1100–1250 К определены составы шпинельных фаз, находящихся в равновесии с делафосситом и Fe_2O_3 .

Работа выполнена по гранту РФФИ.

Литература

1. Третьяков Ю.Д. Термодинамика ферритов. – Л.: Химия, 1967. – 304 с.
2. Залазинский А.Г., Балакирев В.Ф., Чуфаров Г.И. Диаграмма «давление кислорода – состав» в системе Cu–Fe–O при 1000 °C // ЖФХ. – 1969. – Т. 43, № 6. – С. 1636–1637.
3. Jacob K.T., Fitzner K., Alcock C.B. Activities in the spinel Solid Solution. Phase Equilibria and Thermodynamic Properties of Ternary Phases in the System Cu–Fe–O // Metal. Trans. – 1977. – V. 8B, № 3. – P. 451–460.
4. Schaefer S.C., Hundley G.L., Block F.E. and other. Phase Equilibria and X-Ray Diffraction Investigation of the System Cu–Fe–O // Metal. Trans. – 1970. – V. 1, № 9. – P. 2557–2561.
5. Katsutoshi O., Yoshihiro I., Ario Y. and other. The Equilibrium Phase Diagram of Cu–Fe–O System // J. Japan Inst. Metals. – 1972. – V. 36, № 7. – P. 701–704.
6. Paulsson H., Rosen E., Terman R. Thermodynamic Studies of High Temperature Equilibria. Potentiometric determination of the stability of $CuFeO_2$ using a galvanic cell involving stabilized ZrO_2 as solid electrolyte // Chem. Scr. – 1975. – V. 8, № 5. – P. 193–196.
7. Fredriksson M., Rosen E. Thermodynamic Studies of High Temperature Equilibria. On phase relations and equilibrium oxygen pressures in the system Cu– $CuFeO_2$ – Fe_2O_3 – Fe_3O_4 at temperatures 1100–1300 K // Chem. Scr. – 1976. – V. 9, № 3. – P. 118–121.
8. Eriksson G., Tegman R. Thermodynamic Studies of High Temperature Equilibria. Determination of the stability of $CuFe_2O_4$ by emf measurements using stabilized ZrO_2 as solid electrolyte // Chem. Scr. – 1976. – V. 10, № 4. – P. 164–167.
9. Катков А.Э., Лыкасов А.А. Вюститные твердые растворы в системе Cu–Fe–O // Неорганические материалы. – 1999. – Т. 35, № 7. – С. 836–839.
10. Лыкасов А.А., Карел К., Мень А.Н., Варшавский М.Т., Михайлов Г.Г. Физико-химические свойства вюстита и его растворов. – Свердловск: УНЦ АН СССР, 1987. – 230 с.
11. Santander N.H., Kubashewski O. The thermodynamics of the copper-oxygen system // High temperatures-high pressures. – 1975. – V. 7. – P. 573–582.
12. Левинский Ю.В. P–T–x-диаграммы состояния двойных металлических систем. Справочник. Кн. 1. – М.: Металлургия, 1990. – 400 с.

$$G = U^s - TS^s + P^\Psi V^s, \quad (2)$$

используя фундаментальное уравнение Гиббса для гетерогенной системы, можно показать, что дифференциал свободной энергии Гиббса гетерогенной системы, содержащей такой зародыш, равен

$$dG_{\text{тер}} = -S^s dT + V^s dP^\Psi - (P^\Psi - P^\Phi) dV^\Phi + \gamma dA^\Phi + \sum_{j=1}^m \mu_{R_j} dq_{R_j} + \sum_{i=1}^n \mu_i dq_i. \quad (3)$$

Здесь S – энтропия, V – объемы, μ – химические потенциалы соответствующих величин, γ – коэффициент поверхностной энергии, индекс Ψ относится к фазе раствора, Φ – к фазе зародыша, s – ко всей системе.

Применяя к выражению (3) теорему Эйлера [6], можно получить выражение для свободной энергии Гиббса системы, содержащей зародыш, и, соответственно, выражение для изменения свободной энергии Гиббса при образовании такого зародыша. Проведя необходимые математические преобразования и используя формулы для химических потенциалов компонентов зародышей, приведенные в [7], получим выражение изменения свободной энергии Гиббса системы при образовании зародыша продуктов реакций (1):

$$\Delta G(q, c_1, \dots, c_{n-1}) = kT \left(\kappa q^{2/3} - q \sum_{i=1}^n c_i \ln \varepsilon_i \right). \quad (4)$$

Здесь q – количество молекул в зародыше, c_i – молекулярная доля i -го компонента в зародыше, κ – величина, характеризующая поверхностную энергию зародыша и определяемая для сферической частицы выражением

$$\kappa = (36\pi)^{1/3} v^{2/3} \gamma / kT \quad (5)$$

(v – средний молекулярный объем зародыша, зависящий только от его состава), ε_i – пересыщение раствора по i -му компоненту рассматриваемого зародыша:

$$\varepsilon_i = \varepsilon_i^0 / a_i,$$

ε_i^0 – абсолютное пересыщение раствора по i -му компоненту:

$$\varepsilon_i^0 = \prod_{j=1}^m (a_{R_j})^{\nu_{ij}} / K_i, \quad (6)$$

K_i – константа равновесия i -ой химической реакции в (1), a_i – активность i -го компонента зародыша в макрофазе, имеющей тот же состав, что и зародыш, a_{R_j} – активность j -го компонента раствора.

Поведение функции $\Delta G(q, c_1, \dots, c_{n-1})$ оказывает решающее влияние на кинетику зародышеобразования. В связи с этим целесообразно провести ее исследование как математической функции.

Реальный зародыш состоит из фазы продуктов реакций и части переходного слоя, ограниченных поверхностью натяжения. Размер термодинамического зародыша (число молекул q в нем) может принимать любые положительные значения, в том числе и нецелые и меньшие единицы. Доли молекул компонентов зародыша c_i ($i=1, \dots, n-1$), могут выражаться любыми числами, при которых сумма долей равна единице. Считаем также функцию $\kappa(c_1, \dots, c_{n-1})$ дважды дифференцируемой, а активности компонентов $a_i(c_1, \dots, c_{n-1})$ – непрерывно дифференцируемыми функциями, удовлетворяющими в соответствующих предельных случаях законам, аналогичным закону Генри для двухкомпонентного раствора. Можно показать, что из этих условий следует непрерывность функции $\Delta G(q, c_1, \dots, c_{n-1})$ при всех допустимых значениях c_1, \dots, c_{n-1} .

Для дальнейших рассуждений найдем производные функции $\Delta G(q, c_1, \dots, c_{n-1})$. Дифференцируя (4) по q и c_k с учетом уравнения Гиббса–Дюгема [6] получим

$$\frac{\partial \Delta G}{\partial q} = kT \left[\frac{2}{3} \kappa q^{-1/3} - \sum_{i=1}^n c_i \ln \varepsilon_i \right]. \quad (7)$$

$$\frac{\partial \Delta G}{\partial c_k} = kT \left[\kappa'_k q^{2/3} - q (\ln \varepsilon_k - \ln \varepsilon_n) \right] \quad (k=1, \dots, n-1). \quad (8)$$

$$\frac{\partial^2 \Delta G}{\partial q^2} = -\frac{2}{9} kT \kappa q^{-4/3}, \quad (9)$$

$$\frac{\partial^2 \Delta G}{\partial c_k \partial q} = kT \left[\frac{2}{3} \kappa'_k q^{-1/3} - \ln \varepsilon_k + \ln \varepsilon_n \right], \quad (10)$$

$$\frac{\partial^2 \Delta G}{\partial c_k \partial c_l} = kT \left[\kappa''_{kl} q^{2/3} - q \left(\frac{\partial \ln \varepsilon_k}{\partial c_l} - \frac{\partial \ln \varepsilon_n}{\partial c_l} \right) \right], \quad (11)$$

где $\kappa'_k = \frac{\partial \kappa}{\partial c_k}$ и $\kappa''_{kl} = \frac{\partial^2 \kappa}{\partial c_k \partial c_l}$, $k, l = 1, \dots, n-1$.

В термодинамике изменение функции ΔG характеризует возможность самопроизвольного протекания процесса в том или ином направлении. Если переход из начального состояния в конечное сопровождается увеличением свободной энергии Гиббса системы, то для макрообъектов это означает, что самопроизвольное протекание такого процесса невозможно. Однако зародыш является малым объектом, для которого большое значение имеют флуктуации параметров состояния системы. Термодинамика описывает поведение зародыша «в среднем». Поэтому, если она запрещает переход из какого-либо начального состояния в конечное, то для зародыша это означает, что вероятность прямого перехода меньше вероятности обратного, но прямой переход при этом не исключается.

Для дальнейшего анализа исследуем поведение функции $\Delta G(q, c_1, \dots, c_{n-1})$ при условии постоянства состава зародыша. Приравнявая (7) к нулю можно определить критический размер зародыша q^M при заданном постоянном составе:

$$q^M(c_1, \dots, c_{n-1}) = \left[\frac{2}{3} \kappa / \sum_{i=1}^n c_i \ln \varepsilon_i \right]^3, \quad (12)$$

начиная с которого возможен самопроизвольный рост частицы вследствие уменьшения ее свободной энергии Гиббса. При размерах зародыша меньших q^M его рост возможен лишь за счет флуктуаций.

Так как функция $\Delta G(q, c_1, \dots, c_{n-1})$ непрерывна, существует некоторый состав $(c_1^k, \dots, c_{n-1}^k)$ и соответствующее этому составу значение $q^k = q^M(c_1^k, \dots, c_{n-1}^k)$, определяемое выражением (12), при котором значение функции $\Delta G(q^k, c_1, \dots, c_{n-1})$ минимально. Так как в соответствии с общей теорией флуктуаций количество зародышей каждого типа очень резко (обратно экспоненциально) зависит от соответствующего этому типу значения функции ΔG , то основная часть зародышей, которым удается вырасти в макрофазу, будет «проходить» через окрестность точки с координатами $(q^k, c_1^k, \dots, c_{n-1}^k)$. Назовем эту точку критической и определим область критических составов из выражений (8), приравнявая их к нулю:

$$\kappa'_k q^{2/3} - q(\ln \varepsilon_k - \ln \varepsilon_n) = 0, \quad (k = 1, \dots, n-1). \quad (13)$$

Полученные $n-1$ уравнений (13) определяют такой состав (молярные доли компонентов c_1^m, \dots, c_{n-1}^m) для каждого размера q зародыша, при котором величина свободной энергии Гиббса ΔG минимальна. Таким образом, исследуя зависимости $c_1^m(q), \dots, c_{n-1}^m(q)$, можно проследить изменение состава зародышей в процессе роста.

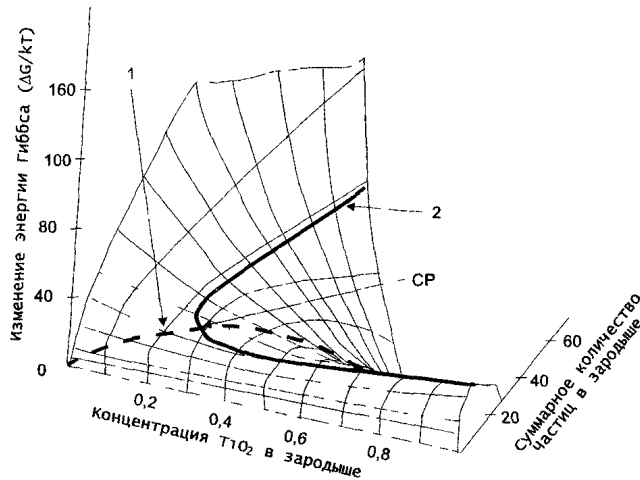
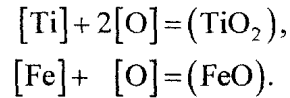
Совокупность выражения (12) и $n-1$ уравнений (13) определяет координаты критической точки, через окрестность которой переход зародышами поверхности ΔG наиболее выгоден энергетически. Численное значение величины ΔG в этой точке можно определить, подставляя выражение (12) в (4):

$$\Delta G^k = \frac{1}{3} kT \kappa(c_1^k, \dots, c_{n-1}^k) \left(q^M(c_1^k, \dots, c_{n-1}^k) \right)^{2/3}. \quad (14)$$

Обобщая вышеизложенное, можно заключить, что для определения координат критической точки, необходимо найти совместное решение уравнения (12), которое в общем случае определяет подмножество значений q^M в зависимости от состава зародыша, и $n-1$ уравнений (13), опреде-

ляющих критические составы для каждого значения q . Если система этих уравнений будет иметь несколько решений, то нужно выбрать то из них, при котором величина ΔG будет наименьшей.

В качестве примера на рис. 1 показана поверхность $\Delta G(q, c_1)$ при образовании зародышей продуктов раскисления металла титаном при 1600 °С и 0,05 % кислорода и 1 % титана. При этом протекают реакции:



Как видно из рис. 1, первоначально образующиеся зародыши состоят из почти чистого оксида железа (линия 1), но по мере роста содержание TiO_2 в них растет. Выросший зародыш будет состоять из почти чистого оксида титана.

На рис. 2 показано изменение состава зародышей продуктов раскисления железа комплексом Al-Si по мере его роста при 0,1 % [Al], 0,6 % [Si], 0,001 % [O]. Первоначально образовавшиеся зародыши обогащены оксидом железа. По мере роста происходит изменение их состава в сторону увеличения содержания оксида алюминия. Выросшие зародыши состоят из почти чистого оксида алюминия.

Рис. 1. Поверхность изменения свободной энергии Гиббса при образовании зародыша продуктов раскисления титаном. 1 – «линия минимумов» – линия наиболее выгодных в энергетическом отношении составов; 2 – «линия максимумов» – линия точек с наивысшей для каждого состава зародыша свободной энергией Гиббса; CP – критическая точка

Алгоритм расчета параметров критического зародыша

В качестве исходных данных необходимо задать начальные концентрации компонентов раствора (c_{R_i}), константы равновесий реакций химического взаимодействия между этими компонентами (K_i), коэффициенты поверхностного натяжения между продуктами химических реакций и исходным раствором, коэффициенты диффузии исходных компонентов в растворе (D_{R_i}) и их молекулярные массы (M_{R_i}).

А1. Определяем абсолютные пересыщения ε_i^0 раствора относительно i -го компонента зародыша:

$$\varepsilon_i^0 = \prod_{j=1}^m (a_{R_j})^{v_{ij}} / K_i, \quad (i=1, 2, \dots, n).$$

А2. Задаемся некоторыми исходными значениями (c_1, c_2, \dots, c_{n-1}) концентраций независимых компонентов в критическом зародыше.

Рис. 2. Изменение состава зародышей продуктов раскисления железа комплексом Al-Si по мере его роста при 0,1 % [Al], 0,6 % [Si], 0,001 % [O]. Числа у кривой – размер зародыша (суммарное количество молекул оксидов в нем); CP – критическая точка

А3. Для состава (c_1, c_2, \dots, c_{n-1}) определяем активности a_i компонентов фазы зародыша, значения величины κ и ее производных:

$$\kappa'_i = \frac{\partial \kappa}{\partial c_i} \quad \text{и} \quad \kappa''_{ik} = \frac{\partial^2 \kappa}{\partial c_i \partial c_k} \quad i, k = 1, \dots, n-1.$$

A4. Определяем пересыщения ε_i раствора по компонентам зародыша:

$$\varepsilon_i = \varepsilon_i^0 / a_i, \quad (i=1, 2, \dots, n).$$

A5. Находим значения вспомогательных функций f_i и их производных f'_{ij} :

$$f_i = 3\kappa'_i \sum_{k=1}^n c_k \ln \varepsilon_k - 2\kappa (\ln \varepsilon_i - \ln \varepsilon_n),$$

$$f'_{ij} = 3\kappa''_{ij} \sum_{k=1}^n c_k \ln \varepsilon_k + 3\kappa'_i (\ln \varepsilon_j - \ln \varepsilon_n) - 2\kappa'_j (\ln \varepsilon_i - \ln \varepsilon_n) + 2\kappa \left(\frac{\partial \ln a_i}{\partial c_j} - \frac{\partial \ln a_n}{\partial c_j} \right),$$

$$(i, j=1, 2, \dots, n-1)$$

Составляем матрицы

$$\hat{f} = \begin{pmatrix} f_1 \\ \dots \\ f_{n-1} \end{pmatrix}, \quad \hat{f}' = \begin{pmatrix} f'_{11} & \dots & f'_{1,n-1} \\ \dots & \dots & \dots \\ f'_{n-1,1} & \dots & f'_{n-1,n-1} \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} c_1 \\ \dots \\ c_{n-1} \end{pmatrix}.$$

A6. Находим матрицу уточненных значений состава критического зародыша:

$$\hat{C}'' = \hat{C} - (\hat{f}')^{-1} \cdot \hat{f}.$$

A7. Если для всех i $0 < c_i'' < 1$, переходим к A9.

A8. $\hat{C}''' = (\hat{C}'' + \hat{C})/2$. Переходим к A7.

A9. Если для всех i $|c_i'' - c_i| < \delta$, где δ – приемлемая точность вычислений, переходим к A11.

A10. $C = \hat{C}'''$. Переходим к A3.

A11. Находим число молекул q^k в критическом зародыше:

$$q^k = \left[\frac{2}{3} \kappa / \sum_{k=1}^n c_k \ln \varepsilon_k \right]^3.$$

Анализ приведенных результатов расчетов показывает, что до определенных значений концентраций раскислителя и кислорода в металле интенсивность зародышеобразования практически равна нулю. Затем, при определенных значениях концентраций раскислителя и кислорода происходит резкий скачок значения интенсивности. Далее, с увеличением концентраций реагентов, интенсивность увеличивается значительно медленнее. Максимальные значения, которые достигает интенсивность зародышеобразования, определяется раскислительными способностями элементов, участвующих в реакциях.

Заключение

Таким образом, получены уравнения, описывающие термодинамику процесса образования критических зародышей, и предложен алгоритм расчета размеров и состава критических зародышей продуктов реакций.

Работа выполнена при поддержке гранта губернатора Челябинской области P2001УЧЕЛ-02-10 и гранта Минобразования 01.200103088.

Литература

1. Гиббс Дж.В. Термодинамика. Статистическая механика. – М.: Наука, 1982. – 584 с.
2. Фольмер М. Кинетика образования новой фазы. – М.: Наука, 1986. – 208 с.
3. Зельдович Я.Б. Избранные труды. Химическая физика и гидродинамика. – М.: Наука, 1984. – 376 с.
4. Френкель Я.И. Кинетическая теория жидкостей. – Л.: Наука, 1975. – 592 с.
5. Turnbull D., Fisher J.C. Rate of Nucleation in Condensed System // J. Chem. Phys. – 1949. – V. 17, 11. – P. 71–73.
6. Пригожин И., Дефэй Р. Химическая термодинамика. – Новосибирск: Наука, 1966. – 510 с.
7. Дрозин А.Д. Химические потенциалы компонентов малых тел // Расплавы. – 1990. – № 4. – С. 65–69.

**ВЕСТНИК
ЮЖНО-УРАЛЬСКОГО
ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА**

№ 3 (12) 2002

**Серия
«МАТЕМАТИКА, ФИЗИКА, ХИМИЯ»
Выпуск 2**

Издательство Южно-Уральского государственного университета

ИД № 00200 от 28.09.99. Подписано в печать 5.11.2002. Формат 60×84 1/8. Печать трафаретная.
Усл. печ. л. 10,67. Уч.-изд. л. 11,16. Тираж 150 экз. Заказ 374/106.

Группа МЭНП Издательства. 454080, г. Челябинск, пр. им. В.И. Ленина, 76.