TARTU RIIKLIKU ÜLIKOOLI
TOIMETISED

УЧЕНЫЕ ЗАПИСКИ
ТАРТУСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS

500

ARVUTUSMEETODID DIFERENTSIAAL- JA INTEGRAALVÕRRANDITE LAHENDAMISEKS

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ РЕШЕНИЯ ДИФФЕРЕНЦИАЛЬНЫХ И ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ

Matemaatika- ja mehhaanikaalaseid töid
Труды по математике и механике

XXV

TARTU RIIKLIKU ÜLIKOOLI TOIMETISED
УЧЕНЫЕ ЗАПИСКИ
ТАРТУСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS
ALUSTATUD 1893.a. VIHIK 500 ВЫПУСК ОСНОВАНЫ В 1893.г.

ARVUTUSMEETODID DIFERENTSIAAL- JA INTEGRAALVÕRRANDITE LAHENDAMISEKS

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ РЕШЕНИЯ ДИФФЕРЕНЦИАЛЬНЫХ И ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ

Matemaatika-ja mehhaanikaalaseid töid
Труды по математике и механике
XXV

ТАРТУ 1979

Настоящее издание является межвузовским сборником высших учебных заведений Эст. ССР

Уч.зап.Тартуск.ун-та,1979,500,3-23

ДИСКРЕТИЗАЦИЯ АБСТРАКТНОГО ГИПЕРБОЛИЧЕСКОГО УРАВНЕНИЯ

С. Пискарев
Тартуский государственный университет

В данной работе рассмотрена общая теория аппроксимации для дифференциального уравнения второго порядка в банаховом пространстве. Несколько отступая от традиционных методов теории разностных схем, исследование аппроксимации уравнения гиперболического типа сводится к изучению аппроксимации косинус и синус-функций. Как показано в [14-16, 23] в случае корректно поставленной задачи для дифференциального уравнения второго порядка косинус-функция играет ту же роль, что и C_0-полугруппа e^{tA} для задачи Коши $u'(t)=Au(t)$, $u(0)=u^0$. Следуя взглядам работы [8], § 2 посвящен полудискретной аппроксимации (здесь доказана теорема, аналогичная теореме Троттера—Като для полугрупп), когда уравнения рассматриваются в разных пространствах; а § 3, 4 - полной дискретизаци задачи. Используя введенный в § 3 дискретный косинус, удалось получить оценки скорости сходимости, не предполагая дополнительной гладкости решения по времени. Устанавливаются условия устойчивости методов.

Пусть E и E_n - банаховы пространства. Всюду далее мы предполагаем, что заданы линейные непрерывные отображения $p_n : E \to E_n$, причем

$$\|p_n x\| \longrightarrow \|x\|, \quad x \in E.$$

На простяжении всей работы используется терминология и понятия дискретной сходимости ($\mathcal{P}$-сходимость); подробные сведения о ней имеются в [1-2, 27]. Через $\mathcal{L}(E)$ обозначается множество непрерывных линейных отображений из E в E. Положим также $R_+=[0,\infty)$ и $\mathcal{R}(\lambda;A)=(\lambda-A)^{-1}$ резольвента оператора A.

§ 1. Косинус и синус-функции

В банаховом пространстве E рассмотрим задачу Коши

$$\frac{d^2}{dt^2}u(t)=Au(t),\ u(0)=u^0,\ u'(0)=u^1, \tag{1}$$

где A - замкнутый линейный оператор в E с непустым резольвентным множеством $\rho(A)\neq\emptyset$. Решением задачи (1) называют дважды непрерывно дифференцируемую функцию $u(t)$ такую, что $u(t)\in\mathcal{D}(A), t\in[0,T]$ и выполняется (1).

Говорят, что задача Коши (1) равномерно-корректно поставлена, если

1^0 Найдется плотное подмножество $\mathcal{D}\subset E$ такое, что для любых $u^0, u^1\in\mathcal{D}$ существует решение $u(t)$ задачи (1) со свойством

$$u^{(k)}(0)=\lim_{t\to 0+}u^{(k)}(t)=u^k,\quad k=0,1.$$

2^0 В случае $u^0\to 0$ и $u^1\to 0$ имеем $u(t)\to 0$ равномерно по $t\in[0,T]$.

Также как с решением корректной задачи Коши для уравнения первого порядка связано семейство полугрупп операторов (экспоненциальная оператор-функция), с решением задачи (1) можно связать косинус-функцию (косинус оператор-функцию).

Однопараметрическое семейство $C(t), t\in R$ линейных ограниченных отображений $C(t): E\to E$ называется косинус-функцией, если

$$1^0\quad C(t+h)+C(t-h)=2C(t)C(h)\quad \forall t,h\in R; \tag{2}$$

2^0 $C(0)=I$ (единичный оператор);

3^0 $C(t)x$ сильно непрерывна по $t\in R$ для любого $x\in E$.

Следуя [24–25, 15] мы ассоциируем с косинус-функцией синус-функцию по формуле

$$S(t)=\int_0^t C(s)\,ds,\quad t\in R. \tag{3}$$

Инфинитезимальный оператор (производящий оператор, генератор) косинус-функции определяется как

$$A=\frac{d^2}{dt^2}C(0) \tag{4}$$

Известно, что A – замкнутый оператор с областью определения $\mathcal{D}(A)=\{x\in E:\ C(t)x$ дважды непрерывно дифференцируема по $t\in R\}$. Как показано в [14–15, 23] задача Коши (1) равномерно-корректна тогда и только тогда, когда A является инфинитезимальным оператором косинус-функции. Определим также множество

$$\mathcal{D}^1=\{x\in E:\ C(t)x \text{ непрерывно дифференцируема по } t\}.$$

Решение равномерно-корректной задачи (1) при $u^0\in\mathcal{D}(A)$ и $u^1\in\mathcal{D}^1$ дается по формуле $u(t)=C(t)u^0+S(t)u^1$. Если $u^0\notin\mathcal{D}(A)$ или $u^1\notin\mathcal{D}^1$, то выражение $C(t)u^0+S(t)u^1$

будем называть обобщенным решением задачи (1).

Укажем некоторые необходимые в дальнейшем свойства косинус-функции.

Предложение 1 (см. [5, 23, 25]). Пусть $C(t)$ – косинус-функция и $S(t)$ – ассоциированная с ней синус-функция, тогда

$$\|C(t)\| \leq M e^{\omega |t|}, \ t \in R; \tag{5}$$

$$S(t+h) + S(t-h) = 2 S(t) C(h); \tag{6}$$

если $x \in \mathcal{D}(A)$, то $C(t)x \in \mathcal{D}(A)$ и

$$C''(t)x = AC(t)x = C(t)Ax, \ t \in R; \tag{7}$$

$$C(t)^2 - AS(t)^2 = 1; \tag{8}$$

$$\lambda \mathcal{R}(\lambda^2; A)x = \int_0^\infty e^{-\lambda t} C(t) x\, dt, \ x \in E, \ \mathrm{Re}\lambda > \omega, \tag{9}$$

$$\mathcal{R}(\lambda^2; A)x = \int_0^\infty e^{-\lambda t} S(t) x\, dt, \ x \in E, \ \mathrm{Re}\lambda > \omega; \tag{10}$$

Наконец, укажем аналогичный теореме Хилле—Иосиды—Филлипса-Миядеры результат, полученный в работах Сова [23], Да Прато и Джусти [14] и Фатторини [16].

Теорема 1. Замкнутый, плотно определенный оператор A в E является инфинитезимальным генератором косинус-функции $C(t), t \in R$ тогда и только тогда, когда резольвента $\mathcal{R}(\lambda^2; A)$ существует при $\lambda > \omega$ и

$$\left\| \frac{d^m}{d\lambda^m}\left(\lambda \mathcal{R}(\lambda^2; A)\right) \right\| \leq \frac{M\, m!}{(\lambda-\omega)^{m+1}}, \ \lambda > \omega, \ m = 0, 1, 2, \dots \tag{11}$$

Замечание 1. Как показано в [15, 18, 21] из (11) следует, что оператор A генерирует аналитическую полугруппу e^{-tA}. В тоже время, как видно из примеров [18, 21], обратное утверждение не верно.

§ 2. Полудискретная аппроксимация

Рассмотрим в банаховых пространствах E_n задачи Коши

$$\frac{d^2}{dt^2} u_n(t) = A_n u_n(t), \ u_n(0) = u_n^0, \ u_n'(0) = u_n^1, \tag{12}$$

где операторы A_n аналогичны по свойствам оператору A, а именно, A_n – операторы, генерирующие косинус-функции.

Определение 1. Говорят, что $C_n(t) \longrightarrow C(t)$ равномерно на $t \in [0,T]$, если

$$\max_{t\in[0,T]} \| C_n(t)x_n - p_n C(t)x \| \to 0$$

как только $x_n \to x$.

Введем следующие условия:

(A) Существует число ω такое, что
$$\mathcal{R}(\lambda^2; A_n) \to \mathcal{R}(\lambda^2; A) \text{ при } \operatorname{Re}\lambda > \omega;$$

(B) Существуют константы M и ω такие, что
$$\| C_n(t) \| \leq M e^{\omega t}, \; t \in R_+;$$

(C) Для любого конечного $T \geq 0$ имеем
$$C_n(t) \to C(t) \text{ равномерно на } t \in [0,T];$$

(A$'$) Существует $\lambda \in \bigcap_n \varrho(A_n) \cap \varrho(A)$ такое, что
$$\mathcal{R}(\lambda; A_n) \to \mathcal{R}(\lambda; A);$$

(B$'$) Существуют константы M и ω такие, что
$$\left\| \frac{d^m}{d\lambda^m}\left(\lambda \mathcal{R}(\lambda^2; A_n)\right) \right\| \leq \frac{M\, m!}{(\lambda-\omega)^{m+1}}, \; \lambda > \omega, m = 0,1,\dots; \tag{13}$$

(C$'$) Для любого конечного $T \geq 0$ имеем
$$S_n(t) \to S(t) \text{ равномерно по } t \in [0,T].$$

Замечание 2. Условия (B) и (B$'$) эквивалентны.

Действительно, импликация $(B) \Rightarrow (B')$ следует из соотношений

$$\left\| \frac{d^m}{d\lambda^m}\left(\lambda \mathcal{R}(\lambda^2; A_n)\right) \right\| = \left\| \frac{d^m}{d\lambda^m} \int_0^\infty e^{-\lambda t} C_n(t)\,dt \right\| = \left\| \int_0^\infty e^{-\lambda t}(-t)^m C_n(t)\,dt \right\|$$
$$\leq M \int_0^\infty e^{-(\lambda-\omega)t} t^m\,dt = \frac{M\,m!}{2}\left(\frac{1}{(\lambda+\omega)^{m+1}} + \frac{1}{(\lambda-\omega)^{m+1}} \right).$$

Для доказательства обратного утверждения воспользуемся формулой обращения Уиддера—Поста (см. [12], стр. 241). А именно, пусть

$$F(f)(\lambda) = \int_0^\infty e^{-\lambda t} f(t)\,dt, \; \operatorname{Re}\lambda > \omega,$$

где ω - тип оператор-функции f, тогда

$$\lim_{m\to\infty} \frac{(-1)^m}{m!}\left(\frac{m}{t}\right)^{m+1} F(f)^{(m)}\left(\frac{m}{t}\right) = f(t) \tag{14}$$

равномерно на каждом компакте $t \in [\delta, T]$, $0 < \delta < T$. Учитывая (9), и воспользовавшись выражением (14), находим

$$C_n(t) = \lim_{m\to\infty} \frac{(-1)^m}{m!}\left(\frac{m}{t}\right)^{m+1} \frac{d^m}{d\lambda^m}\left(\lambda \mathcal{R}(\lambda^2; A_n)\right)\Big|_{\lambda = m/t} \tag{15}$$

равномерно по компакту $[\delta, T]$. Записав $(m/t)^{m+1} = (1-\omega t/m)^{-(m+1)}(m/t-\omega)^{m+1}$, мы можем правую часть в (15) переписать в виде

$$\lim_{m\to\infty} \frac{(-1)^m}{m!}\left(1-\frac{\omega t}{m}\right)^{-(m+1)} \mathcal{P}_{m,n,t},$$

где

$$\mathcal{P}_{m,n,t} = \frac{(m/t-\omega)^{m+1}}{m!} \frac{d^m}{d\lambda^m}\left(\lambda \mathcal{R}(\lambda^2, A_n)\right)\Big|_{\lambda=m/t}.$$

Из оценок (11) следует, что $\mathcal{P}_{m,n,t}$ ограничена на компакте $[\delta,T]$. В силу равномерной сходимости $(1-\omega t/m)^{-(m+1)} \to e^{\omega t}$ на $[\delta,T]$ устанавливаем теперь, что $\|C_n(t)\| \le M$ при $t \in [\delta,T]$. Распространение последней оценки на $[0,T]$ легко получить, положив в уравнении Даламбера (2) $h \to t$, где $t \in [\delta,T]$ и $t+h \in [\delta,T]$. Доказательство справедливости утверждения (B) можно завершить, например, так. Пусть M и ω такие константы, что

$$\|C_n(t)\| \le M e^{\omega t} \quad \text{и} \quad 2\|C_n(1)\| e^{-\omega} + e^{-2\omega} \le 1 \quad \text{при } t \in [0,1].$$

Предположим, что (B) имеет место при $t \le \ell$, тогда

$$\|C_n(t+1)\| \le 2\|C_n(t)\|\|C_n(1)\| + \|C_n(t-1)\| \le$$
$$\le 2\|C_n(1)\| M e^{\omega t} + M e^{\omega(t-1)} \le M e^{\omega(t+1)}.$$

Итак, оценка в утверждении (B) справедлива при $t \in R_+$. Следующий результат аналогичен теореме Троттера—Като для полугрупп операторов.

Теорема (ABC). Условия (A) и (B) эквивалентны условию (C).

Доказательство. Справедливость импликации (C)$\Rightarrow$(A),(B) устанавливается с учетом выражения типа (9) и теоремы Лебега о мажорантной сходимости. Для доказательства обратного утверждения воспользуемся представлением $(\mu > \omega)$

$$C(t)x = \frac{1}{2\pi i} \int_{\mu-i\infty}^{\mu+i\infty} e^{\lambda t} \frac{\mathcal{R}(\lambda^2; A)}{\lambda} Ax\, d\lambda + Ix \tag{16}$$

справедливым, когда $x \in \mathcal{D}(A)$. Используя эквивалентность утверждений (B) и (B′), а также выражения типа (9), находим, что $\mathcal{R}(\lambda^2; A_n) \to \mathcal{R}(\lambda^2; A)$ для каждого λ из контура интегрирования в (16). Условие (A) позволяет для любого $x \in \mathcal{D}(A)$ выбрать $x_n \in \mathcal{D}(A_n)$ со свойством $x_n \to x$ и $A_n x_n \to Ax$, поэтому $C_n(t)x_n \to C(t)x$ для любого $t \in [0,T]$, где $T>0$ – произвольное фиксированное число. Пусть теперь найдутся такие $x \in E$ и последовательности $\{t_n\} \subset [0,T]$, $\{x_n\}$, $x_n \in E_n$, что $x_n \to x$, но $\|C_n(t_n)x_n - p_n C(t_n)x\| \ge \varepsilon > 0$, $n \to \infty$. (17)

К противоречию с (17) приходим из следующей цепочки неравенств

$$\|C_n(t_n)x_n - p_n C(t_n)x\| \le \|C_n(t_n)(x_n - y_n)\| + \|C_n(t_n)y_n - C_n(t_0)y_n\| +$$
$$+ \|C_n(t_0)y_n - p_n C(t_0)y\| + \|p_n(C(t_0) - C(t_n))y\| + \|p_n(C(t_n)y - C(t_n)x)\|, \quad (18)$$

где $y \in \mathcal{D}(A)$, $\{y_n\}$, $y_n \in \mathcal{D}(A_n)$ со свойством $y_n \to y$ и $A_n y_n \to Ay$, а $t_n \to t_0$. Выбрав величину $\|y - x\|$ достаточно малой, получим, начиная с некоторого n, что правая часть в (18) меньше ε. Теорема доказана.

Теорема АВС в случае $E = E_n$ и $p_n = I$ другими методами доказана в [19, 20].

Замечание 3. Условие (А) в теореме АВС можно заменить на (А′). Действительно, представление (9) позволяет установить ограниченность $\mathcal{R}(\lambda^2; A_n)$ при λ, изменяющемся по контуру интегрирования в (16). Условие (А) теперь следует из соотношений $\mathcal{R}(\lambda^2; A_n)p_n x - p_n \mathcal{R}(\lambda^2; A)x = \mathcal{R}(\lambda^2; A_n)[p_n A - A_n p_n]\mathcal{R}(\lambda^2; A)x =$

$$= \mathcal{R}(\lambda^2; A)(p_n Ay - A_n y_n) + \mathcal{R}(\lambda^2; A_n)(A_n y_n - A_n p_n y),$$

где $y_n \in E_n$ – элементы из условия согласованности $y_n \to y = \mathcal{R}(\lambda^2; A)x$, $(A_n - \lambda^2)y_n \to (A - \lambda^2)y = x$.

Замечание 4. При выполнении условия (В) условие (А) можно заменить на условие согласованности: для любого $x \in \mathcal{D}(A)$ существуют $x_n \in \mathcal{D}(A_n)$ такие, что $x_n \to x$ и $A_n x_n \to Ax$ (Ср. [19], теорема 2.)

Замечание 5. Очевидно, что из (С) вытекает равномерная сходимость $S_n(t) \to S(t)$, $t \in [0, T]$.

В силу теоремы АВС из условий (А) и (В) следует, что обобщенные решения задачи (12) будут сходится к обобщенному решению задачи (1), если $u_n^0 \to u^0$ и $u_n^1 \to u^1$. Для получения оценок сходимости предположим гладкость начальных условий.

Для элементов $x \in \mathcal{D}(A)$ и $x_n \in \mathcal{D}(A_n)$ определим величины

$$e_n^1(\Gamma; \alpha; x_n, x) = \sup_{\zeta \in \Gamma} \|\mathcal{R}(\zeta^2; A_n)A_n x_n - p_n \mathcal{R}(\zeta^2; A)Ax\| \, |\zeta|^{\alpha}, \quad (19)$$

$$e_n^0(\Gamma; x_n, x) = \sup_{\zeta \in \Gamma} \|\mathcal{R}(\zeta^2; A_n)A_n x_n - p_n \mathcal{R}(\zeta^2; A)Ax\|, \quad (20)$$

где $\Gamma = \{\zeta : \zeta = \mu + i\nu, -\infty < \nu < \infty, \mu > \omega\}$ и $0 < \alpha < 1$.

Предложение 2. Пусть выполнены условия (A) и (B). Пусть для $x \in \mathcal{D}(A)$ и $x_n \in \mathcal{D}(A_n)$ выполняется $x_n \to x$ и $A_n x_n \to Ax$, тогда $e_n^0(\Gamma; x_n, x) \to 0$, $e_n^1(\Gamma; \alpha; x_n, x) \to 0$ при любом $0<\alpha<1$.

Доказательство проведем от противного. Пусть $e_n^1(\Gamma; \alpha; x_n, x) \geq \varepsilon > 0$ при некотором $0<\alpha<1$. Тогда из определения (19) следует, что найдется такая последовательность $\{\zeta_n\}$, что

$$\|\mathcal{R}(\zeta_n^2; A_n)A_n x_n - p_n\mathcal{R}(\zeta_n^2; A)Ax\| \cdot |\zeta_n|^\alpha \geq \varepsilon > 0 . \qquad (21)$$

Предположим, что $\zeta_n \to \zeta_0$, тогда для получения противоречия воспользуемся неравенством

$$\|\mathcal{R}(\zeta_n^2; A_n)A_n x_n - p_n\mathcal{R}(\zeta_n^2; A)Ax\| \leq \|\mathcal{R}(\zeta_n^2; A_n)A_n x_n - \mathcal{R}(\zeta_n^2; A_n)p_n Ax\|$$
$$+\|\mathcal{R}(\zeta_n^2; A_n)p_n Ax - \mathcal{R}(\zeta_0^2; A_n)p_n Ax\| + \|\mathcal{R}(\zeta_0^2; A_n)p_n Ax - p_n\mathcal{R}(\zeta_0^2; A)Ax\| +$$
$$+\|p_n(\mathcal{R}(\zeta_0^2; A)Ax - \mathcal{R}(\zeta_n^2; A)Ax)\|$$

и для второго и четвертого члена справа тождеством Гильберта. Пусть теперь $|\zeta_n| \to \infty$. В этом случае, используя оценку типа (9), приходим к противоречию с (21). Первое утверждение доказывается аналогично. Предложение доказано.

Обозначим

$$\varepsilon_n = $$
$$= \|u_n^0 - p_n u^0\| + t\|u_n^1 - p_n u^1\| + c_1 e^{\omega t} e_n^1(\Gamma; \alpha; u_n^0, u^0) + c_2 e^{\omega t} e_n^0(\Gamma; u_n^1, u^1).$$

Теорема 2. Пусть $u^0, u^1 \in \mathcal{D}(A)$ и $u_n^0, u_n^1 \in \mathcal{D}(A_n)$. Тогда для разности точных решений задач (1) и (12) справедливы следующие оценки:

$$\|u_n(t) - p_n u(t)\| \leq \varepsilon_n . \qquad (22)$$

Доказательство. Используя для (10) формулу обращения, находим аналогично (16), что

$$S(t)x = \frac{1}{2\pi i}\int_{\mu-i\infty}^{\mu+i\infty} e^{\lambda t}\mathcal{R}(\lambda^2; A)x\,d\lambda =$$

$$= \frac{1}{2\pi i}\int_{\mu-i\infty}^{\mu+i\infty} e^{\lambda t}\mathcal{R}(\lambda^2; A)\frac{Ax}{\lambda^2}\,d\lambda + \frac{1}{2\pi i}\int_{\mu-i\infty}^{\mu+i\infty} e^{\lambda t}\frac{d\lambda}{\lambda^2}\,x,$$

при $x \in \mathcal{D}(A)$. Последний интеграл легко вычисляется и равен tIx. Учитывая последнее выражение, а также (16), получим

$$u_n(t)-p_n u(t) = u_n^0 - p_n u^0 + t(u_n^1 - p_n u^1) + \\ + \frac{1}{2\pi i}\int_{\mu-i\infty}^{\mu+i\infty} e^{\lambda t}\frac{R(\lambda^2;A_n)A_n u_n^0 - p_n R(\lambda^2;A)Au^0}{\lambda}\,d\lambda + \\ + \frac{1}{2\pi i}\int_{\mu-i\infty}^{\mu+i\infty} e^{\lambda t}\left(R(\lambda^2;A_n)A_n u_n^1 - p_n R(\lambda^2;A)Au^1\right)\frac{1}{\lambda^2}\,d\lambda .$$

Отсюда легко установить оценку (22). Теорема доказана.

§ 3. Дискретная косинус и синус-функция и полная дискретизация задачи

1. Решение задач (12) можно проводить, дискретизируя временную переменную, в этом случае говорят о полной дискретизации задачи (1).

Определение 2. Множество ограниченных операторов $\{K(n)\}_{n=-\infty}^{\infty}$ в пространстве E называют дискретной косинус-функцией, если выполняются следующие условия

$$K(m+n)+K(m-n)=2K(m)K(n),\quad K(0)=I,\quad m,n=0,\pm1,\dots \tag{23}$$

Оператор $K(1)$ назовем ведущим оператором дискретной косинус-функции.

Как и в случае дискретных полугрупп (см. [5], стр. 629) с величиной $K(1)$ мы связываем шаг по времени τ ; пишем $K(1)=\mathcal{C}(\tau)$ и $K(m)=\mathcal{C}(m\tau)$. Примером дискретный косинус-функции является сужение косинус-функции на сетку $\{k\tau\}_{-\infty}^{\infty}$

Замечание 7. Пусть задан произвольный ограниченный оператор $K=\mathcal{C}(\tau)$ в E . Вычислим $\mathcal{C}(n\tau)$ по рекуррентному соотношению

$$\mathcal{C}((n+1)\tau)=2\mathcal{C}(n\tau)\mathcal{C}(\tau)-\mathcal{C}((n-1)\tau),\quad \mathcal{C}(0)=I. \tag{24}$$

Оказывается, что тогда $\{\mathcal{C}(n\tau)\}$ - дискретная косинус-функция.

Определение 3. Оператор

$$\mathcal{A}=2(\mathcal{C}(\tau)-1)/\tau^2 \tag{25}$$

называется производящим оператором (генератором) дискретной косинус-функции.

Соотношение (25) позволяет по оператору $\mathcal{A}$ строить ведущий оператор $\mathcal{C}(\tau)=1+\tau^2\mathcal{A}/2$ дискретной косинус-функции, а следовательно и дискретную косинус-функцию.

Можно рассматривать целое семейство дискретных косинус-функций $\{\mathcal{C}_n(k\tau_n)\}$, $\mathcal{C}_n(k\tau_n): E_n \longrightarrow E_n$ с шагами по времени $\tau_n \longrightarrow 0$ и производящими операторами $\mathcal{O}\!\iota_n$. Для удобства определим $\mathcal{C}_n(t) = \mathcal{C}_n([t/\tau_n]\tau_n)$, t - любое вещественное число, $[t]$ - целая часть числа t.

Определение 4. Говорят, что последовательность $\{\mathcal{C}_n(k\tau_n)\}$ аппроксимирует косинус-функцию $C(t)$ в точке t_0, если

$$\mathcal{C}_n(k_n\tau_n) \longrightarrow C(t_0) \tag{26}$$

для любой последовательности $\{k_n\}$ целых чисел со свойством $k\tau_n \longrightarrow t_0$. Если (26) имеет место для любого $t \in [0,T]$, то будем говорить, что $\mathcal{C}_n(t)$ аппроксимирует $C(t)$ на отрезке $[0,T]$.

Лемма 1. Если $\{\mathcal{C}_n(t)\}$ аппроксимирует $C(t)$ на какомлибо отрезке $[0,T]$, то $\{\mathcal{C}_n(t)\}$ аппрокимирует $C(t)$ и на $(-\infty,\infty)$; пишем $\mathcal{C}_n(t) \longrightarrow C(t)$.

Доказательство. Для произвольного $h \in [0,T]$ имеем $\mathcal{C}_n(T) \to C(T)$, $\mathcal{C}_n(T-h) \to C(T-h)$, $\mathcal{C}_n(h) \longrightarrow C(h)$, поэтому в силу соотношений (2) и (23) получим $\mathcal{C}_n(T+h) \longrightarrow$ $\to C(T+h)$. Тем самым сходимость доказана на $[0,2T]$ и т.д. Лемма доказана.

Лемма 2. Пусть $\mathcal{C}_n(t) \longrightarrow C(t)$, тогда

$$\|\mathcal{C}_n(t)\| \leq M e^{\omega|t|}, \quad t \in (-\infty,\infty). \tag{27}$$

Доказательство. Во-первых, из условия леммы следует, что $\|\mathcal{C}_n(t)\| \leq M_1$, $t \in [0,1]$. Действительно, пусть $\|\mathcal{C}_n(t_n)x_n\| \to \infty$ при $\|x_n\| = 1$, $t_n \in [0,1]$. Тогда на сходящихся последовательностях $t_n \to t_0$ и $y_n = x_n/\|\mathcal{C}_n(t_n)x_n\| \to 0$ имеем

$$\mathcal{C}_n(t_n)y_n \longrightarrow C(t_0)\theta = \theta.$$

в противоречии с $\|\mathcal{C}_n(t_n)y_n\| = 1$. Во-вторых, пусть ω такое число, что $2\|\mathcal{C}_n(1)\|e^{-\omega} + 2e^{-2\omega} \leq 1$. Тогда

$$\|\mathcal{C}_n(t+1)\| \leq 2\|\mathcal{C}_n(t)\|\|\mathcal{C}_n(1)\| + \|\mathcal{C}_n(t-1)\| \leq$$
$$\leq 2\|\mathcal{C}_n(1)\| M_1 e^{\omega t} + M_1 e^{\omega(t-1)} \leq M_1 e^{\omega(t+1)}.$$

Утверждение леммы для $t \geq 0$ теперь следует по индукции; для $t < 0$ следует из равенства $\mathcal{C}_n(t) = \mathcal{C}_n(-t)$. Лемма доказана.

Лемма 3. Для того, чтобы $\mathcal{C}_n(t) \longrightarrow C(t)$ необходимо, чтобы $\mathscr{C}_n(t) \longrightarrow C(t)$ равномерно на любом конечном отрезке $[0,T]$ и достаточно, чтобы это имело место для какого-нибудь отрезка $[0,\delta]$.

Доказательство. Необходимость доказывается от противного с помощью неравенства

$$\|\mathscr{C}_n(t_n)x_n - p_n C(t_n)x\| \le \|\mathscr{C}_n(t_n)x_n - p_n C(t_0)x\| + \|p_n(C(t_n) - C(t_0))x\|.$$

Достаточность устанавливается с использованием леммы 1. Лемма доказана.

Ввиду (6) заданной дискретной косинус-функции $\{K(m)\}$ естественно сопоставить дискретную синус-функцию по формуле

$$S(m+n) + S(m-n) = 2S(m)K(n),\ S(0) = 0,\ m,n = 0,\pm 1,\ldots \quad (28)$$

Оператор $S(1)$ назовем ведущим оператором дискретной синус-функции. Как и выше будем писать $S(1) = \mathscr{S}(\tau)$ и $S(m) = \mathscr{S}(m\tau)$.

Замечание 8. Пусть задана дискретная косинус-функция $\{\mathscr{C}(m\tau)\}$. Взяв ведущий оператор $\mathscr{S}(\tau)$ произвольным, вычислим $\{\mathscr{S}(m\tau)\}$ по рекуррентному соотношению

$$\mathscr{S}((m+1)\tau) = 2\mathscr{S}(m\tau)\mathscr{C}(\tau) - \mathscr{S}((m-1)\tau),\ \mathscr{S}(0) = 0. \quad (29)$$

Можно доказать, что построенное множество операторов $\{\mathscr{S}(m\tau)\}$ удовлетворяет соотношению (28), т.е. является дискретной синус-функцией. В этом случае будем говорить, что дискретная синус-функция $\{\mathscr{S}(m\tau)\}$ ассоциирована с дискретной косинус-функцией $\{\mathscr{C}(m\tau)\}$.

Замечание 9. В формулировки лемм 1 – 3 можнодобавить соответствующее утверждения о сходимости и ограниченности дискретных синус-функций.

Решение рекуррентного соотношения (24), соответственно (29), имеет вид (см. [13], стр. 37)

$$\mathscr{C}(k\tau) = \alpha\, w(\tau)^k + \beta\, v(\tau)^k, \quad (30)$$

$$\mathscr{S}(k\tau) = \gamma\, w(\tau)^k + \eta\, v(\tau)^k, \quad (31)$$

где $w(\tau) = \mathscr{C}(\tau) + (\mathscr{C}(\tau)^2 - 1)^{1/2}$, $v(\tau) = \mathscr{C}(\tau) - (\mathscr{C}(\tau)^2 - 1)^{1/2}$, а коэффициенты α, β, γ и η выбираются из условий $\alpha + \beta = I$, $\gamma + \eta = 0$, $\alpha w(\tau) + \beta v(\tau) = \mathscr{C}(\tau)$, $\gamma w(\tau) + \eta v(\tau) = \mathscr{S}(\tau)$. Учитывая определение производящего оператора дискретной косинус-функции, находим

$$\mathscr{C}(k\tau) = (w(\tau)^k + v(\tau)^k)/2, \quad (32)$$

$$\mathscr{S}(k\tau) = \mathscr{S}(\tau)(v(\tau)^k - w(\tau)^k)/[2(\mathscr{C}(\tau)^2 - 1)^{1/2}]. \quad (33)$$

Замечание 10. Величины $\mathcal{W}(\tau)$ и $\mathcal{V}(\tau)$ имеют смысл, когда существует оператор $(\mathcal{C}(\tau)^2-1)^{1/2}$. В рассматриваемой далее ситуации, указанный оператор, как легко видеть, определен.

Замечание 11. В силу (8) и представления (33), если существует $\mathcal{A}^{-1/2}$, естественно положить $\mathcal{S}(\tau)=(\mathcal{C}(\tau)^2-1)^{1/2}\mathcal{A}^{-1/2}=(\tau^2+\tau^4\mathcal{A}/4)^{1/2}$. На протяжении следующего пункта считаем, что сделан именно такой выбор ведущего оператора для дискретной синус-функции.

Корень $\mathcal{A}^{1/2}$ можно, например, определить по формуле $\mathcal{A}^{1/2}=i(-\mathcal{A})^{1/2}$, где $(-\mathcal{A})^{1/2}$ определен естественным образом, если $\mathcal{A}$ генерирует аналитическую полугруппу. (См. замечание 1 и [4, 6]).

2. Понятие дискретной косинус и синус-функций тесно связано с задачей

$$\mathcal{U}_n(k\tau_n+\tau_n)-2\mathcal{U}_n(k\tau_n)+\mathcal{U}_n(k\tau_n-\tau_n)=\tau_n^2 A_n\mathcal{U}_n(k\tau_n),$$

$$\mathcal{U}_n(0)=u_n^0,\ \mathcal{U}_n(\tau_n)=\mathcal{C}_n(\tau_n)u_n^0+\mathcal{S}_n(\tau_n)u_n^1, \qquad (34)$$

аппроксимирующей задачу (12). В самом деле, в силу линейности задачи (34), ее решение дается формулой

$$\mathcal{U}_n(k\tau_n)=\mathcal{C}_n(k\tau_n)u_n^0+\mathcal{S}_n(k\tau_n)u_n^1,\quad k=0,\pm1,\ldots$$

До конца параграфа предполагаем, что спектры операторов A и A_h лежат в левой полуплоскости, так что в (11) можно считать $\omega<0$. Более того, будем предполагать, что E и E_h – гильбертовы пространства, а операторы $-A$ и $-A_h$ самосопряженные положительно определенные. Считаем также, что A_h ограниченный оператор при каждом h. Для удобства записи введем обозначения

$$\mathcal{W}(\tau_n,\lambda)=1+\tau_n^2\lambda/2+\left(\tau_n^2\lambda+\frac{\tau_n^4\lambda^2}{4}\right)^{1/2},\ \mathcal{V}(\tau_n,\lambda)=1+\frac{\tau_n^2\lambda}{2}-\left(\tau_n^2\lambda+\frac{\tau_n^4\lambda^2}{4}\right)^{1/2},$$

$$\mathcal{M}_n(\tau_n)=C_n(\tau_n)+(C_n(\tau_n)^2-1)^{1/2},\ \mathcal{N}_n(\tau_n)=C_n(\tau_n)-(C_n(\tau_n)^2-1)^{1/2}.$$

Теорема 3. Пусть $\|\tau_n^2A_n\|\le\delta<4$, тогда для дискретной синус-функции и косинус-функции с генератором A_h имеют место оценки:

$$\|\mathcal{C}_n(k\tau_n)\|\le1,\quad \|\mathcal{S}_n(k\tau_n)\|\le1/|\omega|^{1/2}. \qquad (35)$$

Доказательство использует хорошо известный факт о совпадении нормы и спектрального радиуса для самосопряженного оператора (см. (32), (33))

$$\|\mathcal{C}_n(k\tau_n)\| = \max_{\lambda\in\sigma(A_n)} \frac{1}{2}\left| \mathcal{W}(\tau_n,\lambda)^k + \mathcal{V}(\tau_n,\lambda)^k \right|,$$

$$\|\mathcal{S}_n(k\tau_n)\| = \max_{\lambda\in\sigma(A_n)} \frac{1}{2|\lambda|^{1/2}} \left| \mathcal{V}(\tau_n,\lambda)^k - \mathcal{W}(\tau_n,\lambda)^k \right|$$

и условие $\sigma(A_n)\subset(-\infty,\omega)$. Теорема доказана.

Данный результат устанавливает устойчивость метода (34). Напомним, что в случае ограниченного оператора A_n, косинус и синус-функции задаются по формулам

$$C_n(t)=\sum_{k=0}^{\infty} t^{2k}A_n^k/(2k)!\,,\quad S_n(t)=\sum_{k=0}^{\infty} t^{2k+1}A_n^k/(2k+1)!\,. \tag{36}$$

Предложение 3. Пусть выполнено условие устойчивости

$$\|\tau_n^2A_n\|\le\delta<4,$$

причем $u_n^0\in\mathcal{D}(A_n^2),\ u_n^1\in\mathcal{D}(A_n)$, тогда

$$\|(C_n(\tau_n)-\mathcal{C}_n(\tau_n))u_n^0\|\le c_1\tau_n^4\|A_n^2u_n^0\|, \tag{37}$$

$$\|(S_n(\tau_n)-\mathcal{S}_n(\tau_n))u_n^1\|\le c_2\tau_n^3\|A_nu_n^1\|.$$

Если же начальное условие $u_n^0\in\mathcal{D}(A_n^{3/2})$, то

$$\|(C_n(\tau_n)-\mathcal{C}_n(\tau_n))u_n^0\|\le c\tau_n^3\|A_n^{3/2}u_n^0\|. \tag{38}$$

Доказательство. Утверждения (37) доказываются с использованием (36), а именно

$$\|(C_n(\tau_n)-\mathcal{C}_n(\tau_n))u_n^0\|=\Big\|\Big(\sum_{k=0}^{\infty}\tau_n^{2k}A_n^k/(2k)!-1-\frac{\tau_n^2A_n}{2}\Big)u_n^0\Big\|\le c_1\tau_n^4\|A_n^2u_n^0\|,$$

$$\|(S_n(\tau_n)-\mathcal{S}_n(\tau_n))u_n^1\|=\max_{\lambda\in\sigma(A_n)}\Big|\frac{1}{\lambda}\sum_{k=0}^{\infty}\tau_n^{2k+1}\frac{\lambda^k}{(2k+1)!}-\frac{\tau_n}{\lambda}\Big(1+\frac{\tau_n^2\lambda}{4}\Big)^{1/2}\Big|\times$$

$$\times\|A_nu_n^1\|\le\max_{\lambda\in\sigma(A_n)}\Big|\frac{1}{\lambda}\Big(\tau_n+\frac{\tau_n^3\lambda}{3!}+\ldots\Big)-\frac{\tau_n}{\lambda}\Big(1+\frac{\tau_n^2\lambda}{8}+\ldots\Big)\Big|\,\|A_nu_n^1\|\le$$

$$\le c_2\tau_n^3\|A_nu_n^1\|.$$

Предложение доказано.

Теорема 4. Пусть выполнено условие устойчивости

$$\|\tau_n^2A_n\|\le\delta<4,$$

причем $u_n^0\in\mathcal{D}(A_n^{3/2}),\ u_n^1\in\mathcal{D}(A_n)$, тогда для разности решений задач (12) и (34) справедливы оценки

$$\|U_n(k\tau_n)-u_n(k\tau_n)\|\le c\tau_n^2t\big(\|A_n^{3/2}u_n^0\|+\|A_nu_n^1\|\big),\ t=k\tau_n. \tag{39}$$

Доказательство. В силу того, что $C_n(t)$ и $S_n(t)$, суженные на сетку $\{k\tau_n\}_{k=-\infty}^{\infty}$ удовлетворяют соответственно уравнениям (24) и (29), можно получить следующие выражения:

$$C_n(k\tau_n) = \left(\mathfrak{M}_n(\tau_n)^k + \mathfrak{N}_n(\tau_n)^k\right)/2,$$

$$S_n(k\tau_n) = A_n^{-1/2}\left(\mathfrak{N}_n(\tau_n)^k - \mathfrak{M}(\tau_n)^k\right)/2. \qquad (40)$$

Учитывая (32) и (33), находим

$$\mathcal{U}_n(k\tau_n) - u_n(k\tau_n) = \frac{1}{2}\left(\mathcal{W}_n(\tau_n)^k + \mathcal{V}_n(\tau_n)^k\right)u_n^0 + \frac{1}{2}A_n^{-1/2}\left(\mathcal{V}_n(\tau_n)^k - \mathcal{W}_n(\tau_n)^k\right)u_n^1 -$$

$$- \frac{1}{2}\left(\mathfrak{M}_n(\tau_n)^k + \mathfrak{N}_n(\tau_n)^k\right)u_n^0 - \frac{1}{2}A_n^{-1/2}\left(\mathfrak{N}_n(\tau_n)^k - \mathfrak{M}_n(\tau_n)^k\right)u_n^1.$$

Так как $\mathcal{W}_n(\tau_n)^k - \mathfrak{M}_n(\tau_n)^k = \left[\mathcal{W}_n(\tau_n) - \mathfrak{M}_n(\tau_n)\right]\sum\limits_{i=0}^{k} \mathcal{W}_n(\tau_n)^i\, \mathfrak{M}_n(\tau_n)^{k-i}$,

$$\mathcal{V}_n(\tau_n)^k - \mathfrak{N}_n(\tau_n)^k = \left[\mathcal{V}_n(\tau_n) - \mathfrak{N}_n(\tau_n)\right]\sum_{i=0}^{k} \mathcal{V}_n(\tau_n)^i\, \mathfrak{N}_n(\tau_n)^{k-i},$$

а степени $\mathcal{V}_n^j, \mathcal{W}_n^j, \mathfrak{N}_n^j, \mathfrak{M}_n^j$ ограничены, то достаточно оценить разности

$$\mathcal{W}_n(\tau_n) - \mathfrak{M}_n(\tau_n),\ \mathcal{V}_n(\tau_n) - \mathfrak{N}_n(\tau_n),\ A_n^{-1/2}\left(\mathfrak{M}_n(\tau_n) - \mathcal{W}_n(\tau_n)\right),\ A_n^{-1/2}\left(\mathcal{V}_n(\tau_n) - \mathfrak{N}_n(\tau_n)\right).$$

Имеем

$$\mathcal{W}_n(\tau_n) - \mathfrak{M}_n(\tau_n) = \mathcal{C}_n(\tau_n) + \left(\mathcal{C}_n(\tau_n)^2 - 1\right)^{1/2} - C_n(\tau_n) - \left(C_n(\tau_n)^2 - 1\right)^{1/2} =$$

$$= \mathcal{C}_n(\tau_n) - C_n(\tau_n) + A_n^{1/2}\mathcal{S}_n(\tau_n) - A_n^{1/2}S_n(\tau_n).$$

Аналогично расписываются и другие разности. Оценки (39) следуют теперь из предложения 3. Теорема доказана.

Замечание 12. Практически вычисление по рекуррентным соотношениям типа (24) и (29) проводится приближенно, а именно, например, дискретная косинус-функция $\{K_n(k\tau_n)\}$ вычисляется по формуле

$$K_n(k\tau_n + \tau_n) = 2K_n(\tau_n k)\,\mathcal{C}_n(\tau_n) - K_n(k\tau_n - \tau_n) + r_{n,k}, \qquad (41)$$

где $r_{n,k}$ – ошибка вычислений на шаге $k+1$ для операторов в пространстве E_n. Решение (41) записывается в виде (см. [11], стр. 60)

$$K_n(k\tau_n) = \frac{1}{2}\left(\mathcal{W}_n(\tau_n)^k + \mathcal{V}_n(\tau_n)^k\right) + \sum_{i=0}^{k-2} \frac{\mathcal{W}_n(\tau_n)^{k-i-1} - \mathcal{V}_n(\tau_n)^{k-i-1}}{\mathcal{W}_n(\tau_n) - \mathcal{V}_n(\tau_n)}\cdot r_{n,i}.$$

Преобразуя $\left(\mathcal{W}_n(\tau_n)^l - \mathcal{V}_n(\tau_n)^l\right)/\left(\mathcal{W}_n(\tau_n) - \mathcal{V}_n(\tau_n)\right) =$
$= \sum\limits_{i=0}^{l} \mathcal{W}_n(\tau_n)^i\,\mathcal{V}_n(\tau_n)^{l-i}$ и, учитывая ограниченность степеней

$W_n(\tau_n)^k$ и $V_n(\tau_n)^k$, находим

$$\| K_n(k\tau_n) \| \leq 1 + \frac{k(k+1)}{2} \max_{0 \leq i \leq k} \| z_{n,i} \|.$$

Ведя вычисление, например, с точностью $\| z_{n,i} \| = O(\tau_n^4)$, получим

$$\| K_n(t) \| \leq 1 + c\tau_n^2 t, \; t = k\tau_n,$$

и аналогично

$$\| S_n(t) \| \leq 1/|\omega|^{1/2} + c\tau_n^2 t, \; t = k\tau_n.$$

Теорема 5. Пусть выполнено условие

$$\| \tau_n^2 A_n \| \leq \varepsilon < 4$$

тогда при $u_n^0 \in \mathcal{D}(A_n^{3/2})$, $u_n^1 \in \mathcal{D}(A_n)$, $u^0 \in \mathcal{D}(A)$, $u^1 \in \mathcal{D}(A)$ справедливы оценки

$$\| u_n(t) - p_n u(t) \| \leq c\tau_n^2 t \left(\| A_n^{3/2} u_n^0 \| + \| A_n u_n^1 \| \right) + \varepsilon_n, \quad t = k\tau_n.$$

Доказательство следует из теоремы 2 и 4.

Фундаментальные исследования по устойчивости и сходимости разностных схем приведены в [7, 9, 10]. Данный подход использует метод теории полугрупп операторов.

§ 4. Другие схемы полной дискретизации задачи (1)

Будем говорить, что косинус-функция $C(t)$ с инфинитезимальным оператором A удовлетворяет в банаховом пространстве E условию (F), если

1° Существует замкнутый оператор B в E такой, что $B^2 = A$ и B коммутирует с любым оператором из $\mathcal{L}(E)$ коммутирующим с A;

2° $S(t)$ отображает E в $\mathcal{D}(B)$ при любом $t \in R$,

3° $BS(t)x$ непрерывна по $t \in R$ для любого $x \in E$.

Справедливо следующее

Предложение 4 (см. [24], стр. 12). Пусть A и B линейные операторы в E и $B^2 = A$, причем B коммутирует с каждым оператором из $\mathcal{L}(E)$, коммутирующим с A. Пусть $0 \in \rho(B)$. Тогда указанные ниже условия 1° и 2° эквивалентны.

1° Оператор A является инфинитезимальным генератором косинус-функции $C(t), t \in R$, удовлетворяющей условию (F),

2° Оператор B является инфинитезимальным генератором непрерывной группы e^{tB}, $t \in R$.

Фатторини [16] доказал, что при подходящем сдвиге $A + cI$ оператора A, каждая косинус-функция в банаховом пространстве L_p, $1<p<\infty$ удовлетворяет условию (F). Как показывают контрпримеры [24, 22], в произвольном банаховом пространстве условие (F) может не иметь места.

Замечание 13. Пусть $-A$ – положительно определенный самосопряженный оператор в гильбертовом пространстве, тогда A является производящим оператором косинус-функции, удовлетворяющей условию (F).

До конца параграфа предполагаем, что косинус-функции $C(t)$ и $C_n(t)$ удовлетворяют условию (F) и, более того, существуют такие константы M_1 и ν, что (ср. [3], стр. 669)

$$\|(\lambda-B)^{-k}\| \le M_1/(\mathrm{Re}\,\lambda-\nu)^k, \quad n,k=0,1,\dots, \quad |\mathrm{Re}\,\lambda|>\nu,$$

$$\|B_n^{-1}\| \le \mathrm{const}. \tag{42}$$

Как уже отмечалось операторы A_n генерируют аналитические полугруппы, поэтому сделав соответствующий сдвиг A_n+cI_n можно определить корень $(A_n+cI_n)^{1/2}=i\,(-(A_n+cI_n))^{1/2}$, причем (см. [6], стр. 144, 146) справедлива оценка

$$\|(-(A_n+cI_n))^{-1/2}\| \le \mathrm{const}.$$

Замечание 14. Условие (42) можно заменить на следующее эквивалентное: существует константа L такая, что

$$\|B_n S_n(t)\| \le L, \quad t\in[0,1], \quad n=1,2,\dots \tag{43}$$

Доказательство эквивалентности (42) и (43) почти дословно повторяет рассуждения на стр. 97–98 в [15], и здесь не приводится.

Рассмотрим следующую схему аппроксимации уравнения (12):

$$U_{n,\ell}(k\tau_n+\tau_n)-2U_{n,\ell}(k\tau_n)+U_{n,\ell}(k\tau_n-\tau_n)=\tau_n^2 A_n U_{n,\ell}(k\tau_n+\tau_n),$$

$$U_{n,\ell}(0)=u_n^0, \quad U_{n,\ell}(\tau_n)=(1-\tau_n^2A_n)^{-1}(u_n^0+\tau_n u_n^1). \tag{44}$$

Решение (44) дается по формуле $U_{n,\ell}(k\tau_n)=\mathcal{C}_{n,\ell}(k\tau_n)u_n^0+$ $+\mathcal{S}_{n,\ell}(k\tau_n)u_n^1$, где $\mathcal{C}_{n,\ell}$ и $\mathcal{S}_{n,\ell}$ удовлетворяют рекуррентным соотношениям

$$\mathcal{C}_{n,\ell}(k\tau_n+\tau_n)=2(1-\tau_n^2A_n)^{-1}\mathcal{C}_{n,\ell}(k\tau_n)-(1-\tau_n^2A_n)^{-1}\mathcal{C}_{n,\ell}(k\tau_n-\tau_n),$$

$$\mathcal{C}_{n,\ell}(0)=I_n, \quad \mathcal{C}_{n,\ell}(\tau_n)=(1-\tau_n^2A_n)^{-1}, \tag{45}$$

3

$$\mathcal{S}_{n,l}(k\tau_n+\tau_n) = 2(1-\tau_n^2 A_n)^{-1}\mathcal{S}_{n,l}(k\tau_n) - (1-\tau_n^2 A_n)^{-1}\mathcal{S}_{n,l}(k\tau_n-\tau_n),$$
$$\mathcal{S}_{n,l}(0)=0,\ \mathcal{S}_{n,l}(\tau_n)=\tau_n(1-\tau_n^2 A_n)^{-1} \quad (46)$$

Теорема 6. Существуют константы $M>0$ и $\eta\geqslant 0$ такие, что
$$\|\mathcal{C}_{n,l}(t)\| \leq M e^{\eta t},\ \|\mathcal{S}_{n,l}(t)\| \leq M e^{\eta t},\ t=k\tau_n. \quad (47)$$

Доказательство. Решение рекуррентных уравнений (45) и (46) можно записать в виде
$$\mathcal{C}_{n,l}(k\tau_n) = \frac{1}{2}\left[(1+\tau_n B_n)^{-k} + (1-\tau_n B_n)^{-k}\right], \quad (48)$$
$$\mathcal{S}_{n,l}(k\tau_n) = \frac{1}{2} B_n^{-1}\left[(1-\tau_n B_n)^{-k} - (1+\tau_n B_n)^{-k}\right] \quad (49)$$

Утверждение теоремы вытекает из того факта, что оператор B_n генерирует группу. Теорема доказана.

Наконец, пусть уравнение (12) аппроксимируется схемой
$$\mathcal{U}_{n,f}(k\tau_n+\tau_n) - 2\mathcal{U}_{n,f}(k\tau_n) + \mathcal{U}_{n,f}(k\tau_n-\tau_n) = \tau_n^2 A_n \mathcal{U}_{n,f}(k\tau_n-\tau_n),$$
$$\mathcal{U}_{n,f}(0)=u_n^0,\ \mathcal{U}_{n,f}(\tau_n)=(1+\tau_n^2 A_n)(u_n^0+\tau_n u_n^1). \quad (50)$$

Соответствующие рекуррентные соотношения имеют вид
$$\mathcal{C}_{n,f}(k\tau_n+\tau_n) = 2\mathcal{C}_{n,f}(k\tau_n) - (1-\tau_n^2 A_n)\mathcal{C}_{n,f}(k\tau_n-\tau_n),$$
$$\mathcal{C}_{n,f}(0)=I_n,\ \mathcal{C}_{n,f}(\tau_n)=1+\tau_n^2 A_n. \quad (51)$$

а также
$$\mathcal{S}_{n,f}(k\tau_n+\tau_n) = 2\mathcal{S}_{n,f}(k\tau_n) - (1-\tau_n^2 A_n)\mathcal{S}_{n,f}(k\tau_n-\tau_n),$$
$$\mathcal{S}_{n,f}(0)=0,\ \mathcal{S}_{n,f}(\tau_n)=\tau_n(1+\tau_n^2 A_n). \quad (52)$$

Как и в предыдущем случае решение (50) дается по формуле
$$\mathcal{U}_{n,f}(k\tau_n) = \mathcal{C}_{n,f}(k\tau_n)u_n^0 + \mathcal{S}_{n,f}(k\tau_n)u_n^1.$$

Теорема 7. Пусть выполнено условие устойчивости
$$\tau_n\|A_n\| \leq const, \quad (53)$$

тогда с некоторыми константами $M>0$ и $\eta>0$ справедливы оценки
$$\|\mathcal{C}_{n,f}(k\tau_n)\| \leq M e^{\eta t},\ \|\mathcal{S}_{n,f}(k\tau_n)\| \leq M e^{\eta t},\ t=k\tau_n. \quad (54)$$

3т

Доказательство. Аналогично теореме 5 для (51) и (52) имеем

$$\mathcal{C}_{n,f}(k\tau_n) = \frac{1}{2}\left[(1+\tau_n B_n)^{k+1} + (1-\tau_n B_n)^{k+1}\right], \tag{55}$$

$$\mathcal{S}_{n,f}(k\tau_n) = \frac{1}{2}\left[(1+\tau_n^2 A_n) B_n^{-1}\left((1+\tau_n B)^k - (1-\tau_n B_n)^k\right)\right] \tag{56}$$

Учитывая тождества $1 \pm \tau_n B_n = (1-\tau_n^2 A_n)/(1 \mp \tau_n B_n)$, неравенства $\|(1-\tau_n^2 A_n)^k\| \le (1+\tau_n^2\|A_n\|)^k \le e^{\tau_n k\tau_n \|A_n\|}$ и рассуждения, проведенные в доказательстве теоремы 5, устанавливаем нужные оценки. Теорема доказана.

Теоремы 5 и 7 дают устойчивость методов (44) и (50). В гильбертовом пространстве для самосопряженных положительно-определенных операторов для схемы (50) условие устойчивости (53) без оценок скорости сходимости получено другим методом в работе [26]. Отметим также, что данный подход позволяет получить аналог дискретного варианта ABC теоремы (см. [26], стр. 193) т.е. доказать сходимость обобщенных решений задач (44) и (50) к обобщенному решению задачи (1).

Лемма 4. Пусть $u_n^0 \in \mathcal{D}(A_n)$, тогда

$$\left\|\left((1\pm\tau_n B_n)^{-1} - e^{\mp\tau_n B_n}\right)u_n^0\right\| \le C\tau_n^2 \|A_n u_n^0\|. \tag{57}$$

Если же дополнительно выполнено условие устойчивости (53), то

$$\left\|\left((1\pm\tau_n B_n) - e^{\pm\tau_n B_n}\right)u_n^0\right\| \le C\tau_n^2 \|A_n u_n^0\|. \tag{58}$$

Доказательство. Воспользуется равенством

$$(1 \pm \tau_n B_n)^{-1} = (1 \mp \tau_n B_n)/(1-\tau_n^2 A_n)$$

Имеем

$$\left[(1\pm\tau_n B_n)^{-1} - e^{\mp\tau_n B_n}\right]u_n^0 = (1-\tau_n^2 A_n)^{-1}\left[(1\mp\tau_n B_n) - (1-\tau_n^2 A_n)e^{\mp\tau_n B_n}\right]u_n^0 =$$
$$= (1-\tau_n^2 A_n)^{-1}\left[\tau_n^2 e^{\mp\tau_n B_n} A_n u_n^0 + \left(1\mp\tau_n B_n - e^{\mp\tau_n B_n}\right)u_n^0\right].$$

Доказательство (57) завершается использованием тождества (см. [5], стр. 630)

$$\left(1\pm\tau_n B_n - e^{\pm\tau_n B_n}\right)u_n^0 = -\int_0^{\tau_n} (\tau_n - s)\, e^{\pm s B_n} A_n u_n^0\, ds.$$

Для установления (58) достаточно разложить $e^{\pm\tau_n B_n}$ в ряд и воспользоваться неравенством $\|\tau_n B_n\|\le\|\tau_n A_n\|\,\|B_n^{-1}\|\le const$. Лемма доказана.

Напомним, что в наших условиях (см. предложение 4)

$$C_n(t)=\frac{1}{2}\left(e^{-tB_n}+e^{tB_n}\right),\quad S_n(t)=\frac{1}{2}B_n^{-1}\left(e^{tB_n}-e^{-tB_n}\right). \tag{59}$$

Теорема 8. Пусть $u_n^0\in\mathcal{D}(A_n)$ и $u_n^1\in\mathcal{D}(B_n)$. Тогда для разности решений задач (12) и (44), а в случае выполнения условия устойчивости (53) и для разности решений задач (12) и (50), справедливы оценки

$$\|\mathcal{U}_{n,f}(t)-u_n(t)\|\le\tau_n t M e^{\eta t}\left(\|A_n u_n^0\|+\|B_n u_n^1\|\right),\ t=k\tau_n; \tag{60}$$

$$\|\mathcal{U}_{n,\ell}(t)-u_n(t)\|\le\tau_n t M e^{\eta t}\left(\|A_n u_n^0\|+\|B_n u_n^1\|\right),\ t=k\tau_n. \tag{61}$$

Доказательство. Используя выражения (59) и (48), (49) соответственно (55), (56), находим $(t=k\tau_n)$

$$\mathcal{U}_{n,f}(k\tau_n)-u_n(k\tau_n)=\frac{1}{2}\left[(1+\tau_nB_n)^{-k}+(1-\tau_nB_n)^{-k}\right]u_n^0+\frac{1}{2}\left[(1-\tau_nB_n)^{-k}-\right.$$

$$\left.-(1+\tau_nB_n)^{-k}\right]B_n^{-1}u_n^1-\frac{1}{2}\left(e^{-tB_n}+e^{tB_n}\right)u_n^0-\frac{1}{2}\left(e^{tB_n}-e^{-tB_n}\right)B_n^{-1}u_n^1,$$

$$\mathcal{U}_{n,\ell}(t)-u_n(t)=\frac{1}{2}\left[(1+\tau_nB_n)^{k+1}-(1-\tau_nB_n)^{k+1}\right]u_n^0+\frac{1}{2}\left[(1+\tau_n^2A_n)(1+\tau_nB_n)^{k}-\right.$$

$$\left.-(1+\tau_n^2A_n)(1-\tau_nB_n)^{k}\right]B_n^{-1}u_n^1-\frac{1}{2}\left(e^{-tB_n}+e^{tB_n}\right)u_n^0-\frac{B_n^{-1}}{2}\left(e^{tB_n}-e^{-tB_n}\right)u_n^1.$$

Выпишем равенства для выражений, к оценке которых сводится доказательство (60) и (61)

$$(1\pm\tau_nB_n)^{-k}-e^{\mp\tau_nkB_n}=\left[(1\pm\tau_nB_n)^{-1}-e^{\pm\tau_nB_n}\right]\sum_{i=0}^{k}(1\pm\tau_nB_n)^{i}\left(e^{\mp\tau_nB_n}\right)^{k-i},$$

$$(1\pm\tau_nB_n)^{k+1}-e^{\pm\tau_nkB_n}=\left[(1\pm\tau_nB_n)-e^{\pm\tau_nB_n}\right]\sum_{i=0}^{k}(1\pm\tau_nB_n)^{i}\left(e^{\mp\tau_nB_n}\right)^{k-i}\pm$$

$$\pm\tau_n(1+\tau_nB_n)^{k}\,\frac{B_n\cdot B_n}{B_n}.$$

Учитывая экспоненциальную ограниченность степеней операторов $(1\pm\tau_nB_n)^{\pm i}$ и $e^{\pm\tau_nB_n i}$, а также лемму 4, получаем доказываемые неравенства. Теорема доказана.

Схема типа (44) с начальными условиями $\mathcal{U}_n(0)=u_n^0$, $\mathcal{U}_n(\tau_n)=u_n^0+\tau_nu_n^1$ и оператором $A_n(t)$, зависящим от времени, подробно изучена в [11a].

Замечание 15. Рассуждения замечания 2 можно провести и для схем (55), (56). Здесь точность вычислений должна быть порядка $O(\tau_n^3)$.

Положим

$$\varepsilon_{n,1} = \varepsilon_n + \tau_n t M e^{\eta t}\left(\|A_n u_n^0\| + \|A_n u_n^1\|\right).$$

Приведем резюмирующую теорему, вытекающую из теорем 2 и 8.

Теорема 9. Пусть выполнены условия (B), (F), а также предположения (42). Пусть $u_n^0, u_n^1 \in \mathcal{D}(A_n)$, $u^0, u^1 \in \mathcal{D}(A)$. Тогда справедливы оценки

$$\|U_{n,b}(t) - p_n u(t)\| \le \varepsilon_{n,1},$$

а при условии $\tau_n \|A_n\| \le const$ имеем также

$$\|U_{n,f}(t) - p_n u(t)\| \le \varepsilon_{n,1}.$$

Автор признателен проф. Г.М.Вайникко за помощь и постоянное внимание к работе.

Литература

1. Вайникко Г., Анализ дискретизационных методов. Тарту, 1976.
2. Вайникко Г.М., Пискарев С., О регулярно согласованных операторах. Изв. ВУЗ. Математика, 1977, № 10, 25-36.
3. Данфорд Н., Шварц Дж., Линейные операторы. Общая теория. Москва, 1962.
4. Иосида К., Функциональный анализ. Москва, 1967.
5. Като Т., Теория возмущений линейных операторов. Москва, 1972.
6. Крейн С.Г., Линейные дифференциальные уравнения в банаховом пространстве. Москва, 1967.
7. Марчук Г.И., Методы вычислительной математики. Москва, 1977.
8. Пискарев С., Об аппроксимации голоморфных полугрупп. Уч. зап. Тартуск. ун-та, 1979, 492, 3-14.
9. Самарский А.А., Теория разностных схем. Москва, 1977.
10. Самарский А.А., Гулин А.В., Устойчивость разностных схем. Москва, 1973.
11. Самарский А.А., Николаев Е.С., Методы решения сеточных уравнений. Москва, 1978.

11[a]. С о б о л е в с к и й П.Е., Ч е б о т а р е в а Л.М., Приближенное решение методом прямых задачи Коши для абстрактного гиперболического уравнения. Изв. ВУЗ. Математика, 1977, № 5, 105–116.

12. Х и л л е Э., Ф и л л и п с Р.С., Функциональный анализ и полугруппы. Москва, 1962.

13. Х о л л М., Комбинаторика. Москва, 1970.

14. D a P r a t o, G., G i u s t i, E. Una caratterizzazione dei generatori di funzioni coseno astratte. Bol. Unione Mat. ital., 1967, 22, 357-362.

15. F a t t o r i n i, H.O., Ordinary differential equations in linear topological spaces, I. J. Different. Equat., 1968, 5, 72-105.

16. F a t t o r i n i, H.O., Ordinary differential equations in linear topological spaces, II, J. Different. Equat., 1969, 6, 50-70.

17. G o l d s t e i n, J.A., Semigroups and hyperbolic equations. J. Funct. Anal., 1969, 4, 50-70.

18. G o l d s t e i n, J.A., On a connection between first and second order differential equations in Banach spaces, J. Math. Anal. and Appl., 1970, 30, 246-251.

19. G o l d s t e i n, J.A., On the convergence and approximation of cosine functions. Aequat. Math., 1974, 11, 201-205.

20. K o n i s h i, Y., Cosine functions of operators in locally convex spaces. J. Fac. Sci. Univ.Tokyo, 1971/72, Sec. I A Math., 18, 443-463.

21. N a g y, B., On the generators of cosine operator functions. Publicationes Math., 1974, 21, 151-154.

22. N a g y, B., Cosine operator functions and the abstract Cauchy problem. Periodica Math. Hung., 1976, 7,3-4, 213-217.

23. S o v a, M., Cosine operator functions. Rozprawy Matematiycane, 1966, 49, 1-46.

24. T r a v i s, C.C., W e b b, G.F., Second Order differential Equations in Banach Space. Preprint.

25. T r a v i s, C.C., W e b b, G.F., Cosine families and abstract nonlinear second order differential equations. Acta Math. Scient. Hung., 1978, 32, № 3-4, 75-96.

26. U s h i j i m a, T., Approximation theory for semigroups of linear operators and its application to approximation of wave equations. Japan. J. Math., 1975, 1, № 1, 185-224.
27. V a i n i k k o, G., Approximative methods for nonlinear equations (two approaches to the convergence problem). Nonlinear Analysis. Theory, Methods and Appl., 1978, 2, № 6, 647-687.

Поступило
19 I 1979

DISCRETIZATION OF ABSTRACT HYPERBOLIC EQUATION

S. Piskarjov

Summary

Let E and E_n be Banach spaces such that there are continuous linear maps $p_n : E \longrightarrow E_n$ with the property

$$\| p_n x \| \longrightarrow \| x \| .$$

In this paper the problem of discretization to the second order differential equations (1) is considered.

Some time ago Sova [23], Da Prato—Guisti [14] and Fattorini [16] showed that the Cauchy problem for this equation is uniformly well-posed if and only if A generates a cosine function. As a continuation of oue previous work [8] the present paper is a study of semi and full discretization of (1) from the viewpoint of the operator theory. The theorem is proved which is the cosine function analog of the Trotter—Kato theorem if operators A_n and A are definedin different spaces E_n and E. As for the full discrete approximation where the time variable is also discretisized we shall deal with a notion of discrete cosine function and two other discretization schemes. Stability condition and error estimates for explicit and implicit methods are given.

Уч.зап.Тартуск.ун-та,1979,500,24-32

РЕШЕНИЕ НЕОДНОРОДНОГО АБСТРАКТНОГО ЛИНЕЙНОГО ГИПЕРБОЛИЧЕСКОГО УРАВНЕНИЯ

С. Пискарев

Тартуский государственный университет

Введение

Настоящая работа посвящена аппроксимации неоднородного дифференциального уравнения второго порядка. Приводятся оценки сходимости в случаях, когда свободный член один и два раза непрерывно дифференцируем. В § 5 изучается вопрос компактной сходимости синус-функций.

Детальное обсуждение понятий при дискретной сходимости можно найти в [1].

§1. Основная задача

В банаховом пространстве E рассмотрим задачу Коши

$$\frac{d^2}{dt^2}u(t) = Au(t) + f(t), \; u(0) = u^0, \; u'(0) = u^1, \tag{1}$$

где A является инфинитезимальным генератором косинус-функции, а $f(t)$ – некоторая непрерывная функция со значениями в пространстве E. Как обычно сильным решением задачи (1) назовем функцию $u(t)$ такую, что она дважды непрерывно дифференцируема и удовлетворяет (1). Известно (см. [6], стр.560), что если функция $u(t)$ является сильным решением, то

$$u(t) = C(t)u^0 + S(t)u^1 + \int_0^t S(t-s)f(s)\,ds, \tag{2}$$

где $S(t) = \int_0^t C(s)\,ds$ – синус-функция, ассоциированная с косинус-функцией $C(t)$. Обратное вообще говоря неверно. Поэтому функцию, определенную по (2), будем называть слабым решением задачи (1). Обозначим

$$\mathcal{D}^1 = \{x \in E : C(t)x \text{ – один раз непрерывно дифференцируемая}\}.$$

Укажем для информации

<u>Предложение 1</u>. (см. [2], стр. 30, [6], стр. 563-564). Пусть A – инфинитезимальный генератор косинус-функции $C(t)$ и $u^0 \in \mathcal{D}(A)$, $u^1 \in \mathcal{D}^1$, тогда следующие условия эквивалентны:

1° $u(t)$ – является сильным решением задачи (1);

2° функция $g(t)=\int_0^t S(t-s)f(s)ds$ дважды непрерывно дифференцируема по t;

3° функция $g(t)\in\mathcal{D}(A)$ и $Ag(t)$ непрерывна по t.

Если же косинус-функция $C(t)$ удовлетворяет условию (F) (см. [4], стр. 16) и $0\in\rho(B)$, то для существования сильного решения задачи (1) достаточно, чтобы функция $f(t)$ обладала одним из свойств:

1) функция $f(t)$ непрерывно дифференцируема;

2) функция $Bf(t)$ определена и непрерывна по t.

Существование сильных и слабых решений уравнения (2) в случае, когда $f(t)=F(t,u(t),u'(t))$, рассматрено в [7].

Всюду далее будем предполагать, что спектр оператора A лежит в левой полуплоскости и существует ограниченный обратный A^{-1}.

§ 2. Полудискретная аппроксимация

В банаховых пространствах E_n рассмотрим задачи

$$\frac{d^2}{dt^2}u_n(t)=A_n u_n(t)+f_n(t),\ u_n(0)=u_n^0,\ u_n'(0)=u_n^1, \qquad (3)$$

где A_n – являются инфинитезимальным генераторами косинус-функций, а $f_n(t)$ – непрерывные функции. Будем предполагать как и в §1, что существуют операторы A_n^{-1}, причем $\|A_n^{-1}\|=O(1)$

Введем следующие условия:

(A) Существует такое число $\lambda\in\bigcap_n\rho(A_n)\cap\rho(A)$, что

$$\mathcal{R}(\lambda;A_n)\longrightarrow\mathcal{R}(\lambda;A);$$

(B) Существуют такие константы M и ω, что

$$\left\|\frac{d^m}{d\lambda^m}(\lambda\mathcal{R}(\lambda^2;A_n))\right\|\le\frac{M\,m!}{(\lambda-\omega)^{m+1}},\ \lambda>\omega,\ m=0,1,2,\dots.$$

Известно (см. [4], теорема ABC), что при выполнении условий (A) и (B) имеют место сходимости $C_n(t)\longrightarrow C(t)$ и $S_n(t)\to S(t)$ равномерно по любому компакту $[0,T]$. Итак, справедливо

Предложение 1. Пусть выполнены условия (A) и (B), а функции $f_n(t)$ таковы, что $f_n(t)\longrightarrow f(t)$, $t\in[0,T]$ и

$$\|S_n(t-s)f_n(s)\|\le L(t,s),\ \int_0^T L(t,s)ds<\infty,\ t\in[0,T]$$

тогда, если $u_n^0\to u^0$ и $u_n^1\to u^1$ - то слабые решения задачи

(3) сходятся к слабому решению задачи (1), а именно:

$$u_n(t) \to u(t) \text{ равномерно по } t \in [0,T]$$

Как и в случае однородной задачи (1) для получения оценок сходимости будем предполагать гладкость начальных условий и свободного члена.

Обозначим

$$h_{n,0}(\Gamma;\alpha;f_n,f) = \sup_{\lambda\in\Gamma} \|R(\lambda^2;A_n)f_n(0) - p_n R(\lambda^2;A)f(0)\| \cdot |\lambda|^{\alpha},$$

$$h_{n,1}(\Gamma;\alpha;f_n,f) = \max_{t\in[0,T]} \sup_{\lambda\in\Gamma} \|R(\lambda^2;A_n)f_n(t) - p_n R(\lambda^2;A)f(t)\| \cdot |\lambda|^{\alpha},$$

$$h_{n,2}(\Gamma;f_n,f) = \max_{t\in[0,T]} \sup_{\lambda\in\Gamma} \|R(\lambda^2;A_n)f_n(t) - p_n R(\lambda^2;A)f(t)\|,$$

$$h_{n,3}(f_n,f) = \max_{t\in[0,T]} \|A_n^{-1}f_n(t) - p_n A^{-1}f(t)\|,$$

$$h_{n,4}(\Gamma;f_n,f) = \sup_{\lambda\in\Gamma} \|R(\lambda^2;A_n)f_n(0) - p_n R(\lambda^2;A)f(0)\|,$$

где $\Gamma = \{\zeta : \zeta = \mu + i\nu, -\infty < \nu < \infty, \mu > \omega\}$, $0 < \alpha < 1$.

Предложение 2. Пусть выполнены условия (A) и (B), а функции $f_n(t)$ и $f(t)$ – непрерывно дифференцируемы на $[0,T]$, причем $f_n'(t) \to f'(t)$ равномерно по $t \in [0,T]$, тогда

$$h_{n,0}(\Gamma;\alpha;f_n,f) \to 0,\ h_{n,1}(\Gamma;\alpha;f_n',f') \to 0,\ h_{n,2}(\Gamma;f_n,f) \to 0$$

$$h_{n,3}(f_n,f) \to 0,\ h_{n,4}(\Gamma;f_n',f') \to 0 \text{ при любом } 0<\alpha<1.$$

Доказательство опирается на равенство

$$R(\lambda_n^2;A_n)f_n(t_n) - p_n R(\lambda_n^2;A)f(t_n) = R(\lambda_n^2;A_n)f_n(t_n) - R(\lambda_n^2;A_n)f_n(t_0) +$$

$$+ R(\lambda_n^2;A_n)f_n(t_0) - R(\lambda_0^2;A_n)f_n(t_0) + R(\lambda_0^2;A_n)f_n(t_0) - p_n R(\lambda_0^2;A)f(t_0) +$$

$$+ p_n\big(R(\lambda_0^2;A) - R(\lambda_n^2;A)\big)f(t_0) + p_n\Big(R(\lambda_n^2;A)\big(f(t_0) - f(t_n)\big)\Big)$$

и проводится от противного, аналогично [4], предложение 2. Предложение доказано.

Для удобства записи положим

$$\varepsilon_{n,0} = h_{n,0}(\Gamma;\alpha;f_n,f) + 2h_{n,3}(f_n,f) + t\,h_{n,3}(f_n',f') + t\,h_{n,1}(\Gamma;\alpha;f_n',f').$$

Предложение 3. Пусть выполнено условие (B), а функции $f_n(t)$ и $f(t)$ непрерывно дифференцируемы. Пусть $u_n^0 = u_n^1 = \theta$ и $u^0 = u^1 = \theta$. Тогда

$$\|u_n(t) - p_n u(t)\| \le \varepsilon_{n,0}. \tag{4}$$

Доказательство. Справедливо тождество

$$\int_0^t S(t-s)f(s)ds = C(t)A^{-1}f(0) - A^{-1}f(t) + \int_0^t C(t-s)A^{-1}f'(s)ds \tag{5}$$

Учитывая (2), для доказательства предложения достаточно оценить

$$\Delta_{n,0} = \int_0^t S_n(t-s)f_n(s)ds - p_n \int_0^t S(t-s)f(s)ds.$$

Используя выражение (5) и утверждение из [4], теорема 2, находим $\|\Delta_{n,0}\| \le \varepsilon_{n,0}$. Предложение доказано.

Для сокращения записи считаем, что

$$\varepsilon_{n,1} = h_{n,\beta}(\Gamma;\alpha;f_n,f) + 2h'_{n,3}(f_n,f) + h_{n,\mu}(\Gamma;f'_n,f') +$$
$$+ t\,h_{n,2}(\Gamma;f''_n,f'') + t\,h_{n,3}(f''_n,f'') + t\|A_n^{-1}f'_n(0) - p_n A^{-1}f'(0)\|.$$

Предложение 4. Пусть выполнено условие (B), а функции $f_n(t)$ и $f(t)$ дважды непрерывно дифференцируемы. Пусть $u_n^0 = u_n^1 = \theta$, $u^0 = u^1 = \theta$. Тогда

$$\|u_n(t) - p_n u(t)\| \le \varepsilon_{n,1}.$$

Доказательство проводится аналогично предложению 3 с учетом тождества

$$\int_0^t S(t-s)f(s)ds = C(t)A^{-1}f(0) - A^{-1}f(t) + S(t)A^{-1}f'(0) - \int_0^t S(t-s)A^{-1}f''(s)ds.$$

Замечание 1. Оценки сходимости для полудискретной аппроксимации с ненулевыми начальными условиями и нулевой правой частью указаны в [4] и здесь не приводятся.

§ 3. Полная дискретизация задачи (1)

В этом параграфе дополнительно предположим, что операторы A_n удовлетворяют условию (F), причем

$$\|(\lambda - B_n)^{-k}\| \le L/(\mathrm{Re}\lambda - \omega)^k, \ |\mathrm{Re}\lambda| > \omega \quad \text{и} \quad \|B_n^{-1}\| = O(1), \tag{6}$$

где L – некоторая константа, не зависящая от n. Точное решение задачи (3) (сильное или слабое) выражается в виде:

$$u_n(t) = C_n(t)u_n^0 + S_n(t)u_n^1 + \int_0^t S_n(t-s)f_n(s)ds. \tag{7}$$

Для аппроксимации выражения в (7) мы определим следующие семейства операторов:

$$\mathcal{C}_{n,\ell}(k\tau_n+\tau_n) = 2(1-\tau_n^2A_n)^{-1}(\mathcal{C}_{n,\ell}(k\tau_n) - \mathcal{C}_{n,\ell}(k\tau_n-\tau_n)),\ \mathcal{C}_{n,\ell}(0)=I_n,\ \mathcal{C}_{n,\ell}(\tau_n)=(1-\tau_n^2A_n)^{-1}$$

$$\mathcal{S}_{n,\ell}(k\tau_n+\tau_n) = 2(1-\tau_n^2A_n)^{-1}(\mathcal{S}_{n,\ell}(k\tau_n) - \mathcal{S}_{n,\ell}(k\tau_n-\tau_n)),\ \mathcal{S}_{n,\ell}(0)=0,\ \mathcal{S}_{n,\ell}(\tau_n)=\tau_n(1-\tau_n^2A_n)^{-1}$$

$$\mathcal{C}_{n,f}(k\tau_n+\tau_n)=2\mathcal{C}_{n,f}(k\tau_n)-(1-\tau_n^2A_n)\mathcal{C}_{n,f}(k\tau_n-\tau_n),\ \mathcal{C}_{n,f}(0)=I_n,\ \mathcal{C}_{n,f}(\tau_n)=1+\tau_n^2A_n$$

$$\mathfrak{S}_{n,f}(k\tau_n+\tau_n)=2\mathfrak{S}_{n,f}(k\tau_n)-(1-\tau_n^2A_n)\mathfrak{S}_{n,f}(k\tau_n-\tau_n),\ \mathfrak{S}_{n,f}(0)=0,\ \mathfrak{S}_{n,f}(\tau_n)=\tau_n(1+\tau_n^2A_n)$$

Построим функции

$$\mathcal{U}_{n,f}(k\tau_n)=\mathcal{C}_{n,f}(k\tau_n)u_n^0+\mathfrak{S}_{n,f}(k\tau_n)u_n^1+\sum_{j=0}^{k-1}\mathfrak{S}_{n,f}((k-j)\tau_n)f_n(j\tau_n)\tau_n, \quad (8)$$

$$\mathcal{U}_{n,b}(k\tau_n)=\mathcal{C}_{n,b}(k\tau_n)u_n^0+\mathfrak{S}_{n,b}(k\tau_n)u_n^1+\sum_{j=0}^{k-1}\mathfrak{S}_{n,b}((k-j)\tau_n)f_n(j\tau_n)\tau_n. \quad (9)$$

Нам понадобятся следующие обозначения:

$$F_1(t,f_n)=\max_{0\le s\le t}\|f_n(s)\|,\quad F_2(t,f_n)=F_1(t,B_nf_n),$$

$$F_3(t,f_n)=F_1(t,A_nf_n),\quad \mathfrak{F}_k(t,f_n)=\sum_{i=0}^{k}F_i(t,f_n).$$

<u>Предложение 5</u>. Пусть выполнено условие (B), $u_n^0=u_n^1=\theta$, $u^0=u^1=\theta$ и f_n непрерывно дифференцируемая функция, причем $f_n(s)\in\mathcal{D}(B_n)$ при $s\in[0,t]$. Тогда существуют константы M и η такие, что

$$\|\mathcal{U}_{n,b}(t)-u_n(t)\|\le\tau_n t Me^{\eta t}\big(F_1(t,f_n')+\mathfrak{F}_2(t,f_n)\big),\ t=k\tau_n, \quad (10)$$

а при выполнении условия

$$\tau_n\|A_n\|\le const \quad (11)$$

имеем также

$$\|\mathcal{U}_{n,f}(t)-u_n(t)\|\le\tau_n t Me^{\eta t}\big(F_1(t,f_n')+\mathfrak{F}_2(t,f_n)\big),\ t=k\tau_n.$$

<u>Доказательство</u>. Докажем, например, оценку (10). Рассмотрим равенство

$$\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}S_n(t-s)f_n(s)ds-\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}\mathfrak{S}_{n,b}(t-i\tau_n)f_n(i\tau_n)ds=$$

$$=\Big[\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}S_n(t-s)f_n(s)ds-\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}S_n(t-s)f_n(i\tau_n)ds\Big]+$$

$$+\Big[\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}S_n(t-s)f_n(i\tau_n)ds-\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}S_n(t-i\tau_n)f_n(i\tau_n)ds\Big]+$$

$$+\Big[\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}S_n(t-i\tau_n)f_n(i\tau_n)ds-\sum_{i=0}^{k-1}\int_{i\tau_n}^{(i+1)\tau_n}\mathfrak{S}_{n,b}(t-i\tau_n)f_n(i\tau_n)ds\Big].$$

Первая квадратная скобка справа оценивается $t\,\mathcal{O}(\tau_n)$ в силу представления

$$f_n(s)-f_n(i\tau_n)=\int_{i\tau_n}^{s} f_n'(\xi)\,d\xi .$$

Для второй используем тождество

$$S_n(t-s)-S_n(t-i\tau_n)=\int_{t-i\tau_n}^{t-s} C_n(\xi)\,d\xi . \qquad (12)$$

Доказательство завершается применением результата [4], теорема 8. Предложение доказано.

Сформулируем общий результат.

Теорема 1. Пусть выполнены условие (B) и предположение (6). Пусть $u_n^0=u_n^1=\theta$ и $u^0=u^1=\theta$, а функции $f_n(t), f(t)$ непрерывно дифференцируемы, причем $f_n\in\mathcal{D}(B_n)$. Тогда

$$\|\mathcal{U}_{n,\ell}(t)-p_n u(t)\|\le \varepsilon_{n,o}+\tau_n t M e^{\eta t}\left(F_1(t,f_n')+\mathcal{F}_2(t,f_n)\right), \qquad (13)$$

а в случае выполнения условия (11).

$$\|\mathcal{U}_{n,f}(t)-p_n u(t)\|\le \varepsilon_{n,o}+\tau_n t M e^{\eta t}\left(F_1(t,f_n')+\mathcal{F}_2(t,f_n)\right),\ t=k\tau_n . \qquad (14)$$

Доказательство вытекает из предложений 3 и 5.

§ 4. Аппроксимация с помощью дискретной синус-функции

До конца параграфа предполагаем, что E и E_n – гильбертовы пространства, а операторы $-A$ и $-A_n$ самосопряженные положительно определенные, причем спектры $\sigma(A_n)$ отделены от нуля расстоянием $\nu>0$. Дискретные синус и косинус-функции определим рекурентно по формулам

$$\mathcal{C}_n(k\tau_n+\tau_n)=2\,\mathcal{C}_n(k\tau_n)\mathcal{C}_n(\tau_n)-\mathcal{C}_n(k\tau_n-\tau_n),\ \mathcal{C}_n(0)=I_n,\ \mathcal{C}_n(\tau_n)=1+\tau_n^2A_n/2\ ;$$

$$\mathfrak{S}_n(k\tau_n+\tau_n)=2\,\mathfrak{S}_n(k\tau_n)\mathcal{C}_n(\tau_n)-\mathfrak{S}_n(k\tau_n-\tau_n),\ \mathfrak{S}_n(0)=0,\ \mathfrak{S}_n(\tau_n)=(\tau_n^2+\tau_n^4A_n/4)^{1/2}.$$

Выражение (7) аппроксимируем следующим образом:

$$\mathcal{U}_n(k\tau_n)=\mathcal{C}_n(k\tau_n)u_n^0+\mathfrak{S}_n(k\tau_n)u_n^1+\frac{\tau_n}{2}\mathfrak{S}_n(t)f_n(0)+\sum_{i=0}^{k-1}\mathfrak{S}_n(t-i\tau_n)f_n(i\tau_n)\tau_n . \qquad (15)$$

Предложение 6. Пусть выполнены условия (B) и

$$\tau_n^2\|A_n\|\le\delta<4 .$$

Пусть, далее, $u_n^0=u_n^1=\theta$ и $u^0=u^1=\theta$, а функции $f_n(t), f(t)$ дважды непрерывно дифференцируемы, причем $f_n(s)\in\mathcal{D}(A_n)$, $s\in[0,t]$. Тогда

$$\|\mathcal{U}_n(t)-u_n(t)\|\le\tau_n^2 t M e^{\eta t}\left(F_1(t,f_n'')+F_3(t,f_n)\right),\ t=k\tau_n . \qquad (16)$$

Доказательство. Рассмотрим равенство

$$\int_0^t S_n(t-s)f_n(s)ds - \frac{\tau_n}{2}\mathfrak{S}_n(t)f_n(0) - \sum_{i=0}^{k-1}\mathfrak{S}_n(t-i\tau_n)f_n(i\tau_n)\tau_n =$$

$$=\left\{\int_0^t S_n(t-s)f_n(s)ds - \frac{\tau_n}{2}S_n(t)f_n(0) - \sum_{i=0}^{k-1}S_n(t-i\tau_n)f_n(i\tau_n)\tau_n\right\}+$$

$$+\left\{\frac{\tau_n}{2}S_n(t)f_n(0)+\sum_{i=0}^{k-1}S_n(t-i\tau_n)f_n(i\tau_n)\tau_n-\frac{\tau_n}{2}\mathfrak{S}_n(t)f_n(0)-\sum_{i=0}^{k-1}\mathfrak{S}_n(t-i\tau_n)f_n(i\tau_n)\tau_n\right\}.$$

Учитывая, что

$$\left(S_n(t-s)f_n(s)\right)''_s = S_n(t-s)f_n''(s) - 2C_n(t-s)f_n'(s) + S_n(t-s)A_n f_n(s),$$

первую фигурную скобку справа в предыдущем равенстве легко оценить по формуле трапеций (см. [5], стр. 245). Для оценки второй фигурной скобки достаточно воспользоваться в [4] теоремой 4. Предложение доказано.

Укажем резюмирующую теорему, вытекающую из предложений 4 и 6.

Теорема 2. Пусть выполнены условия (B) и

$$\tau_n^2\|A_n\| \le \delta < 4.$$

Пусть $u_n^0 = u_n^1 = \theta$ и $u^0 = u^1 = \theta$, а функции $f_n(t), f(t)$ дважды непрерывно дифференцируемы, причем $f(s)\in\mathcal{D}(A), f_n(s)\in\mathcal{D}(A_n)$, при $s\in[0,t]$. Тогда $(t=k\tau_n)$

$$\|\mathcal{U}_n(t)-p_n u(t)\| \le \varepsilon_{n,1} + \tau_n^2 t M e^{\eta t}\left(F_1(t,f_n'')+F_3(t,f_n)\right). \qquad (17)$$

§ 5. Компактная сходимость синус-функций

Напомним следующее

Предложение (см. [5], стр. 557). Синус-функция $S(t)$ компактна тогда и только тогда, когда $\mathcal{R}(\lambda;A)$ компактна хоть для одного, а следовательно любого $\lambda\in\rho(A)$.

Здесь будет доказан результат, аналогичный утверждению, справедливому для полугрупп (см. [3], стр. 73). Определение компактной сходимости можно найти в[1].

Теорема 3. Пусть выполнены условия (A) и (B), тогда компактная сходимость

$$\mathcal{R}(\lambda;A_n)\longrightarrow\mathcal{R}(\lambda;A)$$

имеет место тогда и только тогда, когда

$$S_n(t)\longrightarrow S(t) \text{ компактно при } t\in(-\infty,\infty).$$

Доказательство. Пусть $S_n(t)\longrightarrow S(t)$ компактно, тогда учитывая выражение

$$\mathcal{R}(\lambda^2;A_n)=\int_0^\infty e^{-\lambda t}\,S_n(t)\,dt\,,$$

легко получить, что $\mu((\mathcal{R}(\lambda^2;A_n)x_n))=0$ для любой ограниченной последовательности $\{x_n\}$. Здесь μ – дискретная мера некомпактности (см. [1], стр. 17).

Обратно, пусть $\mathcal{R}(\lambda;A_n)\to\mathcal{R}(\lambda;A)$ компактно. Достаточно доказать, что $\mu((S_n(t)x_n))=0$ лишь только $\{x_n\}$ ограничена. Рассмотрим оператор-функции $\mathcal{F}_n(\lambda)=\lambda^2\mathcal{R}(\lambda^2;A_n)S_n(t)$ (до конца доказательства переменную t фиксируем). В силу ограниченности $\{S_n(t)x_n\}$ последовательность $\{\mathcal{F}_n(\lambda)x_n\}$ компактна. Имеем

$$\mu((S_n(t)x_n))\le\mu((\mathcal{F}_n(\lambda)x_n))+\mu(((\mathcal{F}_n(\lambda)-S_n(t))x_n))\,.$$

Первый член справа равен нулю и мы завершим доказательство теоремы, показав, что

$$\|S_n(t)-\mathcal{F}_n(\lambda_n)\|\to 0 \quad \text{при} \quad \lambda_n\to\infty.$$

Доказательство последнего утверждения повторяет рассуждения в работе [6], приведем его. Так как $\|C_n(t)\|\le Me^{\omega t}$, то $\|S_n(t)\|\le(M/\omega)e^{\omega t}$. Вычислим

$$\mathcal{F}_n(\lambda)-S_n(t)=\int_0^\infty\lambda e^{-\lambda s}C_n(s)S_n(t)\,ds-S_n(t)=$$

$$=\int_0^\infty\lambda e^{-\lambda s}\,(S_n(t+s)-S_n(t-s))/2\,ds-S_n(t)=$$

$$=\int_0^\infty\lambda/2\;e^{-\lambda s}\,(S_n(t+s)-2S_n(t)+S_n(t-s))\,ds.$$

Последний интеграл разбивается на два: $I_1=\int_0^{\delta_n}(\lambda/2)e^{-\lambda s}(S_n(t+s)-2S_n(t)+S_n(t-s))\,ds$ и $I_2=\int_{\delta_n}^\infty\frac{\lambda}{2}e^{-\lambda s}(S_n(t+s)-2S_n(t)+S_n(t-s))\,ds$.

За счет выбора δ_n, в силу непрерывности по норме $S_n(t)$ см. (12), величину $\|I_1\|$ можно сделать сколь угодно малой, например, меньше произвольного $\varepsilon>0$. Выбрав $\lambda=\lambda_n(\delta_n)$ достаточно большим, сделаем $\|I_2\|$ также меньше ε. Теорема доказана.

Литература

1. Вайникко Г., Анализ дискретизационных методов. Тарту, 1976.
2. Крейн С.Г., Линейные дифференциальные уравнения в ба-

наховом пространстве. Москва, 1967.
3. Пискарев С., О компактной сходимости резольвент и полугрупп. Материалы конф. методы алгебры и функционалын. анализа при исслед. семейств операторов. Тарту, 1978, 71-74.
4. Пискарев С., Дискретизация абстрактного гиперболического уравнения. Настоящий сборник, стр. 3-23.
5. Võhandu, L., Tamme, E., Luht, L., Arvutusmeetodid I. Tallinn, 1971.
6. Travis, C.C., Webb, G.F., Compactness, regularity, and uniform continuity properties of strongly continuous cosin families. Huston J. Math., 1977, 3, № 4, 555-567.
7. Travis, C.C., Webb, G.F., Cosine families and abstract nonlinear second order differential equations. Acta Math. Scient. Hung., 1978, 32, № 3-4, 75-96.

Поступило
15 I 1979

SOLUTION OF INHOMOGENEOUS ABSTRACT LINEAR HYPERBOLIC EQUATION

S. Piskarjov

Summary

The purpose of the present paper is to obtain some estimates of the semi-discrete as well as full discrete approximation of abstract second order differential inhomogeneous equation. In § 5 we establish that sine functions converge compactly $S_n(t) \longrightarrow S(t)$ if and only if the resolvents $\mathcal{R}(\lambda; A_n) \longrightarrow \mathcal{R}(\lambda; A)$ converge compactly.

Уч.зап.Тартуск.ун-та,1979,500,33-42

КУСОЧНО-ЛИНЕЙНАЯ АППРОКСИМАЦИЯ РЕШЕНИЯ ИНТЕГРАЛЬНОГО УРАВНЕНИЯ С ЛОГАРИФМИЧЕСКОЙ ОСОБЕННОСТЬЮ В ЯДРЕ

А.Педас

Тартуский государственный университет

§1. Введение

В данной работе рассматривается линейное интегральное уравнение

$$x(t) = \int_0^1 \varkappa(|t-s|)x(s)ds + f(t) \quad (0 \le t \le 1) \qquad (1.1)$$

с непрерывно дифференцируемым[1] при $0 < \tau \le 1$ ядром $\varkappa(\tau)$, имеющим при $\tau = 0$ логарифмическую особенность:

$$|\varkappa(\tau)| \le b\,(|\ln\tau| + 1) \quad (0 < \tau \le 1,\ b = const); \qquad (1.2)$$

$$|\varkappa'(\tau)| \le b_1/\tau \qquad (0 < \tau \le 1;\ b_1 = const). \qquad (1.3)$$

Приближенное решение уравнения (1.1) ищется в виде кусочно-линейной функции

$$x_n(t) = \sum_{j=0}^{n} \xi_{j,n}\varphi_{j,n}(t), \qquad (1.4)$$

где $\varphi_{j,n}$ $(j = 0,1,\dots,n)$ – непрерывные на $[0,1]$ функции, линейные на каждом интервале $[(k-1)h, kh]$, $k = 1,2,\dots,n$, $h = 1/n$, и такие, что $\varphi_{j,n}(kh) = \delta_{jk}$, $j,k = 0,1,\dots,n$ ($\delta_{j,k}$ - символ Кронекера). Неизвестные коэффициенты $\xi_{j,n}$ $(j = 0,1,\dots,n)$ определяются методом коллокации, т.е. из условий

$$x_n(ih) = \int_0^1 \varkappa(|ih-s|)x_n(s)ds - f(ih),\ i = 0,1,\dots,n,$$

или, что то же самое, из системы линейных алгебраических уравнений

$$\xi_{i,n} = \sum_{j=0}^{n}\Big[\int_0^1 \varkappa(|ih-s|)\varphi_{j,n}(s)ds\Big]\xi_{j,n} + f(ih)$$
$$(i = 0,1,\dots,n). \qquad (1.5)$$

---

[1] Здесь и всюду в дальнейшем $x'(t) = dx/dt$, $x''(t) = d^2x/dt^2$.

Пусть x^* – точное решение уравнения (1.1), $(\xi^*_{0,n}, \xi^*_{1,n}, \ldots, \xi^*_{n,n})$ – точное решение системы (1.5), x^*_n – приближенное решение уравнения (1.1), получаемое из (1.4) при $\xi_{j,n} = \xi^*_{j,n}$. В [8] показано, что в случае $f \in C^2[0,1]$ (точнее в условиях приведенной ниже теоремы 1) справедлива оценка погрешности

$$\max_{0 \le t \le 1} t^\alpha (1-t)^\alpha |x^*(t) - x^*_n(t)| \le c_1 h^{1+\alpha} |\ln h|, \quad 0 \le \alpha < 1, \tag{1.6}$$

где постоянная $c_1 = c_1(\alpha)$ не зависит от n (от h).

Ввиду малой гладкости решения x^* уравнения (1.1) оценка (1.6) близка к точной. Она согласуется с результатами численного экспериментирования, однако, в случае интегральных уравнений типа Милна [11], в узлах коллокации наблюдалась более быстрая сходимость. Анализ погрешности подтвердил это наблюдение [4]:

$$\max_{0 \le i \le n} |x^*(ih) - x^*_n(ih)| \le c_2 h^2 (\ln h)^2, \tag{1.7}$$

где постоянная c_2 не зависит от n (от h).

Целью настоящей статьи является обобщение оценки (1.7) на класс интегральных уравнений с логарифмической особенностью в ядре (которому принадлежит и интегральное уравнение Милна).

§ 2. Теорема сходимости

Относительно вычислительной схемы (1.4)–(1.5) справедлива следующая

Теорема 1. Пусть ядро $\varkappa(\tau)$ интегрального уравнения (1.1) непрерывно дифференцируемо при $0 < \tau \le 1$, удовлетворяет оценкам (1.2) и (1.3), и пусть существуют конечные пределы

$$\lim_{\tau \to 0+} \varkappa(\tau)/\ln\tau\,; \quad \lim_{\tau \to 0+} \tau\varkappa'(\tau). \tag{2.1}$$

Пусть свободный член $f(t)$ уравнения (1.1) дважды непрерывно дифференцируем при $0 \le t \le 1$. Пусть, наконец, уравнение (1.1) имеет единственное решение $x^*(t)$.

Тогда при достаточно больших n система уравнений (1.5) имеет единственное решение $(\xi^*_{0,n}, \xi^*_{1,n}, \ldots, \xi^*_{n,n})$. При этом приближенные решения $x^*_n(t)$, получаемые из (1.4) при $\xi_{j,n} = \xi^*_{j,n}$, сходятся равномерно на $[0,1]$ к точному решению $x^*(t)$ уравнения (1.1). Справедливы оценки погрешности (1.6) и (1.7).

Ввиду [8] в доказательстве нуждается лишь оценка (1.7); её доказательство будет приведено в последующих параграфах.

Замечание 1. Из непрерывности ядра $\varkappa(\tau)$ при $0<\tau\leq 1$ и неравенства (1.2) вытекает, что интегральный оператор уравнения (1.1) (оператор T в формуле (3.4)) вполне непрерывен в пространстве $C[0,1]$. Поэтому предположение о существовании решения уравнения (1.1) равносильно тому, что соответствующее однородное уравнение (т.е. уравнение $x=Tx$) имеет лишь нулевое решение.

Замечание 2. В условиях теоремы 1 решение $x^*(t)$ уравнения (1.1) негладко. Точнее [8, 10], $x^*\in C[0,1]\cap C^2(0,1)$, причем[2]

$$|x^{*\prime}(t)|\leq c_3(|\ln t|+|\ln(1-t)|)\quad(0<t<1;\ c_3=\mathrm{const}),\tag{2.2}$$

$$|x^{*\prime\prime}(t)|\leq c_4(t^{-1}+(1-t)^{-1})\quad(0<t<1;\ c_4=\mathrm{const}).\tag{2.3}$$

Если $x^*(0)\neq 0$, $x^*(1)\neq 0$, то эти оценки не могут быть улучшены в смысле порядка. Поэтому добится высокого порядка точности приближенных методов для решения уравнения (1.1) в условиях теоремы 1 довольно сложно. Для сравнения напомним, что в случае применения метода механических квадратур с формулой средних прямоугольников непосредственно к уравнению (1.1) получается лишь оценка [3] порядка $O(h|\ln h|)$, а в случае применения этого метода вместе со способом уничтожения особенности в ядре при совпадении аргументов ([6], стр. 115) получается оценка [5] порядка $O(h^2|\ln h|^2)$.

Замечание 3. Теорема 1 допускает обобщение на случай интегральных уравнений со степенной и степенно-логарифмической особенностью в ядре [9, 10].

Замечание 4. В вычислительной схеме (1.4)–(1.5) был рассмотрен случай равноотстоящих узлов. Однако теорема 1 остается справедливой и в случае системы узлов

$$0=t_{0,n}<t_{1,n}<\dots<t_{n,n}=1,$$

---

[2] В [8] вместо (2.3) установлена более грубая оценка.

если только

$$h_n \equiv \max_{1\le j\le n} (t_{j,n} - t_{j-1,n}) \to 0 \text{ при } n \to \infty.$$

(В этом случае в схеме (1.4)-(1.5) и в оценках (1.6)-(1.7) шаг дискретизации h следует заменить на h_n и узлы ih на $t_{i,n}$).

§ 3 . Выбор пространств и вспомогательные результаты

Интегральное уравнение (1.1) рассмотрим как операторное уравнение

$$x = Tx + f \tag{3.1}$$

в банаховом пространстве $E = C \equiv C[0,1]$ непрерывных на $[0,1]$ функций с обычной max-нормой, а систему уравнений (1.5) как операторное уравнение

$$\xi_n = T_n \xi_n + q_n f \tag{3.2}$$

в банаховом пространстве $E_n = m_{n+1}$ векторов вида $\xi_n = (\xi_{0,n}, \xi_{1,n}, \dots, \xi_{n,n})$ с нормой

$$\|\xi_n\| = \max_{0\le i\le n} |\xi_{i,n}|. \tag{3.3}$$

Здесь T и T_n – линейные вполне непрерывные операторы соответственно в пространствах $C[0,1]$ и m_{n+1}, задаваемые формулами

$$(Tx)(t) = \int_0^1 \varkappa(|t-s|)x(s)\,ds, \tag{3.4}$$

$$(T_n\xi_n)_i = \sum_{j=0}^{n}\Big[\int_0^1 \varkappa(|ih-s|)\varphi_{j,n}(s)\,ds\Big]\xi_{j,n} \quad (i=0,1,\dots,n), \tag{3.5}$$

а q_n – линейные непрерывные операторы из $C[0,1]$ в m_{n+1} (связывающие отображения), задаваемые формулами

$$q_n x = (x(0), x(h), \dots, x(nh)). \tag{3.6}$$

Напомним некоторые понятия и результаты из [2]. Пусть E и E_n ($n=1,2,\dots$) – произвольные банаховы пространства, а $q_n \in L(E,E_n)$ – линейные непрерывные операторы[3].

[3] Через $L(E,F)$ обозначается пространство линейных непрерывных операторов из E в F (определенных на всем E).

Определение 1. Последовательность вполне непрерывных операторов $T_n \in L(E_n, E_n)$ компактно аппроксимирует вполне непрерывный оператор $T \in L(E, E)$ по отношению к связывающим отображениям $q_n \in L(E, E_n)$, если выполнены следующие два условия:

1) $\| q_n T x - T_n q_n x \|_{E_n} \to 0$ при $n \to \infty$ для $\forall x \in E$;

2) для любой последовательности $\{\xi_n\}$, $\xi_n \in E_n$, $\|\xi_n\| \leq 1$ $(n = 1, 2, \ldots)$, существуют такие $y_n \in E$, что $q_n y_n = T_n \xi_n$ $(n = 1, 2, \ldots)$ и последовательность $\{y_n\}$ компактна в E.

Теорема 2. Пусть выполнены следующие условия:

1) последовательность вполне непрерывных операторов $\{T_n\}$, $T_n \in L(E_n, E_n)$ компактно аппроксимирует вполне непрерывный опеатор $T \in L(E, E)$ по отношению к связывающим отображениям $q_n \in L(E, E_n)$;

2) уравнение (3.1) имеет единственное решение $x^* \in E$;

3) операторы $q_n \in L(E, E_n)$ $(n = 1, 2, \ldots)$ удовлетворяют условиям

$$q_n E = E_n \quad (n = 1, 2, \ldots);$$

$$\| q_n \| \leq a \quad (n = 1, 2, \ldots;\ a = const);$$

$$\| \xi_n \|_{E_n} \geq d \inf_{x \in E,\, q_n x = \xi_n} \| x \|_E \quad \text{для} \quad \forall \xi_n \in E_n$$
$$(n = 1, 2, \ldots;\ d = const > 0);$$

$$\| q_n x \|_{E_n} \to \| x \|_E \quad \text{при } n \to \infty \text{ для } \forall x \in E.$$

Тогда уравнение (3.2) имеет при достаточно больших n единственное решение $\xi_n^* \in E_n$ и $\| q_n x^* - \xi_n^* \| \to 0$ при $n \to \infty$. Справедлива оценка

$$d_1 \varepsilon_n(x^*) \leq \| q_n x^* - \xi_n^* \|_{E_n} \leq d_2 \varepsilon_n(x^*), \tag{3.7}$$

где d_1, d_2 – некоторые положительные постоянные,

$$\varepsilon_n(x^*) = \| q_n T x^* - T_n q_n x^* \|_{E_n}. \tag{3.8}$$

Заметим, что условие 3) теоремы 2 в нашем случае ($E = C[0,1]$, $E_n = m_{n+1}$, q_n определены в (3.6)) выполнено, причем $a = d = 1$. Так как в теореме 1 предполагается существование решения x^* интегрального уравнения (1.1) (уравнения

(3.1)), то для применения теоремы 2 остается провести проверку условия 1) теоремы 2.

§ 4. Доказательство оценки (1.7)

Мы установим, что для операторов T и T_n, определенных соответственно формулами (3.4) и (3.5), выполняются условия 1) и 2) определения 1. Рассуждения проводятся в условиях теоремы 1, и тем самым мы получим доказательство оценки (1.7) из утверждений теоремы 2 о существовании единственного решения ξ_n^* и оценки (3.7).

Пусть $x \in C[0,1]$. Тогда

$$\varepsilon_n(x) = \| q_n T x - T_n q_n x \|_{m_{n+1}} =$$

$$= \max_{0 \le i \le n} \left| \int_0^1 \varkappa(|ih - s|)[x(s) - (Q_n x)(s)]\,ds \right|, \quad (4.1)$$

где

$$(Q_n x)(t) = \sum_{j=0}^{n} x(jh)\varphi_{j,n}(t) \quad (x \in C[0,1]). \quad (4.2)$$

Легко показать, что

$$\|x - Q_n x\|_C \to 0 \quad \text{при } h \to \infty \text{ для } \forall x \in C[0,1]. \quad (4.3)$$

Так как в силу (1.2)

$$\max_{0 \le i \le h} \int_0^1 |\varkappa(|ih - s|)|\,ds \le c \quad (c = const),$$

то из (4.3) и (4.1) вытекает, что условие 1) определения 1 выполняется:

$$\varepsilon_n(x) \to 0 \quad \text{при } n \to \infty \text{ для } \forall x \in C[0,1].$$

Далее, так как оператор $T \in L(C,C)$ вполне непрерывен, то для любой последовательности $\{\xi_n\}$, $\xi_n = (\xi_{0,n}, \xi_{1,n}, \dots, \xi_{n,n}) \in m_{n+1}$, $\|\xi_n\| \le 1$ $(n = 1, 2, \dots)$, существуют функции

$$y_n(t) = \sum_{j=0}^{n} \left[\int_0^1 \varkappa(|t - s|)\varphi_{j,n}(s)\,ds \right] \xi_{j,n}$$

такие, что $y_n \in C[0,1]$, $q_n y_n = T_n \xi_n$ $(n = 1, 2, \dots)$ и последовательность $\{y_n\}$ компактна в $C[0,1]$. Другими словами, выполняется и условие 2) определения 1 и, следовательно, T_n аппроксимируют T компактно.

Итак теорема 2 применима. По этой теореме получаем, что при достаточно больших n существует единственное решение $\xi_n^* = (\xi_{0,n}^*, \xi_{1,n}^*, \dots, \xi_{n,n}^*)$ системы (1.5) и

$$\max_{0\le i\le n} |x^*(ih) - \xi_{i,n}^*| \le d_2 \max_{0\le i\le n} (A_{i,n} + B_{i,n}), \tag{4.4}$$

где

$$A_{i,n} = \int_{u_i(h)} |\varkappa(|ih-s|)[x^*(s) - (Q_n x^*)(s)]| ds,$$

$$B_{i,n} = \int_{v_i(h)} |\varkappa(|ih-s|)[x^*(s) - (Q_n x^*)(s)]| ds,$$

$$u_i(h) = [(i-1)h, (i+1)h] \cap [0,1]; \; v_i(h) = [0,1] \setminus u_i(h).$$

В силу (1.2)

$$\max_{0\le i\le n} A_{i,n} \le b \|x^* - Q_n x^*\|_C \cdot \max_{0\le i\le n} \int_{(i-1)h}^{(i+1)h} (|\ln|ih-s|| + 1) d$$

$$\le 2bh(|\ln h| + 2) \|x^* - Q_n x^*\|_C, \tag{4.5}$$

а

$$\max_{0\le i\le n} B_{i,n} \le b(|\ln h| + 1) \int_0^1 |x^*(s) - (Q_n x^*)(s)| ds. \tag{4.6}$$

Из (4.2) и (2.2) вытекает, что

$$\|x^* - Q_n x^*\|_{C[0,1]} \le ch|\ln h| \quad (c = const). \tag{4.7}$$

Оценим разность $x^* - Q_n x^*$ в норме пространства $L^1 = L^1[0,1]$

Мы имеем

$$\|x^* - Q_n x^*\|_{L^1} = \left(\int_0^h + \int_h^{1-h} + \int_{1-h}^1 \right) |x^*(s) - (Q_n x^*)(s)| ds \le$$

$$\le 2h \|x^* - Q_n x^*\|_C + \sum_{j=1}^{n-2} \int_{jh}^{(j+1)h} |x^*(s) - (Q_n x^*)(s)| ds \le$$

$$\le 2h \|x^* - Q_n x^*\|_C + \frac{h^2}{2} \sum_{j=1}^{n-2} \left(\frac{1}{j} + \frac{1}{n-j-1} \right). \tag{4.8}$$

На последнем шаге мы воспользовались оценками

$$\int_{jh}^{(j+1)h} |x^*(s) - (Q_n x^*)(s)| ds \le \frac{h^3}{2} \sup_{jh \le s \le (j+1)h} |x^{*\prime\prime}(s)|$$

и (2.3). Ввиду неравенств (4.7) и

$$\sum_{j=1}^{n} 1/j \le \ln n + 1$$

из (4.8) вытекает, что

$$\int_0^1 |x^*(s) - (Q_n x^*)(s)|\, ds \le c h^2 |\ln h| \quad (c = const). \tag{4.9}$$

Оценка (1.7) теперь вытекает из (4.9), (4.7), (4.6), (4.5) и (4.4).

Тем самым завершено доказательство теоремы 1.

§ 5. Численные результаты

Как известно (см., например, [11]), задача переноса излучения в плоско-параллельной среде со сферической индикатрисой рассеивания эквивалентна интегральному уравнению Милна

$$x(t) = \frac{\lambda}{2} \int_0^H E_1(|t-s|)\, x(s)\, ds + d_1 e^{d_2 t}, \quad 0 \le t \le H, \tag{5.1}$$

где x – неизвестная функция (источников), $\lambda \in [0,1]$ – вероятность выживания кванта в акте столкновения,

$$E_1(\tau) = \int_\tau^\infty \mu^{-1} e^{-\mu} d\mu = \int_0^1 \mu^{-1} e^{-\tau/\mu} d\mu \quad (\tau > 0) \tag{5.2}$$

– интегральная экспонента, $H > 0$, d_1, d_2 – некоторые постоянные. Для уравнения (5.1) выполняются все условия теоремы 1 и, следовательно, метод сплайновой коллокации первого порядка (I .4)–(1.5), является сходящимся. Алгоритмическая эффективность метода (1.4)–(1.5) в случае уравнения (5.1) обуславливается тем, что вычисление встречающиеся здесь (в (1.5)) интегралов сводится к вычислению интеральной показательной функции (5.2) в узлах ih, $i = 1, 2, \ldots, n$. Пусть

$$H = 0{,}3; \quad d_1 = 0{,}125 \exp(-0{,}6); \quad d_2 = 2. \tag{5.3}$$

При указанных значениях параметров уравнение (5.1) решалось численно методом (1.4)–(1.5) при различных h. Расчеты проводились в ВЦ Тартуского госуниверситета на ЭВМ "ЕС-I022". Некоторые из результатов вычислений приведены в таблице 1. Для сравнения в первых двух строках таолицы приведены и результаты численного решения указанного уравнения, полученные В.Соболевым и И.Мининым, а также результаты, полученные Дж. Касти и Р. Калабой по методу инвариантного погружения для этой же задачи (см. [7], стр. 93–94). В графах 2–5 таблицы 1 даются значения приближенного решения $x_n(t)$ интегрального уравнения (5.1) в условиях (5.3) соответственно при $t = 0$; $t = 0{,}1$; $t = 0{,}2$; $t = 0{,}3$.

Таблица 1

	Приближенное решение $x_n(t)$ уравнения (5.1) в условиях (5.3)			
t	0	0,1	0,2	0,3
Решение Соболева и Минина	0,083	0,103	0,123	0,141
Инвариантное погружение	0,082 559	0,102 958	0,122 851	0,140 966
Метод (1.4)-(1.5) $n = 3$	0,082 523	0,102 920	0,122 789	0,140 914
$n = 6$	0,082 548	0,102 951	0,122 826	0,140 948
$n = 12$	0,082 558	0,102 960	0,122 837	0,140 962
$n = 24$	0,082 561	0,102 963	0,122 841	0,140 967
$n = 48$	0,082 561	0,102 963	0,122 841	0,140 968

Заметим, что численные результаты согласуются с соответствующими теоретическими оценками погрешности теоремы 1. Кроме того, проведенные расчеты показывают высокую точность метода (1.4)-(1.5) при малых значениях n. При $n = 12$ получается почти полное совпадение с результатами счета по методу инвариантного погружения, однако, вполне удовлетворительное приближение имеем и при $n = 3$.

Литература

1. Алберг Дж., Нильсон Э., Уолш Дж., Теория сплайнов и ее приложения. Москва, 1972.
2. Вайникко Г.М., Компактная аппроксимация операторов и приближенное решение уравнений. Тарту, 1970.
3. Вайникко Г., Педас А., О решении интегральных уравнений с логарифмической особенностью методом механических квадратур. Уч. зап. Тартуск. ун-та, 1971, 281, 201-210.
4. Вайникко Г., Педас А., Кусочно-линейная аппроксимация решения уравнения Милна. Тезисы докладов

республиканского симпозиума по методам решения нелинейных уравнений и задач оптимизации, Таллин, 1978, 23-24.

5. Вайникко Г., Педас А., О методе механических квадратур для решения интегральных уравнений со слабой особенностью. Материалы конф. методы алгебры и функциональн. анализа при исслед. семейств операторов, Тарту, 1978, 58-60.

6. Канторович Л.В., Крылов В.И., Приближенные методы высшего анализа. Москва-Ленинград, 1962.

7. Касти Дж., Калаба Р., Методы погружения в прикладной математике. Москва, 1976.

8. Педас А., О решении интегральных уравнений с логарифмической особенностью сплайн-коллокационным методом первого порядка. Уч. зап. Тартуск. ун-та, 1977, 431, 130-146.

9. Педас А., Кусочно-линейная аппроксимация решения интегрального уравнения со степенной особенностью в ядре. Тезисы докладов республиканского симпозиума по методам решения нелинейных уравнений и задач оптимизации, Таллин, 1978, 79-80.

10. Педас А., О приближенном решении интегральных уравнений со слабой особенностью. Канд. дисс., Тарту, 1978.

11. Соболев В.В., Перенос лучистой энергии в атмосферах свезд и планет. Москва, 1956.

Поступило
19 I 1979

PIECEWISE LINEAR APPROXIMATION OF SOLUTION OF INTEGRAL EQUATION WITH LOGARITHMICAL SINGULAR KERNEL

A. Pedas

Summary

In the paper the spline collocation method (1.4), (1.5) for integral equation (1.1) is investigated. The main result is contained in theorem 1.

Уч.зап.Тартуск.ун-та,1979,500,43-52

ОБ ОДНОЙ ВОЗМОЖНОСТИ ПОСТРОЕНИЯ ИТЕРАЦИОННЫХ МЕТОДОВ

И. Саарнийт

Тартуский государственный университет

I. Введение. В процессе приближенного решения задач вида

$$Ax = y \tag{1}$$

(интегральных уравнений, краевых задач и т.д.) можно выделить два этапа. На первом этапе строится последовательность (как правило, конечномерных) задач

$$A_n x_n = y_n\ ,\quad n = 0,1,\dots, \tag{2}$$

решения x_n^* которых при $n \to \infty$ (при увеличении размерности задачи) сходятся в каком-то смысле к решению x^* задачи (1). Второй этап состоит в решении одной (или нескольких) задач (2) достаточно большой размерности. Часто для этого используется итерационный метод вида

$$x_n^{i+1} = T_n x_n^i\ ,\quad i = 0,1,\dots. \tag{3}$$

В данной статье рассматривается итерационный метод, в котором оба этих этапа в некотором смысле соединены, а именно, в ходе итерационного процесса размерность задачи увеличивается. Рассматриваемый метод является обобщением итерационного метода, рассмотренного в [7]. В статье уточняются результаты, сформулированные ранее в [3].

2. Описание аппроксимационно-итерационного метода. Пусть даны метрические пространства E и E_n, $n = 0,1,\dots$, с метриками $|\cdot,\cdot|$ и $|\cdot,\cdot|_n$ соответственно, а также система $\mathcal{R} = \{r_n\}$, $n = 0,1,\dots$, связывающих отображений $r_n: E \to E_n$. Будем говорить, что последовательность $\{x_n\}$, $n = 0,1,\dots$, с $x_n \in E_n$ $\mathcal{R}$-сходится к $x \in E$, если $\lim |x_n, r_n x|_n = 0$, обозначаем это $\mathcal{R}\text{-}\lim x_n = x$ (подробнее см. [2]).

Пусть $T_n: E_n \to E_n$, $n = 0,1,\dots$, — некоторые отображения метрических пространств E_n в себя, а $\vartheta_n: E_{n-1} \to E_n$, $n = 1,2,\dots$, — отображения перехода из пространств E_{n-1} в пространстве E_n. Пусть в пространстве E_0 задано начальное приближение x_0^0. Итерационный метод

$$z_n = (T_n)^{\alpha_n} \vartheta_n z_{n-1}, \; n=0,1,\ldots; \; \vartheta_0 z_{-1} = x_0^0 \in E_0, \qquad (4)$$

где $\alpha_n \geqslant 0$ – целые числа, будем называть аппроксимационно-итерационным методом (АИМ).

Частным случаем АИМ является проекционно-итерационный метод ($E_{n-1} \subset E_n \subset E$, $\vartheta_n = I$ – единичное отображение в E, $n=1,2,\ldots$, см. напр. [1], [6]). Если кроме r_n даны отображения $p_n : E_n \to E$, то можно определить $\vartheta_n = r_n p_{n-1}$. Такой АИМ рассматривался в [7].

3. Сходимость АИМ. Опишем условия, налагаемые на отображения T_n и ϑ_n. Предположим, что

существуют $x_n^* = T_n x_n^*$, $n=0,1,\ldots$, $\mathcal{R}\text{-}\lim x_n^* = x^* \in E$, (5)

$$|T_n x_n, T_n x_n^*|_n \leq \omega_n(|x_n, x_n^*|_n) \text{ для } \forall x_n \in E_n : |x_n, x_n^*| < \rho_n, \quad (6)$$

где $\omega_n(s)$ – определенные на интервале $[0, \rho_n)$ функции, $0 \leq \omega_n(s) < \rho_n$, $0 < \rho_n \leq \infty$. Пусть

функции $\omega_n(s)$ – неубывающие на $[0, \rho_n)$; (7)

$$\lim_{i\to\infty} \omega_n^i(s) = 0 \text{ при } \forall s \in [0, \rho_n). \qquad (8)$$

Здесь $\omega_n^i(s)$ – i-я итерация функции ω_n в точке s: $\omega_n^0(s) = s$, $\omega_n^i(s) = \omega_n(\omega_n^{i-1}(s))$, $i>0$. В частности, $\omega_n^i(s)$ – тоже неубывающие на $[0, \rho_n)$ функции. Заметим, что условия (7), (8) гарантируют сходимость итерационного процесса (3), если $|x_n^0, x_n^*|_n < \rho_n$ (см. напр., [5]). Пусть

$$|x_0^0, x_0^*|_0 \leq d_0 < \rho_0. \qquad (9)$$

Об отображениях ϑ_n предположим, что

$$|\vartheta_n x_{n-1}, \vartheta_n y_{n-1}|_n \leq K_n |x_{n-1}, y_{n-1}|_{n-1}, \qquad (10)$$

$x_{n-1}, y_{n-1} \in E_{n-1}$, $K_n = const$,

$$|x_n^*, \vartheta_n x_{n-1}^*|_n \leq d_n < \rho_n, \; n>1. \qquad (11)$$

Какие требования на отображения ϑ_n налагает последнее условие, видно подробнее из неравенства

$$|x_n^*, \vartheta_n x_{n-1}^*|_n \leq |x_n^*, r_n x^*|_n + |r_n x^*, \vartheta_n r_n x^*|_n + K_n |x_{n-1}^*, r_{n-1} x^*|_{n-1} \quad (12)$$

Теорема 1. Пусть выполняются условия (5)–(11). Тогда существует последовательность $\{\alpha_n\}$, $\alpha_n \geq 0$, $n=0,1,\ldots$, такая, что АИМ (4) сходится, т.е.

$$\mathcal{R}\text{-}\lim z_n = x^* .$$

Более того, для любой последовательности $\{\varepsilon_n\}$, $\varepsilon_n > 0$, $n=0,1,\ldots$, $\lim \varepsilon_n = 0$ найдется такой АИМ (4) (т.е. последовательность $\{\alpha_n\}$), что

$$|z_n, r_n x^*|_n \leq |x_n^*, r_n x^*|_n + \varepsilon_n , \quad n \geq 0.$$

Доказательство. Ввиду неравенства

$$|z_n, r_n x^*|_n \leq |z_n, x_n^*|_n + |x_n^*, r_n x^*|_n$$

нам достаточно доказать существование такой последовательности $\{\alpha_n\}$, что

$$|z_n, x_n^*|_n \leq \varepsilon_n , \quad n=0,1,\ldots . \tag{13}$$

Выберем последовательность $\{\eta_n\}$ так, чтобы

$$\eta_n < \frac{\rho_{n+1} - d_{n+1}}{K_{n+1}} , \quad \eta_n \leq \varepsilon_n , \quad n=0,1,\ldots . \tag{14}$$

Ввиду условия (9) и сходимости (8) существует α_0 такое, что

$$|z_0, x_0^*|_0 = |(T_0)^{\alpha_0} x_0^0, (T_0)^{\alpha_0} x_0^*|_0 \leq \omega_0^{\alpha_0}(|x_0^0, x_0^*|_0) \leq$$
$$\leq \omega_0^{\alpha_0}(d_0) \leq \eta_0 .$$

Пусть теперь $|z_{n-1}, x_{n-1}^*|_{n-1} \leq \eta_{n-1}$. Тогда

$$|z_n, x_n^*|_n = |(T_n)^{\alpha_n} \vartheta_n z_{n-1}, (T_n)^{\alpha_n} x_n^*|_n \leq \omega_n^{\alpha_n}(|\vartheta_n z_{n-1}, x_n^*|_n)$$
$$\leq \omega_n^{\alpha_n}(K_n |z_{n-1}, x_{n-1}^*|_{n-1} + |\vartheta_n x_{n-1}^*, x_n^*|_n) \leq$$
$$\leq \omega_n^{\alpha_n}(K_n \eta_{n-1} + d_n).$$

Согласно неравенству (14) $K_n \eta_{n-1} + d_n < \rho_n$ и поэтому существует α_n такое, что $|z_n, x_n^*|_n \leq \eta_n$. Неравенство (13) следует теперь из второго неравенства в (14). Теорема доказана.

Следствие 1. Доказательство теоремы дает нам рекуррентную оценку погрешности АИМ:

$$|z_n, r_n x^*|_n \leq |x_n^*, r_n x^*|_n + \delta_n , \quad n=0,1,\ldots ,$$

где

$$\delta_0 = \omega_0^{\alpha_0}(d_0), \quad \delta_n = \omega_n^{\alpha_n}(K_n \delta_{n-1} + d_n), \quad n \geq 1 .$$

При этом надо иметь ввиду, что числа δ_n должны удовлетворять неравенствам $\delta_n < (\rho_{n+1} - d_{n+1})/K_{n+1}$.

2. Если АИМ (4) сходится при последовательности $\{\alpha_n\}$, то он сходится и при последовательности $\{\alpha_n'\}$, $\alpha_n' \geq \alpha_n$, $n=0,1,2,\ldots$.

4. <u>Сходимость АИМ с постоянным числом итерации в каждом пространстве E_n</u>. Рассмотрим АИМ

$$z_n = (T_n)^{\alpha} \vartheta_n z_{n-1},\ n=0,1,\ldots;\ \vartheta_0 z_{-1} = x_0^0 \in E_0, \tag{15}$$

где $\alpha = const$ не зависит от n.

Усилим требования, налагаемые на T_n и ϑ_n. Пусть

$$|T_n x_n, T_n x_n^*|_n \leq \omega(|x_n, x_n^*|_n) \text{ для } \forall x_n \in E_n : |x_n, x_n^*|_n < \rho, \tag{6'}$$

где $\omega(s)$ – определенная на $[0,\rho)$ функция, $0 \leq \omega(s) < \rho$, $0 < \rho \leq +\infty$, удовлетворяющая условиям (7) и (8). Кроме того, пусть

$$\omega(s) \text{ непрерывна справа на } [0,\rho). \tag{16}$$

Заметим (напр., [4], [5]), что в этом случае условие (8) равносильно условию

$$\omega(s) < s, \text{ если } s \in (0,\rho).$$

Заменим неравенство (10) тоже более сильным

$$|\vartheta_n x_{n-1}, \vartheta_n y_{n-1}|_n \leq K |x_{n-1}, y_{n-1}|_{n-1}, \tag{10'}$$

$x_{n-1}, y_{n-1} \in E_{n-1}$, $K = const$ не зависит от n.

<u>Теорема 2</u>. Пусть выполняются условия (5), (6'), (7)–(9), (10'), (11) и (16). Пусть, кроме того,

$$\lim d_n = 0 \tag{17}$$

и существуют числа $\sigma \leq \rho$ и $\beta \geq 0$ такие, что

$$K\omega^{\beta}(s) < s \text{ при } s \in (0,\sigma). \tag{18}$$

Тогда существует число $\alpha \geq \beta$, при котором АИМ (15) сходится. При достаточно больших n АИМ (15) будет сходится с $\alpha = \beta$.

<u>Замечание 1</u>. Неравенство (12) дает нам следующее достаточное условие для сходимости (17):

$$\lim |z_n x^*, \vartheta_n z_{n-1} x^*|_n = 0.$$

<u>Доказательство теоремы 2</u>. Рассмотрим сперва АИМ вида (4). Введем обозначения

$$\delta_0 = \omega^{\alpha_0}(d_0),\ \delta_n = \omega^{\alpha_n}(K\delta_{n-1} + d_n),\ n \geq 1. \tag{19}$$

Согласно следствию 1 теоремы 1 АИМ (4) будет сходится, если

$$\delta_n < \frac{\rho - d_{n+1}}{k}, \quad n \geq 0, \tag{20}$$

и $\lim \delta_n = 0$.

Зафиксируем число $\sigma' < \sigma$ и обозначим

$$\zeta = \sigma' - k\,\omega^\beta(\sigma') > 0. \tag{21}$$

Из сходимости (17) следует существование такого числа $m = m(\sigma') \geq 0$, что

$$d_n \leq \zeta \quad \text{при} \quad n \geq m. \tag{22}$$

Выберем $\alpha_0, \ldots, \alpha_{m-1}$ так, чтобы выполнялись неравенства (20). Число α_m определим исходя из неравенства

$$\delta_m \leq \frac{\sigma' - \zeta}{k}. \tag{23}$$

Заметим, что ввиду выбора числа σ' и неравенства (22) величина δ_m будет тоже удовлетворять соотношению (20). При $n > m$ положим $\alpha_n = \beta$. Докажем, что неравенство (23) (а тем самым и (20)) распространится и на эти n, т.е.

$$\delta_n \leq (\sigma' - \zeta)/k, \quad n \geq m. \tag{24}$$

Действительно, при $n = m$ неравенство (24) выполняется. Если же (24) выполняется при каком-нибудь n, то ввиду (22) мы имеем, что

$$\delta_{n+1} = \omega^\beta(k\delta_n + d_{n+1}) \leq \omega^\beta(\sigma' - \zeta + d_{n+1}) \leq \omega^\beta(\sigma'),$$

откуда согласно равенству (21) следует, что неравенство (24) выполняется и при $n+1$.

Покажем теперь, что $\lim \delta_n = 0$. Для этого выберем подпоследовательность $\{\delta_{n_j}\}$ сходящуюся к конечному верхнему пределу δ^* последовательности $\{\delta_n\}$. Этот предел будет удовлетворять условию

$$\delta^* \leq (\sigma' - \zeta)/k < \sigma/k$$

Предположим, что последовательность $\{\delta_n\}$ не сходится к нулю, т.е. $\delta^* > 0$. Применив в точке $s = k\delta^* < \sigma$ неравенство (18), мы получаем, что

$$k\,\omega^\beta(k\delta^*) < k\delta^*, \tag{25}$$

откуда следует, что при достаточно больших j

$$k\delta_{n_j} = k\,\omega^\beta(k\delta_{n_j - 1} + d_{n_j}) > k\,\omega^\beta(k\delta^*).$$

Так как функция $\omega^\beta(s)$ не убывает на $[0, \rho)$, то последнее неравенство возможно лишь при

$$K\delta_{n_j-1} + d_{n_j} > K\delta^*.$$

Отсюда следует, что последовательность $\{K\delta_{n_j-1}\}$ и $\{K\delta_{n_j-1}+d_{n_j}\}$ сходятся к пределу $K\delta^*$, причем последняя последовательность стремится к этому пределу справа. Ввиду непрерывности справа функции $\omega^\beta(\delta)$ отсюда следует, что

$$K\delta^* = \lim K\delta_{n_j} = \lim K\omega^\beta(K\delta_{n_j-1}+d_{n_j}) = K\omega^\beta(K\delta^*).$$

Мы получили противоречие с неравенством (25), а тем самым и с предположением, что $\delta^*>0$. Таким образом построенный нами АИМ сходится. Ввиду следствия 2 теоремы 1 сходиться будет и АИМ (15), если положить

$$\alpha = \max\{\alpha_1, \dots, \alpha_m, \beta\}.$$

Теорема доказана.

5. Скорость сходимости АИМ. Пусть нам известна оценка

$$|x_n^*, \tau_n x^*|_n < \varepsilon_n, \quad n \geq 0,$$

где $\lim \varepsilon_n = 0$. Изучим, когда погрешность АИМ (15) имеет такой же порядок, т.е.

$$|z_n, \tau_n x^*|_n = O(\varepsilon_n).$$

Теорема 3. Пусть выполняются условия (5), (6'), (7)-(9), (10'), (16) а также условие (11) с

$$d_n = d\cdot\varepsilon_n, \quad n \geq 1, \quad d = const. \tag{26}$$

Пусть далее

$$\varepsilon_n > \eta\cdot\varepsilon_{n-1}, \quad n\geq 1, \quad \eta = const, \quad 0<\eta<1. \tag{27}$$

и существуют числа $\sigma' \leq \varrho$, $\beta \geq 0$ и $K_1 > K/\eta$ такие, что

$$K_1\omega^\beta(\delta) \leq \delta, \quad \text{если} \quad \delta\leq\sigma. \tag{28}$$

Тогда существуют $\alpha \geq \beta$ и $c = const$, что

$$|z_n, \tau_n x^*|_n \leq c\,\varepsilon_n, \quad n\geq 0. \tag{29}$$

Замечание 2. Из неравенства (12) следует, что достаточным условием для (26) будет

$$|\tau_n x^*, \vartheta_n \tau_{n-1} x^*|_n \leq \gamma\cdot\varepsilon_n, \quad n\geq 1, \quad \gamma = const.$$

Доказательство теоремы 3. Зафиксируем число $\sigma' < \sigma$ и построим такой же АИМ вида (4), как в доказательстве теоремы 2. Ввиду неравенства (28) выполняется условие (18) теоремы 2 и АИМ сходится, причем для определенных равенствами (19) величин δ_n выполняются неравенства (24).

Введем последовательность $\{c_n\}$, $n = 0, 1, \dots,$

$$c_n = \begin{cases} \delta_n/\varepsilon_n & \text{при } 0 \le n \le m, \\ \frac{1}{K_1}\left(\frac{K}{\eta}c_{n-1} + d\right) & \text{при } n > m. \end{cases}$$

Покажем, что $\delta_n \le c_n \varepsilon_n, n \ge 0$. Действительно, при $n \le m$ это неравенство выполняется. Если же $\delta_n \le c_n \varepsilon_n$ при некотором $n \ge m$, то согласно условиям (29) и (28) такое неравенство будет выполняется и при $n+1$:

$$\delta_{n+1} = \omega^\beta(K\delta_n + d_{n+1}) \le (K\delta_n + d_{n+1})/K_1 \le$$
$$\le (K c_n \varepsilon_n + d\varepsilon_{n+1})/K_1 \le c_{n+1}\varepsilon_{n+1}.$$

Для завершения доказательства теоремы нам достаточно доказать, что последовательность $\{c_n\}$ равномерно ограничена. Предположим обратное. Тогда найдутся c_n и c_{n+1}, $n \ge m$ такие, что

$$c_{n+1} > c_n > d\eta/(K_1\eta - K) > 0.$$

Но из $c_{n+1} > c_n$ следует, что

$$c_n(1 - K/(\eta K_1)) < d/K_1, \quad \text{т.е.} \quad c_n < d\eta/(K_1\eta - K).$$

Таким образом для построенного нами АИМ (4) справедлива оценка (29). Такая же оценка будет справедлива и для АИМ (15), если положить

$$\alpha = \max\{\alpha_1, \dots, \alpha_m, \beta\}.$$

Теорема доказана.

Следствия 1. Пусть $\rho = \infty$. Тогда АИМ (15) будет сходится при $\alpha = \beta$, где β определяется из условия (18). Если же определить $\alpha = \beta$ из условия (28), то АИМ (15) будет сходится со скоростью $O(\varepsilon_n)$.

2. Пусть функция $\omega^\beta(s)$ удовлетворяет добавочно условию: для любых $\alpha \ge 0$ и $\tau \in [0, \rho)$ из неравенства $\varkappa\omega^\alpha(\tau) \le \tau$ ($\varkappa$ - некоторая постоянная) следует $\varkappa\omega^\alpha(s) \le s$ при всех $s \in [0, \tau]$. Тогда существует α, при котором для АИМ (15) имеет место оценка (29).

3. Пусть функция $\omega^\beta(s)$ при некотором $\beta \ge 1$ удовлетворяет условию: для любой постоянной $\varkappa > 0$ существует $\tau > 0$ такое, что

$$\varkappa\omega^\beta(s) \le s \quad \text{при} \quad 0 \le s \le \tau.$$

Тогда, начиная с достаточно большого n, АИМ (4) с $\alpha_n=\beta$ будет сходится со скоростью порядка $O(\varepsilon_n)$.

6. Пример. Рассмотрим интегральное уравнение

$$x(t)=\int_0^1 ts\,x(s)\,ds+e^t-t$$

с точным решением $x^*(t)=e^t$. Построим АИМ на базе метода механических квадратур, например, используя формулу трапеций. Приближенное решение уравнения будем искать в пространстве E_n сеточных функций $x_n(t),\ t\in\Omega_n=\{ih_n\}\subset[0,1]$, где $h_n=1/(5\cdot 2^n)$, $n=0,1,\dots$. Метрику в E_n определим равенством

$$|x_n,y_n|_n=\max_{t\in\Omega_n}|x_n(t)-y_n(t)|.$$

Неравенство (6') будет тогда выполняться с функцией $\omega(s)=s/2$, а $\varrho=\infty$. Пусть $E=C[0,1]$, а $(\tau_n x)(t)=x(t),\ t\in\Omega_n$. Известно (напр., [2]), что погрешность метода механических квадратур $|x_n^*,\tau_n x^*|_n$ в данном примере будет иметь порядок $O(h_n^2)$.

Для перехода из пространств E_{n-1} в пространства E_n используем два различных оператора:

$$(\vartheta_n^1 x_{n-1})(t)=\begin{cases} x_{n-1}(t), & t\in\Omega_{n-1},\\ [x_{n-1}(t-h_n)+x_{n-1}(t+h_n)]/2, & t\in\Omega_n\setminus\Omega_{n-1}\end{cases}$$

$$(\vartheta_n^2 x_{n-1})(t)=\begin{cases} x_{n-1}(t), & t\in\Omega_{n-1},\\ x_{n-1}(t-h_n), & t\in\Omega_n\setminus\Omega_{n-1}.\end{cases}$$

Оператор ϑ_n^1 удовлетворяет условию (10') с $K=1$, а расстояние $|\tau_n x^*,\vartheta_n^1\tau_{n-1}x^*|_n$ оценивается величиной $O(h_n^2)$. Согласно вышеизложенному АИМ (15) будет сходится при $\alpha>1$. При $\alpha\geq 3$ скорость сходимости будет порядка $O(h_n^2)$. Расстояние $|\tau_n x^*,\vartheta_n^2\tau_{n-1}x^*|_n$ оценивается величиной порядка $O(h_n)$. Такой же порядок будет иметь согласно теореме 3 и скорость сходимости АИМ.

Результаты вычислений приводятся в таблице. Начальным приближением x_0^0 выбирался нулевой вектор.

n	h_n	Погрешность $\|z_n, r_n x^*\|_n$ АИМ (15)				
		$\vartheta_n = \vartheta_n^1$			$\vartheta_n = \vartheta_n^2$	
		$\alpha = 1$	$\alpha = 3$	$\alpha = 10$	$\alpha = 3$	$\alpha = 10$
0	0.2	1.00000	0.09581	0.02232	0.09581	0.02232
1	0.1	0.32882	0.00202	0.00555	0.00733	0.00556
2	0.05	0.10820	0.00148	0.00138	0.00309	0.00139
3	0.025	0.03569	0.00040	0.00034	0.00190	0.00034
4	0.0125	0.01181	0.00009	0.00008	0.00085	0.00008

Результаты вычислений согласуются с теоретическими выкладками. Пример показывает, что скорость сходимости АИМ довольно существенно зависит от выбора оператора ϑ_n. Эта зависимость исчезает лишь при достаточно больших α, т.е. тогда, когда z_n практически совпадает с x_n^* (при $\alpha = 10$).

В заключение отметим, что для достижения итерационным методом (3) такой же точности как АИМ (15) при $\alpha = 3$ (погрешность полученного приближения практически совпадает с погрешностью квадратурного метода) потребовалось при том же начальном приближении в три раза больше арифметических операций.

Литература

I. Балашова С.Д., Федоренко Ю.Д., Проекционно-итерационный принцип решения нелинейных уравнений, основанный на линеаризации. - В сб.: Вычислительная и прикладная математика, К., Изд-во КГУ, 1975, вып. 25, с. 102-110.

2. Вайникко Г., Анализ дискретизационных методов. Тарту, 1976.

3. Саарниийт И., О сходимости аппроксимационно-итеративного метода. Материалы конф. методы алгебры и функциональн. анализа при исслед. семейств операторов. Тарту, 1978, 68-70.

4. Browder, F.E., On the convergence of successive approximations for nonlinear functional equations. "Indagationes math.", 1968, 30, № 1, 27-35.

5. F u r i, M., Un teorema di punto fisso per transformazioni di uno spazio metrico completo in se. "Atti Acad. naz. Lincei, Rend. Cl. sci. fis., mat. e natur.", 1968, 45, № 5, 207-211.
6. G a j e w s k i, H., K l u g e, R., Projektions-Iterationsverfahren und nichtlineare Probleme mit monotonen Operatoren. Monatsber. DAW, 12, 1970, 98-115.
7. K l u g e, R., On a class of iteration methods. "Math. Nachr.", 1976, 73, 7-18.

Поступило
7 III 1979

ÜBER EINE KONSTRUKTIONSMÖGLICHKEIT DES ITERATIONSVERFAHRENS

I. Saarniit

Zusammenfassung

Im Artikel wird das sogenannte Approximations-Iterationsverfahren untersucht. Bei diesem Verfahren ist das Iterieren mit dem Übergang aus einem metrischen Raum in den anderen verbunden (die Dimension des Raumes wächst dabei). Im Artikel wird die Konvergenz des Verfahrens bewiesen und die Konvergenzgeschwindigkeit abgeschätzt.

Уч. зап. Тартуск. ун-та, 1979, 500, 53-68

МЕТОД ПОДОБЛАСТЕЙ ДЛЯ ПЕРИОДИЧЕСКОЙ ЗАДАЧИ

Р. Керге

Тартуский государственный университет

Сходимость метода подобластей для обыкновенных дифференциальных уравнений исследована в [13] и [14], а обыкновенных интегродифференциальных уравнений - [6] и [11]. В [8] уточнялись оценки быстроты сходимости в случае узлов Чебышева 1 рода и результаты были перенесены на нелинейные задачи. Метод подобластей в проблеме собственных значений был исследован в [9], а устойчивость метода подобластей была рассмотрена в [10].

В настоящей статье изучается метод подобластей в случае периодической задачи, притом рассматриваются также проблема собственных значений и вопрос устойчивости.

§ 1. Сходимость метода подобластей для линейного периодического дифференциального уравнения

Обозначим через $\mathcal{L}_p$, $p>1$, пространство 2π-периодических функций, через $\mathcal{C}$ -пространство 2π-периодических непрерывных функций.

Рассмотрим линейное уравнение

$$Lx \equiv x^{(m)}(t) + Bx = y(t), \tag{1}$$

где $y(t) \in \mathcal{L}_p$, $p>1$, а B - дифференциальный оператор

$$Bx \equiv \sum_{j=0}^{m-1} p_j(t) x^{(j)}(t)$$

с коэффициентами $p_j \in \mathcal{L}_p$, $p>1$ $(j=0,1,\ldots,m-1)$. Разыскивается 2π-периодическое решение, т.е. удовлетворяющее краевым условиям

$$x(0) = x(2\pi),\ x'(0) = x'(2\pi), \ldots,\ x^{(m-1)}(0) = x^{(m-1)}(2\pi). \tag{2}$$

Приближенное решение задачи (1), (2) будем искать в виде тригонометрического полинома

$$x_n(t) = a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt), \tag{3}$$

причем коэффициенты $a_k = a_k^{(n)}$, $b_k = b_k^{(n)}$ определим по методу подобластей, т.е. из системы уравнений

$$\int_{t_i}^{t_{i+1}} (Lx_n - y)\,dt = 0 \quad (i=0,\dots,2n). \tag{4}$$

Узлы $t_i = 2i\pi/(2n+1)$, $i=0,\dots,2n+1$ здесь и далее предполагаются равноотстоящими.

Справедлива следующая теорема.

<u>Теорема 1</u>. Пусть однородная задача $Lx = 0$ имеет в $\mathcal{L}_p$, $p>1$, лишь нулевое решение. Тогда задача (1), (2) имеет единственное решение $x^*(t)$, причем $x^* \in C^{m-1}$, $x^{*(m)} \in \mathcal{L}_p$, $p>1$, а задача (3), (4) имеет при $n \geq n_0$ единственное решение $x_n^*(t)$. Последовательность решений $x_n^*(t)$ стремится к $x^*(t)$ со скоростью

$$\|x_n^* - x^*\|_{C^{m-1}} \leq c(1+\ln n)e_n(x^{*(m-1)}). \tag{5}$$

Если $y(t)$, $p_j(t) \in C$ $(j=0,\dots,m-1)$, то $x^{*(m)} \in C$ и справедлива оценка

$$\|x_n^* - x^*\|_{C^m} \leq c(1+\ln n)e_n(x^{*(m)}). \tag{6}$$

Здесь

$$e_n(z) = \inf_{a_0,\dots,b_n} \|z(t) - a_0 - \sum_{k=1}^{n} a_k\cos(kt) + b_k\sin(kt)\|_C.$$

<u>Доказательство</u> теоремы опирается на проверку условий общей теоремы (см. [2], стр. 47), которую приводим лишь с нужной нам степенью общности.

<u>Теорема 2</u>. Пусть A и A_n $(n=1,2,\dots)$ – линейные непрерывные операторы из банахова пространства E в банахово пространство F, причем уравнение $Ax=0$ имеет лишь нулевое решение, а операторы A_n фредгольмовы с нулевым индексом. Пусть $\|A_n x - Ax\| \to 0$ при $n \to \infty$ для любого $x \in E$ и пусть выполнено следующее условие регулярности:
$\|x_n\| \leq const$, $\{A_n x_n\}$ компактна в $F \Rightarrow \{x_n\}$ компактна в E. Тогда операторы A, A_n $(n \geq n_0)$ обратимы, и, если $\|v_n - v\| \to 0$ при $n \to \infty$, то решения x_n^* уравнений $A_n x_n = v_n$ сходятся к решению x^* уравнения $Ax = v$ со скоростью

$$c_1\|A_n x^* - v_n\|_F \leq \|x_n^* - x^*\|_E \leq c_2\|A_n x^* - v_n\|_F.$$

Дадим задачам (1), (2) и (3), (4) операторный вид. Пусть функционал q сопоставляет каждой $z \in \mathcal{L}_p$ число $qz = (2\pi)^{-1}\int_0^{2\pi} z(s)\,ds$ и пусть оператор Q таков, что $Qz = z - qz$.

Пусть $(\mathcal{J}z)(t)=\int_0^t z(s)\,ds$. Легко проверить, что $\mathcal{J}Qz$ будет периодической функцией и что $\mathcal{J}$ – непрерывный оператор из $\mathcal{L}_p$ в $\mathcal{C}^{\gamma}$, $\gamma=1-1/p$.

Задача (1), (2) эквивалентна операторному уравнению

$$Ax=v\,, \tag{7}$$

где

$$A=(\mathcal{J}QL,\,qB)\,,$$
$$v=(\mathcal{J}Qy,\,qy)\,.$$

Оператор $\mathcal{J}QL$ действует из $\mathcal{C}^{m-1}$ в $\mathcal{C}_0$, $\mathcal{C}_0=\{z\in\mathcal{C}\mid z(0)=0\}$, а функционал qB действует из $\mathcal{C}^{m-1}$ в R.

Покажем, что задача (3) и (4) эквивалента операторному уравнению

$$A_n x_n = v_n\,, \tag{8}$$

где

$$A_n=(\mathcal{A}_n,\,qB)\,,$$
$$v_n=(P_n\mathcal{J}Qy,\,qy)\,.$$

Здесь P_n проектор, сопоставляющий каждой функции Qz её тригонометрический интерполяционный полином степени не выше n, построенный по значениям $(Qz)(t_i)$ в узлах t_i $(i=0,1,\dots,2n)$, а оператор $\mathcal{A}_n$ действует также из $\mathcal{C}^{m-1}$ в $\mathcal{C}_0$ по формуле

$$\mathcal{A}_n x_n = x_n^{(m-1)}(t)-x_n^{(m-1)}(0)+P_n\mathcal{J}QBx_n\,. \tag{9}$$

Для доказательства этой равносильности заметим, что суммируя условия (4), получаем

$$\sum_{i=0}^{2n}\int_{t_i}^{t_{i+1}}(Lx_n-y)\,dt=\int_0^{2\pi}(Lx_n-y)\,dt=0,$$

т.е. $q(Lx_n-y)=0$ и, так как $qx^{(m)}=0$, будет $qBx_n=qy$. Следовательно, для вторых компонент (8) справедливо.

Ввиду того, что $Q+q=I$ (I – единичный оператор), условия (4) можем переписать в виде

$$\int_{t_i}^{t_{i+1}}(q+Q)(Lx_n-y)\,dt=0\quad(i=0,1,\dots,2n)$$

и, следовательно, так как $q(Lx_n-y)=0$, то

$$\int_{t_i}^{t_{i+1}}Q(Lx_n-y)\,dt=0\qquad(i=0,\dots,2n) \tag{10}$$

Введем теперь оператор Ψ_n, который каждой функции $Qz \in \mathcal{L}_p$ относит тригонометрический полином степени $\le n$, так что

$$\int_{t_i}^{t_{i+1}} \Psi_n Q z\, dt = \int_{t_i}^{t_{i+1}} Q z\, dt \quad (i=0,\dots,2n).$$

Оператор Ψ_n связан с проектором P_n следующим соотношением:

$$\Psi_n Q = \frac{d}{dt} P_n \mathcal{J} Q. \qquad (11)$$

Действительно

$$\int_{t_i}^{t_{i+1}} \Psi_n Q z\, dt = \int_{t_i}^{t_{i+1}} \frac{d}{dt} P_n \mathcal{J} Q\, dt = (P_n \mathcal{J} Q z)(t_{i+1}) - (P_n \mathcal{J} Q z)(t_i)$$

$$= (\mathcal{J} Q z)(t_{i+1}) - (\mathcal{J} Q z)(t_i) = \int_{t_i}^{t_{i+1}} Q z\, dt \quad (i=0,\dots,2n),$$

что и доказывает равенство (11). Оператор Ψ_n является проектором на подпространстве $Q\mathcal{L}_p$:

$$\Psi_n^2 Q z = \frac{d}{dt} P_n \mathcal{J} \frac{d}{dt} P_n \mathcal{J} Q z = \frac{d}{dt} \mathcal{J} \frac{d}{dt} P_n \mathcal{J} Q z = \Psi_n Q z.$$

Итак, вместо (10) можем написать

$$\int_{t_i}^{t_{i+1}} \Psi_n Q (L x_n - y)\, dt = 0 \quad (i=0,\dots,2n),$$

откуда следует, что $\Psi_n Q (L x_n - y) = 0$. Поскольку $\Psi_n Q x_n^{(m)} = \Psi_n x_n^{(m)} = x_n^{(m)}$ имеем

$$x_n^{(m)} + \Psi_n Q B x_n = \Psi_n Q y,$$

откуда, проинтегрировав, получим

$$x_n^{(m-1)}(t) - x_n^{(m-1)}(0) + (P_n \mathcal{J} Q B x_n)(t) - (P_n \mathcal{J} Q B x_n)(0) =$$

$$= (P_n \mathcal{J} Q y)(t) - (P_n \mathcal{J} Q y)(0).$$

Из этого уравнения ввиду того, что $0 = t_0$ является узлом интерполяции и $(\mathcal{J} z)(0) = 0$, получаем

$$x_n^{(m-1)}(t) - x_n^{(m-1)}(0) + P_n \mathcal{J} Q B x_n = P_n \mathcal{J} Q y.$$

Сравнивая это уравнение с (8) и (9), видим, что уравнение (8) справедливо и для первых компонент.

Покажем теперь, что операторы A_n фредгольмовы с нулевым индексом. Для этого запишем $A_n x$ в следующем виде

$$A_n x = S x + T_n x,$$

где

$$S x = \left(x^{(m-1)}(t) - x^{(m-1)}(0),\ 0\right),$$

$$T_n x = \left(P_n \mathcal{J} Q B x,\ q B x\right).$$

Эти операторы действуют также из $\mathcal{C}^{m-1}$ в $\mathcal{C}_0\times R$. Поскольку сумма фредгольмова оператора F_1 и вполне непрерывного оператора F_2 является также фредгольмовым оператором, и $\operatorname{ind}(F_1+F_2)=\operatorname{ind} F_1$ (см. [7], стр. 300), то ввиду полной непрерывности T_n, нам придется исследовать лишь оператор S. Ясно, что область значений оператора S замкнута: $S\mathcal{C}^{m-1}=\mathcal{C}_0\times O$ и $\operatorname{ind} S=\operatorname{ind} A_n=0$.

Переходим к установлению регулярной сходимости последовательности операторов A_n к A. Поскольку вторые компоненты операторов совпадают, покажем лишь, что $\mathcal{A}_n x\to \mathcal{J}QLx$ при каждом $x\in\mathcal{C}^{m-1}$. Имеем

$$\|\mathcal{A}_n x-\mathcal{J}QLx\|_{\mathcal{C}}=\|x^{(m-1)}-x^{(m-1)}(0)+P_n\mathcal{J}QBx-\mathcal{J}Q(x^{(m)}+Bx)\|=$$

$$=\|P_n\mathcal{J}QBx-\mathcal{J}QBx\|_{\mathcal{C}}\le(1+\|P_n\|)\,d(\mathcal{J}QBx,\mathcal{T}_n),$$

где $d(z,\mathcal{T}_n)$ – расстояние в $\mathcal{C}$ от z до подпространства тригонометрических полиномов степени $\le n$. Поскольку для порядка роста нормы проектора P_n справедливо $\|P_n\|_{\mathcal{C}\to\mathcal{C}}=$ $=O(\ln n)$ (см. [3], стр. 31), то, продолжая с помощью теоремы Джексона, выводим

$$\|\mathcal{A}_n x-\mathcal{J}QLx\|_{\mathcal{C}}\le const\,\frac{\ln n}{n^{\gamma}}\longrightarrow 0\qquad (n\longrightarrow\infty).$$

Рассмотрим условие регулярности. Пусть $\|x_n\|_{\mathcal{C}^{m-1}}\le const$, $\{A_n x_n\}$ компактна в $\mathcal{C}$. Требуется установить компактность последовательности $\{x_n\}$ в $\mathcal{C}^{m-1}$. Последовательности $\{x_n^{(j)}\}$ $(j=0,1,\ldots,m-2)$ компактны по теореме Арцела. Далее, оператор B ограничен как оператор из $\mathcal{C}^{m-1}$ в $\mathcal{L}_p$ по предположению, а оператор $\mathcal{J}$ ограничен из $\mathcal{L}_p$ в $\mathcal{C}^{\gamma}$, $\gamma=1-1/p$. Следовательно, последовательность $\{\mathcal{J}QBx_n\}$ ограничена в $\mathcal{C}^{\gamma}$, вместе с тем она компактна в $\mathcal{C}^{\beta}$, $0<\beta<\gamma$ (см. [1], стр. 202). Далее, из теоремы Джексона вытекает, что $\|z-P_n z\|\to 0$ для любой $z\in\mathcal{C}^{\beta}$. Но тогда и $\|\mathcal{J}z_n-P_n\mathcal{J}z_n\|\to 0$ для каждой ограниченной в $Q\mathcal{L}_p$ последовательности $\{z_n\}$. Отсюда следует, что последовательность $\{P_n\mathcal{J}QBx_n\}$ компактна. Из (9) теперь видно, что $\{x_n^{(m-1)}\}$ компактна в $\mathcal{C}$ и, в итоге, $\{x_n\}$ компактна в $\mathcal{C}^{m-1}$, что нам и требовалось доказать.

Ещё требуется проверить сходимость свободных членов $v_n=(P_n\mathcal{J}Qy,qy)$ к $v=(\mathcal{J}Qy,qy)$, т.е. сходимость

$P_n\mathcal{J}Qy$ к $\mathcal{J}Qy$. Имеем

$$\|P_n\mathcal{J}Qy-\mathcal{J}Qy\|\le(1+\|P_n\|)\,d(\mathcal{J}Qy,\mathcal{T}_n)\le c\,\frac{\ln n}{n^{\delta}}\to 0,$$

так как $\mathcal{J}Qy\in\mathcal{C}^{\delta}$.

Итак, все условия теоремы 2 выполнены.

Оценим скорость сходимости:

$$\|x_n^*-x^*\|_{\mathcal{C}^{m-1}}\le$$

$$\le c_2\|A_nx^*-v_n\|_{\mathcal{C}_0\times R}=c_2\|(\mathcal{A}_nx^*,qBx^*)-(P_n\mathcal{J}Qy,qy)\|_{\mathcal{C}\times R}.$$

Поскольку $qBx^*=qy$, то эта скорость зависит только от первых компонент в паре:

$$\|x_n^*-x^*\|_{\mathcal{C}^{m-1}}\le$$

$$\le c_2\|\mathcal{A}_nx^*-P_n\mathcal{J}Qy\|_{\mathcal{C}}=c_2\|\mathcal{J}x^{*(m)}+P_n\mathcal{J}QBx^*-P_n\mathcal{J}QLx^*\|_{\mathcal{C}}=$$

$$=c_2\|\mathcal{J}x^{*(m)}+P_n\mathcal{J}QBx^*-P_n\mathcal{J}Qx^{*(m)}-P_n\mathcal{J}QBx^*\|_{\mathcal{C}}=$$

$$=c_2\|\mathcal{J}x^{*(m)}-P_n\mathcal{J}Qx^{*(m)}\|_{\mathcal{C}}=c_2\|\mathcal{J}x^{*(m)}-P_n\mathcal{J}x^{*(m)}\|_{\mathcal{C}}\le$$

$$\le c_2(1+\ln n)\,d(\mathcal{J}x^{*(m)},\mathcal{T}_n)\le c\,(1+\ln n)\,e_n(\mathcal{J}x^{*(m)}).$$

Оценка (5) установлена.

Установим оценку (6). Скорость $\|x_n^{*(m)}-x^{*(m)}\|$ также зависит только от первых компонент в паре. Пишем

$$\|x_n^{*(m)}-x^{*(m)}\|\le\|x_n^{*(m)}-\Psi_nx^{*(m)}\|+\|\Psi_nx^{*(m)}-x^{*(m)}\| \quad (12)$$

и оценим оба слагаемых отдельно.

Поскольку $(Q+q)x^{*(m)}=(Q+q)(y-Bx^*)$, $qx^{*(m)}=0$ и $q(y-Bx^*)=0$, то $\Psi_nx^{*(m)}=\Psi_nQ(y-Bx^*)$ и для первого слагаемого получим:

$$\|x_n^{*(m)}-\Psi_nx^{*(m)}\|_{\mathcal{C}}=\|\Psi_nQy-\Psi_nQBx_n^*-\Psi_nQy+\Psi_nQBx^*\|_{\mathcal{C}}$$

$$\le\|\Psi_n\|_{\mathcal{C}\to\mathcal{C}}\|Q\|_{\mathcal{C}\to\mathcal{C}}\|B\|_{\mathcal{C}^{m-1}\to\mathcal{C}}\|x_n^*-x^*\|_{\mathcal{C}^{m-1}}\le$$

$$\le\text{const}\,\|\Psi_n\|_{\mathcal{C}\to\mathcal{C}}\|x_n^*-x^*\|_{\mathcal{C}^{m-1}}.$$

Для второго слагаемого в (12) имеем

$$\|x^{*(m)}-\Psi_nx^{*(m)}\|\le\text{const}\,(1+\|\Psi_n\|_{\mathcal{C}\to\mathcal{C}})\,d(x^{*(m)},\mathcal{T}_n).$$

Для Ψ_n получим из (11), что $\|\Psi_n\|=O(n\ln n)$, так как имеет имеет порядок роста $O(\ln n)$ для производной тригонометрического полинома $T_n(t)$ порядка n справедлива оценка $\|T_n'(t)\|_{\mathcal{C}}\le n\|T_n(t)\|$, а оператор $\mathcal{J}$, как оператор из $\mathcal{C}$ в $\mathcal{C}$, ограничен. В итоге можем для (12), учитывая (5), написать:

$$\| x_n^{*(m)} - x^{*(m)} \| \le c(1+n\ln n)[O(\ln n)e_n(\mathcal{J}x^{*(m)}) + e_n(x^{*(m)})] \le$$
$$\le c'(1+n\ln n)\,e_n(x^{*(m)}),$$

так как $e_n(\mathcal{J}x^{*(m)}) \le c\, e_n(x^{*(m)})/n$ (см. [12], стр. 119) и первое слагаемое является малым относительно второго слагаемого. Оценка (6) установлена. Теорема доказана.

Следствие 1. Если $x^* \in \mathcal{C}^{(m+r,\alpha)}$, $r \ge 0$, $\alpha \ge 0$, $r+\alpha > 0$, то по теореме Джексона из оценок (5) и (6) получим

$$\max_{0 \le t \le 2\pi} | x_n^{*(j)} - x^{*(j)}(t) | = O\left(\frac{\ln n}{n^{r+\alpha+1}}\right) \quad (j=0,\dots,m-1),$$

$$\max_{0 \le t \le 2\pi} | x_n^{*(m)}(t) - x^{*(m)}(t) | = O\left(\frac{\ln n}{n^{r+\alpha-1}}\right).$$

Заметим, что $x^* \in \mathcal{C}^{(m+r,\alpha)}$, если коэффициенты p_j $(j=0,1,\dots,m-1)$ и свободный член $y(t)$, например, принадлежат к классу $\mathcal{C}^{(r,\alpha)}$, $r \ge 0$, $\alpha \ge 0$, $r+\alpha > 0$.

Следствие 2. Поскольку последовательность операторов A_n сходится к A регулярно, $N(A) = \{0\}$ и операторы A_n фредгольмовы с нулевым индексом, то (см. [2], стр. 35) можно вывести, что операторы A_n сходятся к A и устойчиво, т.е. существуют обратные A_n^{-1} $(n \ge n_0)$, причём

$$\| A_n^{-1} \| \le const.$$

Следовательно справедливо неравенство

$$\| x_n \|_{\mathcal{C}^{m-1}} \le c \| (P_n \mathcal{J} Q y, \varphi y) \| \le c(\| P_n \mathcal{J} Q y \|_{\mathcal{C}} + |\varphi y|). \quad (13)$$

§ 2. Случай нелинейной задачи

Рассмотрим нелинейное периодическое дифференциальное уравнение

$$Lx \equiv x^{(m)}(t) + f(t, x(t), \dots, x^{(m-1)}(t)) = y(t), \quad (14)$$

которое решается методом подобластей (3), (4).

На основе общей теоремы из [2], стр. 54, справедлив следующий результат, доказательство которого мы опускаем ввиду аналогии с доказательством теоремы 1.

Теорема 3. Пусть

1) существует решение $x^*(t)$ задачи (14), (2);
2) свободный член $y(t)\in\mathcal{L}_p$, $p>1$;
3) функции $f(t,u_0,\ldots,u_{m-1})$, $\partial f/\partial u_j$ $(j=0,1,\ldots,m-1)$ определены и непрерывны при $0\le t\le 2\pi$,
$$|u_j - x^{*(j)}(t)|\le\delta \quad (j=0,1,\ldots,m-1),\ \delta>0;$$
4) линеаризованная задача
$$h^{(m)}+\sum_{j=0}^{m-1} q_j(t)\,h^{(j)}(t)=0$$
с
$$q_j(t)=\frac{\partial f}{\partial u_j}(t,x^*,x^{*\prime},\ldots,x^{*(m-1)}) \quad (j=0,1,\ldots,m-1)$$
имеет лишь нулевое периодическое решение $h(t)\equiv 0$.

Тогда найдется такое число δ_0, $0<\delta_0\le\delta$, что метод подобластей (3), (4) дает при $n\ge n_0$ в шаре $\|x-x^*\|_{C^{m-1}}\le\delta_0$ единственное решение $x_n^*(t)$. При этом быстрота сходимости x_n^* к x^* имеет оценку
$$\|x_n^*-x^*\|_{C^{m-1}}\le c\,(1+\ln n)\,e_n(x^{*(m-1)}),$$
а в случае $y\in C$ (тогда $x^*\in C^m$)
$$\|x_n^*-x^*\|_{C^m}\le c\,(1+n\ln n)\,e_n(x^{*(m)}).$$

§ 3. Проблема собственных значений

Рассмотрим проблему собственных значений для периодической задачи с нелинейным вхождением параметра:
$$L(\lambda)x\equiv x^{(m)}(t)+\sum_{j=0}^{m-1} p_j(t,\lambda)\,x^{(j)}(t)=0, \qquad (15)$$
решение которой ищем в виде (3), причем коэффициенты определяем из системы
$$\int_{t_i}^{t_{i+1}} L(\lambda)x_n(t)\,dt=0 \quad (i=0,1,\ldots,2n). \qquad (16)$$
Задача (15), (2) равносильна операторному уравнению
$$A(\lambda)x=0, \qquad (17)$$
где
$$A(\lambda)=(JQL(\lambda),\ qB(\lambda)).$$
Оператор $A(\lambda)$ действует при фиксированном λ из C^{m-1} в $C_0\times R$. Задача (3), (16) равносильна операторному уравнению (ср. (8))

$$A_n(\lambda)x_n = 0, \tag{18}$$

где $A_n(\lambda)$ действует также при фиксированном λ из $\mathcal{E}^{m-1}$ в $\mathcal{E}_0 \times R$ по формуле

$$A_n(\lambda) = (\mathcal{A}_n(\lambda), q\, B(\lambda))$$

$$\mathcal{A}_n(\lambda)x_n = x_n^{(m-1)}(t) - x_n^{(m-1)}(0) + P_n J Q\, B(\lambda) x_n .$$

Пусть $\Lambda \subset \mathbb{C}$ некоторая область комплексной плоскости. Обозначим через $S(A)$ спектр задачи (15), (2) (и, следовательно, задачи (17)) в Λ и через $S(A_n)$ спектр задачи (3), (16) (и равносильной задачи (18)) в Λ.

Справедлив следующий результат:

Теорема 4. Пусть

1) существует $\lambda \in \Lambda \subset \mathbb{C}$, при котором задача (15), (2) имеет лишь нулевое решение;
2) коэффициенты $p_j(t,\lambda) : R \times \Lambda \to \mathbb{C}$ $(j = 0, 1, \dots, m-1)$ голоморфны по λ при каждом t, $p_j(t,\lambda)$, $\partial p_j(t,\lambda)/\partial\lambda$ непрерывны по совокупности переменных $(t,\lambda) \in R \times \Lambda$ $(j = 0, \dots, m-1)$

Тогда для каждого $\lambda_0 \in S(A)$ существует последовательность $\{\lambda_n\}$, $\lambda_n \in S(A_n)$, $\lambda_n \to \lambda_0$ $(n \geq n_0)$, где λ_n из достаточно малого круга $|\lambda_n - \lambda_0| \leq \delta$ и справедливы оценки

$$|\lambda_n - \lambda_0| \leq c\, \varepsilon_n^{1/\varkappa},$$

$$|\hat{\lambda}_n - \lambda_0| \leq c\, \varepsilon_n,$$

$$d\left(x_n^0, N(A(\lambda_0))\right) \leq c \cdot \varepsilon_n^{1/\varkappa},$$

где

$$c = c(\delta) = \mathrm{const}, \quad \varepsilon_n \leq c \max_{x \in W, \|x\| = 1} (1 + \ln n)\, e_n(x^{*(m-1)}),$$

$\varkappa$ - наибольшая длина корневых цепочек задачи (15), (2), соответствующих λ_0, W - соответствующее корневое подпространство, $\hat{\lambda}_n$ - взвешенное арифметическое среднее собственных значений ,

$$N(A(\lambda_0)) = \{x \in \mathcal{E}^{m-1} \mid A(\lambda_0)x = 0\}.$$

Доказательство этой теоремы опирается на проверку условий теорем из [2], стр. 69, 72, и [5].

§ 4. Об устойчивости метода подобластей

Устойчивость т.н. прямых методов характеризуется двумя факторами. Она зависит, во первых, от влияния неточностей составления системы для определения коэффициентов, а во вторых – от распространения ошибок округления при практическом решении системы. Характеристикой первого фактора является устойчивость в смысле С.Г.Михлина, а второй фактор характеризуется числом обусловленности системы.

Перепишем условия (4) относительно коэффициентов $a_0, a_1, b_1, \dots, a_n, b_n$. Получаем систему линейных уравнений

$$\int_{t_i}^{t_{i+1}} p_0(t)dt \cdot a_0 + \sum_{k=1}^{n} \int_{t_i}^{t_{i+1}} [\cos^{(m)}(kt) + \sum_{j=0}^{m-1} p_j(t)\cos^{(j)}(kt)]dt\, a_k + \tag{19}$$
$$+ \sum_{k=1}^{n} \int_{t_i}^{t_{i+1}} [\sin^{(m)}(kt) + \sum_{j=0}^{m-1} \sin^{(j)}(kt)]dt\, b_k = \int_{t_i}^{t_{i+1}} y(t)dt.$$

Однозначная разрешимость этой системы вытекает из результатов §1. Однако, в вычислительной практике система (19) составляется с некоторыми погрешностями, в результате чего мы решаем не систему (19), а систему с некоторыми возмущениями:

$$\left(\int_{t_i}^{t_{i+1}} p_0(t)dt + \gamma_{i0}\right)\bar{a}_0 + \sum_{k=1}^{n} \left\{\int_{t_i}^{t_{i+1}} [\cos^{(m)}(kt) + \sum_{j=0}^{m-1} p_j(t)\cos^{(j)}(kt)]dt +\right.$$
$$\left. + \gamma_{ik}\right\}\bar{a}_k + \sum_{k=1}^{n} \left\{\int_{t_i}^{t_{i+1}} [\sin^{(m)}(kt) + \sum_{j=0}^{m-1} p_j(t)\sin^{(j)}(kt)]dt + \gamma_{i,n+k}\right\} \times$$
$$\times \bar{b}_k = \int_{t_i}^{t_{i+1}} y(t)dt + \delta_i, \quad j = 0, 1, \dots, 2n \tag{20}$$

и получаем не приближенное решение (3), а "неточное"

$$\bar{x}_n(t) = \bar{a}_0 + \sum_{k=1}^{n} (\bar{a}_k \cos kt + \bar{b}_k \sin kt). \tag{21}$$

Изучим вопрос устойчивости метода подобластей в смысле С.Г.Михлина, т.е. нас будет интересовать влияние возмущении $\gamma_{ik} = \gamma_{ik}^{(n)}$ и $\delta_i = \delta_i^{(n)}$ на $x_n(t) - \bar{x}_n(t)$. В частности, возникает вопрос о допустимых возмущениях, не портящих разрешимость системы (20).

Введем следующие обозначения:

$$a^{(n)} = (a_0, a_1, \dots, a_n, b_n) \in \mathbb{R}^{2n+1},$$
$$\bar{a}^{(n)} = (a_0, a_1, \dots, \bar{a}_n, \bar{b}_n) \in \mathbb{R}^{2n+1},$$
$$A_n = (a_{ik})_{i,k=0}^{2n} : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1},$$

$$a_{ik} = \begin{cases} \int_{t_i}^{t_{i+1}} p_0(t)\,dt & \text{при } k=0, \\ \int_{t_i}^{t_{i+1}} [\cos^{(m)} ht + \sum_{j=0}^{m-1} p_j(t)\cos^{(j)}(ht)]\,dt, & h=\frac{k+1}{2},\ k=2N-1, \\ \int_{t_i}^{t_{i+1}} [\sin^{(m)} ht + \sum_{j=0}^{m-1} p_j(t)\sin^{(j)}(ht)]\,dt, & h=\frac{k}{2},\ k=2N, \end{cases}$$

$\Gamma_n : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ составленная из погрешностей γ_{ik}.

$$\delta^{(n)} = (\delta_0, \delta_1, \ldots, \delta_{2n}) \in \mathbb{R}^{2n+1},$$

$$G^{(n)} z = (G_i z)_{i=0}^{2n} = \Big(\int_{t_i}^{t_{i+1}} z(t)\,dt\Big)_{i=0}^{2n} \in \mathbb{R}^{2n+1}.$$

Теперь можем выразить системы (19) и (20) соответственно в следующих видах:

$$A_n a^{(n)} = G^{(n)} y, \tag{22}$$

$$(A_n + \Gamma_n) a^{(n)} = G^{(n)} y + \delta^{(n)}. \tag{23}$$

Наша последовательность координатных функций $\{1, \cos t, \sin t, \ldots, \cos kt, \sin kt, \ldots\}$ ортогональная. Обозначим через H^ℓ гильбертово пространство периодических функций с нормой

$$\|u\|_{H^\ell} = \Big(\sum_{k=0}^{\ell} \|u^{(k)}\|_{L_2}^2\Big)^{1/2} = \Big(\sum_{k=0}^{\ell} \int_0^{2\pi} |u^{(k)}|^2 dt\Big)^{1/2}$$

и ортонормируем систему $\{1, \ldots, \cos kt, \sin kt, \ldots\}$ в H^ℓ. Имеем

$$\|\cos kt\|_{H^\ell} = \|\sin kt\|_{H^\ell} = \Big(\pi \sum_{i=0}^{\ell} k^{2i}\Big)^{1/2} =: d_k^\ell,$$

$$\|1\|_{H^\ell} = \sqrt{2\pi} =: d_0.$$

Следовательно система функций $\chi = \{1/d_0, \ldots, \cos kt/d_k^\ell, \ldots\}$ ортонормальная в H^ℓ. Обозначим через $\mathcal{D}_{n,\ell}$ диагональную матрицу, переобразующую систему функций $\{1, \ldots, \cos kt, \sin kt, \ldots\}$ в систему χ:

$$\mathcal{D}_{n,\ell} = \begin{pmatrix} 1/d_0 & & & & & & & 0 \\ & 1/d_1^\ell & & & & & & \\ & & 1/d_1^\ell & \cdots & & & & \\ & & & & 1/d_k^\ell & & & \\ & & & & & 1/d_k^\ell & \cdots & \\ & & & & & & 1/d_n^\ell & \\ 0 & & & & & & & 1/d_n^\ell \end{pmatrix}$$

Ясно, что[1] $\|\mathcal{D}_{n,\ell}\| = 1/d_0 = 1/\sqrt{2\pi}$.
Справедлива следующая теорема.

Теорема 5. Найдется такая постоянная C_0, что при

$$\|\Gamma_n\| \le C_0\sqrt{2\pi}/(\sqrt{n}\,\ln n)$$

возмущенная система (21) однозначно разрешима и

$$\|x_n - \bar{x}_n\|_{H^\ell} \le \frac{\sqrt{n}\,\ln n}{C_0}\left(\|\Gamma_n\mathcal{D}_{n,\ell}\|\,\|x_n\|_{H^\ell} + \|\delta^{(n)}\|\right),$$

где ℓ может иметь значения от 0 до $m-1$, а

$$(\Gamma_n\mathcal{D}_{n,\ell})_{ik} = \gamma_{ik}^{(n)}\left(\pi\sum_{j=0}^{\ell}\left[\frac{k+1}{2}\right]^{2j}\right)^{-1/2},$$

квадратные скобки означают целую часть.

Доказательство. Пусть $b^{(n)} = \mathcal{D}_{n,\ell}^{-1}a^{(n)}$, $\bar{b}^{(n)} = \mathcal{D}_{n,\ell}^{-1}\bar{a}^{(n)}$. Тогда нетрудно проверить, что

$$\|x_n\|_{H^\ell} = \|b^{(n)}\|, \tag{24}$$

$$\|x_n - \bar{x}_n\|_{H^\ell} = \|b^{(n)} - \bar{b}^{(n)}\|. \tag{25}$$

Перепишем теперь системы (22) и (23) относительно $b^{(n)}$:

$$A_n\mathcal{D}_{n,\ell}b^{(n)} = G^{(n)}y, \tag{26}$$

$$(A_n+\Gamma_n)\mathcal{D}_{n,\ell}\bar{b}^{(n)} = G^{(n)}y + \delta^{(n)}, \tag{27}$$

откуда получим соотношение

$$(A_n\mathcal{D}_{n,\ell} + \Gamma_n\mathcal{D}_{n,\ell})(b^{(n)} - \bar{b}^{(n)}) = \Gamma_n\mathcal{D}_{n,\ell}b^{(n)} - \delta^{(n)}. \tag{28}$$

Нам надо установить обратимость оператора $A_n\mathcal{D}_{n,\ell} + \Gamma_n\mathcal{D}_{n,\ell}$ и найти оценку его нормы. Для этого сперва оценим норму $\|(A_n\mathcal{D}_{n,\ell})^{-1}\|$ (обратимость оператора $A_n\mathcal{D}_{n,\ell}$ вытекает из однозначной разрешимости системы (19)). Из (26), (24) и (13) находим

$$\begin{aligned}\|(A_n\mathcal{D}_{n,\ell})^{-1}G^{(n)}y\| &= \|b^{(n)}\| = \|x_n\|_{H^\ell} \le c\,\|x_n\|_{C^\ell} \le \\ &\le c\,\|x_n\|_{C^{m-1}} \le c'\left(\|P_nJQy\|_C + |qy|\right).\end{aligned} \tag{29}$$

Оценим по частям:

$$\|P_nJQ\|_C \le \|P_n\|_{C\to C}\max_{0\le i\le 2n}|(JQ)(t_i)| \le$$

---

[1] Все нормы матриц и векторов берутся по евклидовой метрике.

$$\le c \ln n \max_{0\le i\le 2n} \left| \int_0^{t_i} (y - qy)\,dt \right| \le$$

$$\le c \ln n \left(\left| \int_0^{t_{i_0}} y\,ds \right| + \left| \int_0^{t_{i_0}} qy\,ds \right| \right),$$

далее, с помощью неравенства Коши-Буняковского:

$$\left| \int_0^{t_{i_0}} y\,ds \right| = \left| \sum_{i=0}^{i_0-1} \int_{t_i}^{t_{i+1}} y\,ds \right| \le \left(\sum_{i=0}^{i_0-1} \left| \int_{t_i}^{t_{i+1}} y\,ds \right|^2 \right)^{1/2} \cdot \left(\sum_{i=0}^{i_0-1} 1 \right)^{1/2} \le \sqrt{2n+1}\, \| G^{(n)} y \|.$$

$$\left| \int_0^{t_{i_0}} qy\,ds \right| \le |t_{i_0}| |qy| \le 2\pi \cdot \frac{1}{2\pi} \left| \int_0^{2\pi} y\,ds \right| = \left| \sum_{i=0}^{2n} \int_{t_i}^{t_{i+1}} y\,ds \right| \le \sqrt{2n+1}\, \| G^{(n)} y \|.$$

Из последней цепочки видно также, что $|qy| \le (2\pi)^{-1} \sqrt{2n+1}\, \| G^{(n)} y \|$.
В итоге мы можем продолжить цепочку (29)

$$\| (A_n \mathcal{D}_{n,\ell})^{-1} G^{(n)} y \| \le c \ln n \cdot 2 \cdot \sqrt{2n+1}\, \| G^{(n)} y \| + \frac{1}{2\pi} \sqrt{2n+1}\, \| G^{(n)} y \|.$$

Следовательно,

$$\| (A_n \mathcal{D}_{n,\ell})^{-1} \| \le c \sqrt{n} \ln n.$$

Выберем такое c_0, что $\| (A_n \mathcal{D}_{n,\ell})^{-1} \| \le \sqrt{n} \ln n / c_0$ и пусть для возмущений γ_{ik} справедливо предположение теоремы. Тогда по теореме о возмущениях обратимого оператора (см., например, [4], стр. 212), заключаем, что оператор $A_n \mathcal{D}_{n,\ell} + \Gamma_n \mathcal{D}_{n,\ell}$ также обратим, причем для интересующей нас нормы получаем

$$\| (A_n \mathcal{D}_{n,\ell} + \Gamma_n \mathcal{D}_{n,\ell})^{-1} \| \le \frac{\| (A_n \mathcal{D}_{n,\ell})^{-1} \|}{1 - \| (A_n \mathcal{D}_{n,\ell})^{-1} \Gamma_n \mathcal{D}_{n,\ell} \|} \le \frac{\sqrt{n} \ln n}{c_0}. \quad (30)$$

Ввиду обратимости оператора $A_n \mathcal{D}_{n,\ell} + \Gamma_n \mathcal{D}_{n,\ell}$ уравнение (27) (а следовательно и (23), и система (20) однозначно разрешимы. На основе (25), (28), (30) и (24) выводим, что

$$\| x_n - \bar{x}_n \|_{H^\ell} = \| b^{(n)} - \bar{b}^{(n)} \| \le \| (A_n \mathcal{D}_{n,\ell} + \Gamma_n \mathcal{D}_{n,\ell})^{-1} \| \left(\| \Gamma_n \mathcal{D}_{n,\ell} \| \| b^{(n)} \| + \right.$$

$$\left. + \| \delta^{(n)} \| \right) \le \frac{\sqrt{n} \ln n}{c_0} \left(\| \Gamma_n \mathcal{D}_{n,\ell} \| \cdot \| x_n \|_{H^\ell} + \| \delta^{(n)} \| \right).$$

Теорема доказана.

В доказательстве теоремы 5 существенно использована сходимость приближенных решений в пространстве C^{m-1}. Пусть теперь коэффициенты $p_j(t)$ $(j = 0, 1, \dots, m-1)$ и свободный член $y(t)$ непрерывны. Для m-ой производной оценка погрешности по порядку хуже, чем для младших производных. Возникает вопрос, какими могут быть возмущения, чтобы всё же $\bar{x}_n^{(m)}(t)$ остался близким к $x_n^{(m)}(t)$. Приведем соответствующую теорему без доказательства.

Теорема 6. Найдется такое C_m, что при

$$\|\Gamma_n\| \le C_m\sqrt{2\pi}/(n^{3/2}\ln^2 n)$$

справедлива оценка

$$\|x_n-\bar{x}_n\|_{H^m} \le \frac{n^{3/2}\ln^2 n}{C_m}\left(\|\Gamma_n \mathcal{D}_{m,\ell}\|\|x_n\|_{H^m}+\|\delta^{(n)}\|\right).$$

Следствие 3. Из теоремы 5 видно, что элементы a_{ik} матрицы A_n при больших k можно считать с меньшей точностью, чем при малых k, – нужно заботиться лишь о малости матрицы $\Gamma_n \mathcal{D}_{n,\ell}$.

Следствие 4. Для обеспечения близости решений x_n и $\bar{x}_n$ в H^ℓ $(0\le\ell\le m-1)$ следует проводить вычисления с точностью $\|\Gamma_n\mathcal{D}_{n,\ell}\|=o(n^{-1/2}/\ln n)$, $\|\delta^{(n)}\|=o(n^{-1/2}/\ln n)$, а в H^m требуются точности $\|\Gamma_n\mathcal{D}_{n,m}\|=o(n^{-3/2}/\ln^2 n)$, $\|\delta^{(n)}\|=o(n^{-3/2}/\ln^2 n)$

Замечание 1. Пусть координатная система ортонормальная в H^ℓ $(0\le\ell\le m-1)$, т.е. приближенное решение ищется в виде

$$x_n(t)=\frac{a_0}{\sqrt{2\pi}}+\sum_{k=1}^{n}\left(\pi\sum_{i=0}^{\ell}k^{2i}\right)^{-1/2}(a_k\cos kt+b_k\sin kt).$$

Тогда число обусловленности ν_n системы определения коэффициентов $a_0,a_1,b_1,\dots,a_n,b_n$ по методу подобластей оценивается

$$\nu_n\le \ln n\cdot n^{m-\ell+1/2}$$

Следовательно, при $\ell=m-1$ будет число обусловленности наименьшее, а именно

$$\nu_n\le \ln n\cdot n^{3/2}$$

и координатную последовательность целесообразно ортонормировать в H^{m-1}.

Литература

1. Б е р с Л., Д ж о н Ф., Ш е х т е р М., Уравнения с частными производными. Москва, 1966.
2. В а й н и к к о Г., Анализ дискретизационных методов. Тарту, 1976.
3. З и г м у н д А., Тригонометрические ряды, т.2. Москва, 1965.
4. К а н т о р о в и ч Л.В., А к и л о в Г.П., Функциональный анализ. Москва, 1977.

5. К а р м а О.О., Асимптотические оценки скорости сходимости собственных значений при регулярной аппроксимации. Республиканский симпозиум по методам решения нелинейных уравнений и задач оптимизации. Тезисы докладов. Пярну, 1978, 37-38.

6. К а р п и л о в с к а я Э.Б., О сходимости метода подобластей обыкновенных интегродифференциальных уравнений. Ж. вычисл. матем. и матем. физ., 1965, 5, № 1, 124-132.

7. К а т о Т., Теория возмущений линейных операторов. Москва, 1972.

8. К е р г е Р.М., К оценке погрешности метода подобластей. Ж. вычисл. матем. и матем. физ., 1978, 18, № 3, 628-633.

9. К е р г е Р., Метод подобластей в проблеме собственных значений. Материалы конф. методы алгебры и функциональн. анализа при исслед. семейств операторов. Тарту, 1978, 65-67.

10. К е р г е Р., Об устойчивости метода подобластей. Уч. зап. Тартуск. ун-та, 1979, 484, 47-55.

11. К и ш О., О сходимости метода подобластей. Acta math. Acad. Sci. Hung., 1967, 18, № 1-2, 175-179.

12. Н а т а н с о н И.П., Конструктивная теория функций. Москва-Ленинград, 1949.

13. П е т е р с е н И., О сходимости приближенных методов интерполяционного типа для обыкновенных дифференциальных уравнений. Изв. АН Эст. ССР, 1961, 10, № 1, 3-12.

14. Ф а н т а К., К и ш О., О сходимости интерполяционных методов решения граничных задач для обыкновенных дифференциальных уравнений. Magyar tud. akad. Mat. kutato int. Közl., 1964, 9, № 1-2, 89-112.

Поступило
9 III 1979

TEILGEBIETSVERFAHREN FÜR PERIODISCHE AUFGABE

R. Kerge

Zusammenfassung

Die Konvergenz des Teilgebietsverfahrens ist in [13] und [14] für gewöhnliche Differentialgleichungen und in [6] und [11] für gewöhnliche Integrodifferentialgleichungen untersucht worden. In [8] sind die Schätzungen der Konvergenzgeschwindigkeit im Fall der Tschebyscheffschen Knoten präzisiert worden. Dem Eigenwertproblem ist [9] gewidmet, und in [10] ist die Stabilität des Teilgebietsverfahrens im Fall des polynomischen Koordinatensystems beobachtet worden. In diesem Artikel ist das Teilgebietsverfahren für periodische Aufgabe untersucht worden.

Im § 1 wird die lineare periodische Aufgabe (1), (2) mit Hilfe des Teilgebietsverfahrens gelöst. Die Näherungslösung findet man im Form (3), und die Koeffizienten bestimmt man aus dem System (4), die Knoten t_i liegen auf dem Intervall $[0,2\pi]$ gleichmäßig. Im Theorem 1 ist die eindeutige Lösbarkeit des Systems (4) und die Konvergenzgeschwindigkeiten (5) und (6) begründet. Der Beweis des Theorems 1 stützt sich auf das allgemeine Theorem 2 aus [2]. In der Folgerung 1 sind die Schätzungen für eine glatte Lösung gebracht.

Im § 2 ist die nichtlineare periodische Aufgabe betrachtet worden. Die Ergebnisse sind im Theorem 3 zusammengefaßt.

Im § 3 ist das Eigenwertproblem betrachtet und das entsprechende Theorem 4 gebracht worden.

Der § 4 ist dem Stabilitätsproblem gewidmet. In der Rechenpraktik löst man statt des Systems (19) ein gestörtes System (20), und statt der Lösung (3) bekommt man eine gestörte Lösung (21). Der Einfluß der Störungen auf $x - \bar{x}_n$ ist in den Räumen H^l $(0 \leq l \leq m)$ betrachtet worden. Die Stabilitätsbedingungen sind in den Theoremen 5 und 6 dargestellt worden. Zusätzlich ist die Wahl des Koordinatensystems betrachtet und die Stabilitätszahl eingeschätzt worden.

Уч.зап.Тартуск.ун-та,1979,500,69-72

ИССЛЕДОВАНИЕ СХОДИМОСТИ МЕТОДА ИТЕРАЦИИ ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНОЙ РАЗНОСТНОЙ КРАЕВОЙ ЗАДАЧИ

М. Фишер

Тартуский государственный университет

В статье [1] была построена разностная схема для квазилинейной краевой задачи четвёртого порядка и показано, что эта схема сходится со скоростью $O(h^2)$. В этой статье рассматривается решение разностной схемы методом итераций.

Рассмотрим краевую задачу

$$Lu \equiv \sum_{|\alpha|\le 2} (-1)^{|\alpha|} \mathcal{D}^{\alpha} K_{\alpha}(x, \mathcal{D}^{\gamma} u) = f(x),\ x\in\Omega,\ |\gamma|\le 2, \quad (1)$$

$$u\big|_{\Gamma} = 0,\ \frac{\partial^2 u}{\partial n^2}\Big|_{\Gamma} = 0, \quad (2)$$

где $\Gamma = \bar{\Omega}\setminus\Omega$ граница p-мерного единичного куба $\Omega = \{0<x_i<1,\ i=1,\dots,p\}$, $\frac{\partial u}{\partial n}$ – производная по внешней нормали к границе Γ, $|\alpha| = \alpha_1+\dots+\alpha_p$, $x=(x_1,\dots,x_p)$, $\mathcal{D}^{\alpha} = \mathcal{D}_1^{\alpha_1}\dots\mathcal{D}_p^{\alpha_p}$, $\mathcal{D}_k = \frac{\partial}{\partial x_k}$.

Будем предполагать, что имеют место следующие предположения:

I. Условие эллиптичности

$$\sum_{|\alpha|,|\gamma|\le 2} \frac{\partial K_{\alpha}}{\partial p_{\gamma}} t_{\alpha} t_{\gamma} \ge c_0 \sum_{|\alpha|\le 2} t_{\alpha}^2,\ c_0 = const > 0.$$

II. В некоторой окрестности гиперкуба Ω

А) существует решение краевой задачи (1), (2), причём оно в ней достаточно гладкое;

В) функции $K_{\alpha}(x, p_{\gamma})$ достаточно раз непрерывно дифференцируемы и

$$\left|\frac{\partial K_{\alpha}}{\partial p_{\gamma}}\right| \le c_1,\ c_1 = const > 0.$$

Построим сетку

$$\bar{\omega} = \{x:\ x_i = k_i h,\ k_i = -1, 0, \dots, N+1\},\ h = 1/N.$$

и обозначим $\omega = \bar{\omega}\cap\Omega$, $\gamma_h = \bar{\omega}\cap\Gamma$. Аппроксимируем краевую задачу (1), (2) разностной краевой задачей

$$\Lambda y \equiv \frac{1}{2}\sum_{|\alpha|\le 2}(-1)^{|\alpha|}\left[\bar{\partial}^{\alpha}A_{\alpha}(x,\partial^{\gamma}y)+\partial^{\alpha}A_{\alpha}(x,\bar{\partial}^{\gamma}y)\right]=f(x),\ x\in\omega, \quad (3)$$

$$y\big|_{\gamma_h}=0,\ \partial_i\bar{\partial}_i y\big|_{\gamma_h}=0,\ i=1,\dots,p, \quad (4)$$

где $A_{\alpha}(x,p_{\gamma})=k_{\alpha}(x,p_{\gamma})$ в точках сетки.

Определим скалярные произведения и нормы на множестве и на его частях

$$(u,v)=h^p\sum_{x\in\omega}u(x)v(x),\quad \|u\|^2=(u,u),$$

$$(u,v)_{\pm\alpha}=h^p\sum_{x\in\omega_{\pm\alpha}}u(x)v(x),$$

$$\|u\|_2=\frac{1}{2}\sum_{|\alpha|\le 2}\left[(\bar{\partial}^{\alpha}u,\bar{\partial}^{\alpha}u)_{+\alpha}+(\partial^{\alpha}u,\partial^{\alpha}u)_{-\alpha}\right],$$

где

$$\omega_{+\alpha}=\{x:\ x_i=k_ih,\ k_i=1,\dots,N-1+\alpha_i\},$$

$$\omega_{-\alpha}=\{x:\ x_i=k_ih,\ k_i=1-\alpha_i,\dots,N-1\}.$$

Для рассмотрения итерационных методов при решении задачи (3), (4) введём на основе результатов статьи [2] следующие условия.

III. Условие R -монотонности

$$(\Lambda y-\Lambda v,\ y-v)\ge \bar{c}_0\|y-v\|_R^2,\ \bar{c}_0=const>0.$$

IУ. Условие Липшица

$$\|\Lambda y-\Lambda v\|_{-R}\le \bar{c}_1\|y-v\|_R,$$

где R - самосопряженный положительно определённый ограниченный оператор и

$$(y,z)_R=(Ry,z),\quad \|z\|_{-R}=\sup_{y\ne 0}\frac{|(y,z)|}{\|y\|_R}.$$

Для отыскания решения задачи (3), (4) используем итерационный метод

$$\tau^{-1}B(y^{n+1}-y^n)+\Lambda y^n=f, \quad (5)$$

где $B=E+\tau\sigma R+\tau^2\sigma^2Q(\tau)$, $\tau>0$ - итерационный параметр, $\sigma=const>0$, $Q(\tau)$ - самосопряжённый неотрицательный оператор.

На основе статьи [2] имеет место следующая теорема

<u>Теорема 1</u>. Если условия III, IУ выполнены, то задача (3), (4) однозначно разрешима при любой правой части. Решение за-

дачи (3), (4) может быть найдено с помощью итерационного процесса (5), сходящегося при любом начальном приближении, если

$$\tau > 0\,,\quad \sigma \geq c_1^2/(2c_0)\,.$$

Можно указать некоторые конкретные итерационные методы, проверив в каждом случае выполняемость условий III, IУ.

1° Схема типа метода переменных направлений. В этом случае возьмем

$$R_i y = \partial_i^2 \bar{\partial}_i^2 y\,,\ R = \sum_{i=1}^{p} R_i\,,\ B = \prod_{i=1}^{p}(E + \tau\sigma R_i) = E + \tau\sigma R + \tau^2\sigma^2 Q\,.$$

Учитывая доказательство теоремы в статье [1], можно написать

$$(\Lambda y - \Lambda v,\, y - v) \geq c_0 \|y - v\|_2^2\,,$$

Но так как нормы $\|y\|_2$ и $\|y\|_R$ эквивалентны (см. доказательство леммы 6 из [4]), то

$$(\Lambda y - \Lambda v,\, y - v) \geq \bar{c}_0 \|y - v\|_R^2\,.$$

Итак имеет место условие III. Ссылаясь ещё раз на доказательство теоремы из [1] получим

$$|(\Lambda y - \Lambda v, w)| \leq c\, \|y - v\|_2\, \|w\|_2\,, \qquad \text{где } c = const > 0.$$

и на основе эквивалентности норм получим

$$|(\Lambda y - \Lambda v, w)| \leq \bar{c}_1 \|y - v\|_R \cdot \|w\|_R\,.$$

Но это соотношение и есть условие Липшица IУ. Так как операторы R_i положительно определены, самосопряжёны и попарно перестановочны, то оператор Q неотрицателен.

2° Итерационные схемы бегущего счета. При таких схемах в предыдущем пункте рассмотренный оператор R представим в виде $R = \bar{R}_1 + \bar{R}_2$, $\bar{R}_1^* = \bar{R}_2$, где

$$\bar{R}_1 = \sum_{i=1}^{p} (h^{-2}\partial_i^2 - 2h^{-3}\partial_i)\,,\ \bar{R}_2 = \sum_{i=1}^{p} (h^{-2}\bar{\partial}_i^2 + 2h^{-3}\bar{\partial}_i)\,.$$

Из статьи [3] для схемы типа переменных направлений вытекает еще следующий результат

<u>Теорема 2</u>. При оптимальных значениях параметров $\tau = \tau_0 =$

$= (\sigma \sqrt{(\delta/\Delta)})^{-1}$, $\sigma = \sigma_0 = \frac{c_1^2}{c_0^2}\sqrt{(1 + \delta/\Delta)}$ нужное для оценки

$$\|y^n-u\|_R \le \varepsilon \|y^0-u\|_R$$

число итерации является величиной $O(\sqrt{(\Delta/\delta)}\ln\varepsilon^{-1})$ при $\delta/\Delta\to 0$, где $\delta E \le R_i \le \Delta E$, $i=1,2,\dots,p$.

В данном случае постоянные δ и Δ легко вычисляются $\delta = \frac{16}{h^4}\sin^4\left(\frac{\pi h}{2}\right)$, $\Delta=\frac{16}{h^4}\cos^4\left(\frac{\pi h}{2}\right)$. Следовательно число итерации, нужное для получения погрешности ε, имеет порядок $O(h^{-2}\ln 1/\varepsilon)$.

Литература

1. Фишер М., Исследование сходимости разностного метода для квазилинейной краевой задачи четвертого порядка. Уч. зап. Тартуск. ун-та, 1977, 430, 107-111.
2. Карчевский М.М., Итерационные схемы для уравнений с монотонными операторами. Изв. ВУЗ. Математика, 1971, № 5, 32-37.
3. Ляшко А.Д., Разностные схемы для квазилинейных эллиптических уравнений любого порядка. Изв. Вуз. Математика, 1973, № 9, 46-53.
4. Дьяконов Е.Г., О некоторых итерационных методах решения систем разностных уравнений, возникающих при решении методом сеток уравнений в частных производных эллиптического типа. В сб. Вычисл. методы и программирование, ВЦ МГУ, т. 3, 1965, 191-222.

Поступило
13 III 1979

UNTERSUCHUNG DER KONVERGENZ DES ITERATIONSVERFAHRENS BEI DER AUFLÖSUNG DER RANDWERTAUFGABE DER NICHTLINEAREN DIFFERENZENGLEICHUNG

M. Fischer

Zusammenfassung

In diesem Artikel untersucht man zwei verschiedene Iterationsverfahren: die Methode der veränderlichen Richtungen und die abwechselnde dreieckige Methode. Man zeigt, daß mit der Hilfe dieser beiden Verfahren die Folge der erhaltenen Näherungen zur Lösung der Randwertaufgabe der nichtlinearen Differenzengleichung konvergiert.

Уч.зап.Тартуск.ун-та,1979,500,73-91

СТРУКТУРА РЕШЕНИЯ НЕКОТОРЫХ СИСТЕМ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ, СВЯЗАННЫХ С ПРИБЛИЖЕННЫМ РЕШЕНИЕМ ЗАДАЧ ПЕРЕНОСА ИЗЛУЧЕНИЯ

Ю. Князихин

Тартуский государственный университет

Введение

В данной статье изучается вид решения системы интегральных уравнений (ОЗ - общая задача) вида

$$x(t) = \int_0^t A_1 \exp(-B_1(t-s))x(s)\,ds + \int_t^H A_2 \exp(B_2(t-s))x(s)\,ds + f(t), \qquad (1)$$

где[1] $f(t) = A_0 \exp(B_0 t)\mathbf{1}$,

$$A_\omega = (a_{ij}^{\omega})_{ij}, \quad B_\omega = (\delta_{ij} b_{ij}^{\omega})_{i,j}, \quad b_j^{\varkappa} > 0,$$

а $\mathbf{1}$ есть вектор-столбец, составленный из единиц. Сформулированы условия, при которых ОЗ имеет единственное решение и имеет вид

$$x(t) = G_1 \exp(J_1 t)\xi_1 + G_2 \exp(J_2 t)\xi_2 + G_0 \exp(B_0 t)\xi_0,$$

где G_ω некоторые матрицы, $J_\varkappa$ диагональные матрицы, ξ_ω вектора.

К ОЗ можно прийти, решая методом дискретных ординат одно из основных уравнений переноса излучения

$$y(t,\mu,\varphi) = \lambda/4\pi \Big[\int_0^{2\pi}\int_0^1\int_0^t \frac{g(\mu,\varphi,\eta,\psi)}{\eta} \exp\Big(-\frac{t-\tau}{\eta}\Big) y(\tau,\eta,\psi)\,d\psi\,d\eta\,d\tau + \int_0^{2\pi}\int_0^1\int_t^H \frac{g(\mu,\varphi,-\eta,\psi)}{\eta} \exp\Big(\frac{t-\tau}{\eta}\Big) y(\tau,-\eta,\psi)\,d\psi\,d\eta\,d\tau\Big] + u(t,\mu,\varphi), \quad 0 \le \lambda \le 1. \qquad (2)$$

---

[1] Во всей статье, если не оговорено специально, подразумевается следующее изменение индексов: $k, s = \pm 1, \pm 2, \dots, \pm n$; $m, l = 1, 2, \dots, p$; $\varkappa = 1, 2$, $\omega = 0, 1, 2$. Остальные индексы пробегают значения от 1 до n. Символы $(a_{ij})_{ij}$ и $(a_i)_i$ означают матрицу и вектор-столбец соответственно; δ_{ij} - символ Кронекера.

Функция $u(t,\mu,\varphi)$ характеризует излучающую способность среды и является экспонентой относительно t. Функция $g(\mu,\varphi,\eta,\psi)$ называется индикатрисой рассеяния и удовлетворяет следующим соотношениям

$$\int_{-1}^{1}\int_{0}^{2\pi} g(\mu,\varphi,\eta,\psi)\,d\eta\,d\psi = 4\pi, \tag{3}$$

$$g(\mu,\varphi,\eta,\psi) = g(\eta,\psi,\mu,\varphi) = g(\eta,\varphi,\mu,\psi) = g(\mu,\psi,\eta,\varphi) > 0, \tag{4}$$

$$g(\mu,\varphi,\eta,\psi) = g(-\mu,\varphi,-\eta,\psi). \tag{5}$$

Для того, чтобы привести уравнение (2) в виду ОЗ, заменим интегралы по ψ и η в уравнении переноса излучения какими-нибудь квадратурными формулами (метод дискретных ординат)

$$\int_0^{2\pi} \zeta(\psi)d\psi \approx \sum_m \beta_m \zeta(\psi_m),\ 0<\psi_1<\dots<\psi_p\le 2\pi,\ \beta_m>0,\ \sum_m \beta_m = 2\pi,$$

$$\int_0^1 h(\eta)d\eta \approx \sum_i \alpha_i h(\eta_i),\ 0<\eta_1<\dots<\eta_n\le 1,\ \alpha_j>0,\ \sum_j \alpha_j = 1.$$

Тогда, учитывая последнее, получим дискретную задачу для уравнения переноса (ДЗУП)

$$y_{sm}(t) = \int_0^t \sum_i \sum_\ell r_{i\ell sm} \exp(-b_i(t-\tau))\,y_{i\ell}(\tau)\,d\tau +$$

$$+ \int_t^H \sum_i \sum_\ell r_{-i\ell ms} \exp(b_i(t-\tau))\,y_{-i\ell}(\tau)\,d\tau + u_{sm}(t), \tag{6}$$

где

$$b_k = 1/\eta_k,\ \eta_{-i} = -\eta_i,\ g_{kms\ell} = g(\eta_k,\psi_m,\eta_s,\psi_\ell),$$

$$r_{kms\ell} = (\lambda\alpha_{|k|}\beta_m/4\pi\eta_{|k|})\,g_{kms\ell},\ u_{sm}(t) = u(t,\eta_s,\psi_m).$$

Учитывая равенства (3) - (5) и делая в случае необходимости перенормировку, положим

$$\sum_i \sum_\ell (g_{i\ell sm} + g_{-i\ell sm})\alpha_i\beta_\ell = 4\pi. \tag{7}$$

Пусть $I_{n\times n}$ есть единичная матрица размерности $n\times n$; символ e_j^n означает вектор-столбец длины n, у которого на j-том месте стоит единица, а на остальных местах нули, а $\otimes$ есть знак кронекеровского произведения матриц (см. [1], стр. 256). Положим

$$\sigma_i = (\lambda\alpha_i)/(2\eta_i),\ \gamma_m = \beta_m/2\pi,\ \sigma = (\delta_{ij}\sigma_j)_{i,j},\ \Gamma = (\delta_{m\ell}\gamma_\ell)_{m,\ell},$$

$$\Pi = I_{2\times 2}\otimes\sigma\otimes\Gamma,\ w_{ks} = (g_{kms\ell})_{m,\ell},\ V_1 = (w_{ij})_{i,j},\ V_2 = (w_{-i,j})_{i,j},$$

$$w_1 = \left\| \begin{matrix} V_1 & 0 \\ V_2 & 0 \end{matrix} \right\| , \quad w_2 = \left\| \begin{matrix} 0 & V_2 \\ 0 & V_1 \end{matrix} \right\| , \quad A_\varkappa = w_\varkappa \Pi ,$$

$$b = (\delta_{ij} b_j)_{i,j} , \; B = I_{2\times 2} \otimes b \otimes I_{p\times p} , \; y_j = (y_{je})_e , \; u_j(t) = (u_{je})_e ,$$

$$x(t) = \sum_i \left(e^{2n}{}_i \otimes y_i + e^{2n}{}_{i+n} y_{-i} \right), \; f(t) = \sum \left(e^{2n}_i \otimes u_i + e^{2n}_{i+n} \otimes u_{-i} \right)$$

Учитывая введенные обозначения, ДЗУП (6) примет вид

$$x(t) = \int_0^t A_1 \exp(-B(t-s)) x(s) ds + \int_t^H A_2 \exp(B(t-s)) x(s) ds + f(t) . \qquad (8)$$

Но так как функция $u(t,\mu,\varphi)$ в уравнении (2) является экспонентой относительно t, то свободный член $f(t)$ в ДЗУП можно записать в виде $f(t) = A_0 \exp(B_0 t) \mathbb{1}$, где A_0 некоторая матрица, B_0 диагональная матрица $(B_0 = (1/\gamma) I_{2np\times 2np})$. Подробно ДЗУП будет изучаться в §§ 7–8.

§ 1. Существование и единственность решения ОЗ и ДЗУП

Для матрицы $A = (a_{ij})_{ij}$ положим

$$\|A\|_1 = \max_i \sum_j |a_{ij}| , \quad |A| = (|a_{ij}|)_{i,j} .$$

Рассмотрим пространство непрерывных вектор-функций с нормой

$$\|x\|_C = \max_i \max_{0\le t\le H} |x_i(t)| .$$

Тогда норма оператора K такого, что Kx есть сумма интегральных членов в выражении (1) есть (см. [2], стр. 176)

$$\|K\|_{C\to C} = \max_i \max_{0\le t\le H} \Big(\sum_j \int_0^t |a^1_{ij}| \exp(-b^1_j (t-s)) ds +$$
$$+ \int_t^H |a^2_{i,j}| \exp(b^2_j (t-s)) ds \Big) .$$

Пусть максимум по t достигается в точке $t^* \in [0, H]$. Делая замену $u = t^* - s$ в первом интеграле и $u = s - t^*$ во втором, получим

$$\|K\|_{C\to C} = \max_i \Big(\sum_j \int_0^{t^*} |a^1_{ij}| \exp(-b^1_j u) du +$$
$$+ \int_0^{H-t^*} |a^2_{ij}| \exp(-b^2_j u) du \Big) <$$

$$< \max_i \left(\sum_j \int_0^\infty |a^1_{ij}| \exp(-b^1_j u)du + \int_0^\infty |a^2_{ij}| \exp(-b^2_j u)du \right) =$$

$$= \| |A_1| B_1^{-1} + |A_2| B_2^{-1} \|_1 ,$$

Отсюда, учитывая теорему Банаха о единственности решения операторных уравнений в банаховых пространствах, находим, что справедлива

Теорема 1. Пусть выполнено условие

$$\| |A_1| B_1^{-1} + |A_2| B_2^{-1} \|_1 \leq 1. \tag{9}$$

Тогда ОЗ имеет единственное решение при любом свободном члене.

Если уравнение (1) есть ДЗУП, то, учитывая равенство (7) и (4), получим

$$\| |A_1| B_1^{-1} + |A_2| B_2^{-1} \|_1 = \max_{s,m} \sum_{i,l} (1/b_i)(r_{ilsm} + r_{2ilsm}) = \lambda .$$

Следовательно, ДЗУП имеет решение при любом $\lambda \in (0,1]$.

§ 2. Некоторые определения и леммы

Обозначим через $d(A)$ множество диагональных элементов матрицы A. Определим операцию, которая каждой тройке вида

$$(G, J, B) = ((g_{ij})_{i,j}, (\delta_{ij}\lambda_j)_{ij}, (\delta_{ij} b_j)_{i,j}),$$

при условии $d(J) \cap d(-B) = \emptyset$ ставит в соответствие матрицу

$$G[J,B] = (g_{ij}/(\lambda_i + b_j))_{i,j} .$$

Очевидны следующие утверждения:

Лемма 1. Для диагональных матриц J и B таких, что $d(J) \cap d(-B) = \emptyset$ имеет место равенство

$$\int_a^b \exp(Jt)\, G \exp(Bt) dt = \exp(Jt) G(J,B) \exp(Bt) \Big|_a^b .$$

Лемма 2. Для матриц A, B, C и диагональных матриц $\mathcal{D}$, E таких, что $d(\mathcal{D}) \cap d(E) = \emptyset$ имеют место эквивалентности

а) $A\exp(\mathcal{D}t)x + B\exp(Et)y \equiv 0 \sim A\exp(\mathcal{D}t)x \equiv 0,\ B\exp(Et)y \equiv 0;$

б) $A\exp(Ct)x \equiv 0 \sim AC^k x \equiv 0,\ k = 0,1,\ldots \sim x \in \bigcap_{k=0}^{\infty} \operatorname{Ker}(AC^k).$

Эквивалентность б) вытекает из аналитичности вектор-функции $A\exp(Ct)$.

Лемма 3. Пусть для заданной пары $(\mathcal{X}, A)$, где $\mathcal{X}$ есть функция от матрицы и A некоторая матрица, существуют такие матрица G и обратимая диагональная матрица $\mathcal{D}$, что

$$\mathcal{X}(G) = A\mathcal{D}.$$

Тогда имеет место тождество

$$\mathcal{X}(G)\exp(Bt)\xi \equiv -A\exp(Bt)\eta,$$

где B есть диагональная матрица, а $\xi = -\mathcal{D}^{-1}\eta$.

При доказательстве этой леммы надо учесть коммутативность $\mathcal{D}$ и $\exp(Bt)$.

Лемма 4. Для любого $v \in R^{2n}$ и диагональных матриц B, C, $\mathcal{D}$ таких, что $(d(-B) \cup d(C)) \cap d(\mathcal{D}) = \emptyset$ существует матрица S, что

$$v = e_2^1 \otimes (S[B,\mathcal{D}]\mathbb{1}) + e_2^2 \otimes (S[-C,\mathcal{D}]\exp(\mathcal{D}H)\mathbb{1}). \quad (10)$$

Доказательство. Пусть $S = (s_{ij})_{ij}$ и $s_{ij} = 0$ при $j \geq 3$. Записывая выражение (10) в координатном виде, получим

$$\begin{cases} s_{i1}/(b_i + d_1) + s_{i2}/(b_i + d_2) = v_i^1; \\ (s_{i1}\exp(d_1 H))/(-c_i + d_1) + (s_{i2}\exp(d_2 H))/(-c_i + d_2) = v_i^2, \end{cases}$$

где

$$B = (\delta_{ij} b_i)_{ij}, \quad C = (\delta_{ij} c_i), \quad \mathcal{D} = (\delta_{ij} d_j)_{ij}.$$

Определители этих систем порядка 2×2 отличны от нуля. Это дает нам возможность определения s_{i1} и s_{i2}. Лемма доказана.

Определение 1. Мы говорим, что пара матриц $(\mathcal{J}, G)$ принадлежит классу $T(A_1, A_2, B_1, B_2)$ и будем писать

$$(\mathcal{J}, G) \in T(A_1, A_2, B_1, B_2),$$

если матрица $\mathcal{J} = (\delta_{ij}\lambda_j)_{ij}$ такова, что

$$\det(T(\lambda_j) - I_{n\times n}) = 0,$$

а j-тый столбец матрицы G удовлетворяет

$$(T(\lambda_j) - I)g_j = 0,$$

где $T(s) = A_1[sI, B_1] - A_2[sI, -B_2]$.

Пусть $T^2(A_1, A_2, B_1, B_2) = T(A_1, A_2, B_1, B_2) \times T(A_1, A_2, B_1, B_2)$.

Для каждого $\varepsilon = ((\mathcal{J}_1, G_1), (\mathcal{J}_2, G_2))$ такого, что $\varepsilon \in T^2(A_1, A_2, B_1, B_2)$, определим множество

$$\Xi(\varepsilon) = \{\xi \in R^{2n} \mid \xi = e_1^2 \otimes \xi_1 + e_2^2 \otimes \xi_2,\ \xi_{\varkappa} \in R^n, \sum_{\varkappa} G_{\varkappa} \exp(\mathcal{J}_{\varkappa} t)\xi_{\varkappa} = 0\}$$

и целочисленную функцию $\gamma(\varepsilon) = \dim \Xi(\varepsilon)$.

§ 3. Получение некоторых тождеств

Будем искать решение ОЗ в виде

$$x(t) = G_1 \exp(\mathcal{J}_1 t)\xi_1 + G_2 \exp(\mathcal{J}_2 t)\xi_2 + G_0 \exp(B_0 t)\xi_0, \quad (11)$$

где векторы $\xi_\omega = (\xi_i^\omega)_i$, матрицы $G_\omega = (g_{ij}^\omega)_{ij}$ и $\mathcal{J}_{\varkappa} = (\delta_{ij} \lambda_j^{\varkappa})_{ij}$ пока неопределены. Подставим (11) в уравнение (1). Если мы потребуем выполнение условий

$$(\sigma(-B_1) \cup \sigma(B_2)) \cap \sigma(B_0) = \varnothing, \quad (12)$$

$$\sigma(-B_1) \cap (\sigma(\mathcal{J}_1) \cup \sigma(\mathcal{J}_2)) = \varnothing, \quad \sigma(B_2) \cap (\sigma(\mathcal{J}_1) \cup \sigma(\mathcal{J}_2)) = \varnothing, \quad (13)$$

то, используя лемму 1, мы сможем избавиться от интегральных членов. Из полученного после этого выражения можно увидеть, что (11) будет удовлетворять уравнение (1), если при любом $t \in [0, H]$ будут одновременно выполнены равенства

$$(A_1 G_{\varkappa}[B_1, \mathcal{J}_{\varkappa}] - A_2 G_{\varkappa}[-B_2, \mathcal{J}_{\varkappa}] - G_{\varkappa}) \exp(\mathcal{J}_{\varkappa} t)\xi_{\varkappa} = 0, \quad (14)$$

$$(A_1 G_0 [B_1, B_0] - A_2 G_0[-B_2, B_0] - G_0) \exp(B_0 t)\xi_0 = -A_0 \exp(B_0 t)\mathbb{1}, \quad (15)$$

$$\begin{aligned}
&A_1 \exp(-B_1 t) G_0[B_1, B_0]\xi_0 + A_1 \exp(-B_1 t)\Big(\sum_{\varkappa} G_{\varkappa}[B_1, \mathcal{J}_{\varkappa}]\xi_{\varkappa}\Big) = 0, \\
&A_2 \exp(B_2(t-H)) G_0[-B_2, B_0] \exp(B_0 H)\xi_0 + \\
&+ A_2 \exp(B_2(t-H)) \sum_{\varkappa} G_{\varkappa}[-B_2, \mathcal{J}_{\varkappa}] \exp(\mathcal{J}_{\varkappa} H)\xi_{\varkappa} = 0.
\end{aligned} \quad (16)$$

§ 4. Исследование равенств (14) и (15)

Теорема 2. Пусть $((\mathcal{J}_1, G_1), (\mathcal{J}_2, G_2)) \in T^2(A_1, A_2, B_1, B_2)$ и пусть выполнено условие

$$\forall \sigma \in \sigma(-B_1) \cup \sigma(B_2) \cup \sigma(B_0) \Rightarrow \det(\sigma^2 I + \sigma F + P) = 0, \quad (17)$$

где $F = (A_2 - A_1 + B_1 - B_2)$ и $P = A_1 B_1 + A_2 B_2 - B_1 B_2$. Тогда имеют место условия (13) и пары $(\mathcal{J}_1, G_1)$ и $(\mathcal{J}_2, G_2)$ обращают (14) в тождество.

Доказательство. Заметим, что

$$T(s)-I=-(s^2I+sF+P)(sI+B_1)^{-1}(sI-B_2)^{-1}.$$

Пусть $((\mathcal{J}_1,G_1),(\mathcal{J}_2,G_2))\in T^2(A_1,A_2,B_1,B_2)$. Тогда при выполнении условия (17) будет выполнено условие (13).

Тождества (14) будут выполнены, если найдутся такие диагональные матрицы $\mathcal{J}_æ$ и матрицы $G_æ$, что

$$A_1G_æ[B_1,\mathcal{J}_æ]-A_2G_æ[-B_2,\mathcal{J}_æ]-G_æ=0.$$

Пусть $\lambda_j^æ$ и $g_j^æ$ есть j-тый элемент диагонали матрицы $\mathcal{J}_æ$ и j-тый столбец матрицы $G_æ$. Тогда последнее равенство можно записать

$$(T(\lambda_j^æ)-I_{n\times n})g_j^æ=0.$$

Значит, если $((\mathcal{J}_1,G_1),(\mathcal{J}_2,G_2))\in T^2(A_1,A_2,B_1,B_2)$, то тождества (14) будут выполнены. Теорема доказана.

Теорема 3. Пусть выполнено условие (12) и (17). Тогда существуют такие матрица G_0 и вектор ξ_0, что имеет место тождество (15).

Доказательство. По лемме 3 имеем, что тождество (15) будет выполнено, если мы подберем матрицу G_0 и диагональную матрицу $\mathcal{D}$, что

$$A_1G_0[B_1,B_0]-A_2G_0[-B_2,B_0]-G_0=A_0$$

или, обозначив j-тый столбец матриц G и $A_0\mathcal{D}$ через g_j^0 и s_j соответственно, получим

$$(T(b_j^0)-I_{n\times n})g_j^0=s_j,\quad b_j^0\in M(B_0). \tag{18}$$

Условия (12) и (17) гарантируют невырожденность матриц $(T(b_j^0)-I)$. Значит из условий (18) можно определить матрицу G_0. В качестве обратимой диагональной матрицы берем $I_{n\times n}$. Отсюда $\xi_0=-I_{n\times n}\mathbb{1}=-\mathbb{1}$. Теорема доказана.

§ 5. Исследование тождества (16)

Пусть $P_æ$ есть ортопроекторы в R^n, проектирующие на

$$N_æ=\bigcap_{k=0}^{\infty}\mathrm{Ker}(A_æB_æ^k).$$

Тогда, учитывая лемму 2, тождество (16) будет выполнено тогда и только тогда, когда имеет решение система

$$(I-P_1)\Big(\sum_{\varkappa} G_\varkappa[B_1,\mathcal{J}_\varkappa]\xi_\varkappa + G_0[B_1,B_0]\xi_0\Big)=0,$$
$$(I-P_2)\Big(\sum_{\varkappa} G_\varkappa[-B_2,\mathcal{J}_\varkappa]\exp(\mathcal{J}_\varkappa H)\xi_\varkappa + G_0[-B_2,B_0]\exp(B_0 H)\xi_0\Big)=0. \qquad (19)$$

Положим

$$P=\begin{Vmatrix} I-P_1 & 0 \\ 0 & I-P_2 \end{Vmatrix}, \quad F=\begin{Vmatrix} G_1[B_1,\mathcal{J}_1] & G_2[B_1,\mathcal{J}_2] \\ G_1[-B_2,\mathcal{J}_1]\exp(\mathcal{J}_1 H) & G_2[-B_2,\mathcal{J}_2]\exp(\mathcal{J}_2 H) \end{Vmatrix}$$

$$\eta = e_1^2\otimes\big(G_0[B_1,B_0]\mathbb{1}\big) + e_2^2\otimes\big(G_0[B_2,B_0]\exp(B_0H)\mathbb{1}\big),$$

тогда система (19) примет вид

$$PF\xi = P\eta, \qquad (20)$$

где $P: R^{2n}\longrightarrow (I-P_1)R^n \times (I-P_2)R^n$.

Система (20) разрешима для любого $\eta\in R^{2n}$ тогда и только тогда, когда

$$\dim \operatorname{Ker}(PF) = \dim N_1 + \dim N_2 . \qquad (21)$$

Лемма 5. Пусть выполнены условия (9), (12), (17) и

$$\varepsilon = \big((\mathcal{J}_1,G_1),(\mathcal{J}_2,G_2)\big)\in T^2(A_1,A_2,B_1,B_2).$$

Тогда имеет место равенство

$$\Xi(\varepsilon) = \operatorname{Ker}(PF).$$

Доказательство. Отметим, что условие (9) гарантирует однозначность ОЗ, а условия (12) и (17) корректность соотношений (16). Пусть $\xi\in \operatorname{Ker}(PF)$ или, что тоже самое, $PF\xi = 0$. Рассмотрим соответствующее ОЗ однородное уравнение

$$x(t)=\int_0^t A_1\exp(-B_1(t-s))x(s)ds + \int_t^H A_2\exp(B_2(t-s))x(s)ds.$$

Решение этого уравнения есть

$$x(t)=\sum_{\varkappa} G_\varkappa \exp(\mathcal{J}_1 t)\xi_\varkappa .$$

В силу единственности ОЗ получим, что $x(t)\equiv 0$, а это значит, что

$$\xi = e_1^2\otimes\xi_1 + e_2^2\otimes\xi_2 \in \Xi(\varepsilon)$$

Отсюда имеем, что $\operatorname{Ker} PF \subset \Xi(\varepsilon)$.

Покажем теперь включение в другую сторону

$$\operatorname{Ker} PF \supset \Xi(\varepsilon).$$

Предположим обратное: существует такой вектор $\xi \in \Xi(\varepsilon)$, что $\xi \notin \mathrm{Ker}\, PF$, т.е.

$$PF(\xi) = \delta \neq 0, \quad \delta \in R^{2n}.$$

Пусть[2] $\varphi \in R^{2n}$ таков, что $P\varphi = \delta$. Согласно лемме 4, можно подобрать такую матрицу $S = (s_{ij})_{ij}$, что

$$\varphi = e_1^2 \otimes (S[B_1, B_0]\mathbb{1}) + e_2^2 \otimes (S[-B_2, B_0] \exp(B_0 H)\mathbb{1}).$$

Образуем матрицу A_0^1, у которой j-тый столбец выражается через j-тый столбец матрицы S по формуле

$$a_j^0 = (T(b_j^0) - I_{n\times n}) s_j.$$

В силу невырожденности матриц $(T(b_j^0) - I)$, матрица A_0^1 отлична от нулевой. Полученная таким образом матрица A_1^0 такова, что выполнено тождество (15), где вместе A^0 стоит A_1^0, а вместо G_0 стоит S. Таким образом, у нас имеется, что матрицы $J_\varkappa$ и $G_\varkappa$ удовлетворяют тождества (14), по теореме 2; матрица S удовлетворяет тождество (15), где вместо A_0 стоит A_1^0, по построению; вектор ξ удовлетворяет систему (19), где вместо η стоит φ, по предположению. Отсюда

$$x(t) = G_1 \exp(J_1 t)\xi_1 + G_2 \exp(J_2 t)\xi_2 + G_0 \exp(B_0 t)\xi_0$$

будет удовлетворять уравнение

$$x(t) = \int_0^t A_1 \exp(-B_1(t-s)) x(s)\, ds + \int_t^H A_2 \exp(B_2(t-s)) x(s)\, ds + A_0^1 e^{B_0 t} \quad (22)$$

Учитывая, что $\xi \in \Xi(\varepsilon)$ получим, что решение уравнения (22) есть

$$x(t) = S \exp(B_0 t)\mathbb{1}.$$

Подставляя его в (22) и учитывая условия (11) и (16), лемму 1, получим, что необходимо должно выполняться тождество

$$(S - A_1^0 - A_1 S[B_1, B_0] + A_2 S[-B_2, B_0]) \exp(B_0 t)\mathbb{1} + A_1 \exp(-B_1 t) \times$$
$$\times S[B_1, B_0] + A_2 \exp(B_2(t-H)) S[-B_2, B_0] \exp(B_0 H)\mathbb{1} \equiv 0.$$

Учитывая лемму 2, имеем, что последнее тождество будет выполнено тогда и только тогда, когда

$$(S - A_1^0 - A_1 S[B_1, B_0] + A_2 S[-B_2, B_0]) \exp(B_0 t)\mathbb{1} = 0,$$

---

[2] Подобрать такой вектор всегда можно. Можно взять, например, $\varphi = \delta$.

II

$$A_1 \exp(B_1 t) S[B_1,B_0]\mathbb{1} = 0,\quad A_2 \exp(B_2(t-H)) S[-B_2,B_0]\exp(B_0 H)\mathbb{1} = 0 \quad (23)$$

Тождества (23) будут выполнены тогда и только тогда, когда

$$S[B_1,B_0]\mathbb{1} \in \bigcap_{k=0}^{\infty} Ker(A_1 B_1^k),\quad S[-B_2,B_0]\exp(B_0 H)\mathbb{1} \in \bigcap_{k=0}^{\infty} Ker(A_2 B_2^k),$$

а это противоречит тому, что $P\varphi = 0$. Лемма доказана.

Следствие. Пусть выполнены условия (8), (11), (16) и пусть уравнение $\gamma(\varepsilon) = dim\, N_1 + dim\, N_2$ имеет хотя бы одно решение $\varepsilon = ((G_1, J_1),(G_2, J_2)) \in T^2(A_1, A_2, B_1, B_2)$. Тогда решение ОЗ единственно и имеет вид

$$x(t) = G_1 \exp(J_1 t)\xi_1 + G_2 \exp(J_2 t)\xi_2 + G_0 \exp(B_0 t)\xi_0 .$$

Доказательство вытекает из условия (21) и леммы 5.

§ 6. Некоторые свойства функции $\gamma(\varepsilon)$

Пусть S есть некоторое конечное множество. Через $\nu(S)$ обозначим множество, состоящее из попарно различных элементов множества S. Пусть Λ есть множество всех корней уравнения

$$det\,(T(s) - I_{n\times n}) = 0 .$$

Положим

$$M(i) = \{ j \mid b_j^1 = b_i^1,\ b_j^1 \in d(B_1),\ b_i^1 \in \nu(d(B_1))\},\ i = 1,2,\ldots,q,$$

$$Q(i) = \{ j \mid b_j^2 = b_i^2,\ b_j^2 \in d(B_2),\ b_i^2 \in \nu(d(B_2))\},\ i = 1,2,\ldots,r.$$

Пусть A_M есть матрица вида

$$(a_{p(1)}, a_{p(2)}, \ldots, a_{p(k)}),$$

где a_j есть j-тый столбец матрицы A, а

$$p(i) \in M \subset \{1,2,\ldots,n\},\quad p(i) < p(i+1).$$

Теорема 4. Имеет место равенство

$$\min\{\gamma(\varepsilon) \mid \varepsilon \in T^2(A_1,A_2,B_1,B_2)\} = 2n - \sum_{s\in\nu(\Lambda)} dim\, Ker(T(s) - I). \quad (24)$$

Доказательство. Возьмем[3] $\varepsilon \in T^2(A_1,A_2,B_1,B_2)$ и $\xi \in \Xi(\varepsilon)$. Тогда, по определению множества $\Xi(\varepsilon)$ имеем

$$G_1 \exp(J_1 t)\xi_1 + G_2 \exp(J_2 t)\xi_2 \equiv 0 .$$

[3] Здесь и далее $\varepsilon = ((J_1, G_1),(J_2, G_2))$

Запишем это тождество в координатном виде. Приравнивая к нулю коэффициенты при разных $exp(\lambda_j^{ae} t)$, получим

$$\sum_{k \in \Lambda(i)} g_k \xi^k = 0 \ , i = 1, 2, \dots, s, \tag{25}$$

где g_k есть столбец матрицы G_1 или G_2, а ξ^k координата вектора $\xi = e_1^2 \otimes \xi_1 + e_2^2 \otimes \xi_2$,

$$\Lambda(i) = \{ j \mid \lambda_j = \lambda_i, \lambda_j \in d(\mathcal{J}_1) \cup d(\mathcal{J}_2), \lambda_i \in \nu(d(\mathcal{J}_1) \cup d(\mathcal{J}_2)) \}, i = 1, \dots, s.$$

Из системы (25) можно определить вектор ξ, причем значение $dim\ \Xi(\varepsilon)$ совпадает с размерностью ядра системы (25). Учитывая, что g_k удовлетворяет системе (по определению 1)

$$(T(\lambda) - I) g_k = 0, \ \lambda \in d(\mathcal{J}_1) \cup d(\mathcal{J}_2),$$

и что $\Lambda(i) \cap \Lambda(j) = \emptyset$ при $i \neq j$, получим, что максимальный ранг матрицы системы (25) не может превышать

$$\sum_{s \in \nu(\Lambda)} dim\, Ker\, \Xi(\varepsilon).$$

Учитывая, что $dim\ \Xi(\varepsilon)$ совпадает с ядром системы (25), получим формулу (24). Теорема доказана.

Теорема 5. Для любого $\varepsilon \in T^2(A_1, A_2, B_1, B_2)$ имеет место неравенство

$$\gamma(\varepsilon) \geq dim\, N_1 + dim\, N_2.$$

Доказательство. Из очевидного неравенства

$$rang\, PF \leq dim\, PR^{2n},$$

следует

$$dim\, Ker\, PF \geq dim\, N_1 + dim\, N_2.$$

Отсюда, по лемме 5, следует

$$dim\, \Xi(\varepsilon) \geq dim\, N_1 + dim\, N_2.$$

Теорема доказана.

Теорема 6. Имеют место равенства

$$dim\, N_1 = \sum_{i=1}^{q} dim\, Ker(A_{1m(i)}),$$

$$dim\, N_2 = \sum_{i=1}^{2} dim\, Ker(A_{2Q(i)}).$$

Доказательство аналогично доказательству теоремы 4.

Из теорем 3–6 следует

Теорема 7. Пусть выполнены условия (9), (12), (17) и

$$\sum_{i=1}^{q} rang\,(A_{1M(i)}) + \sum_{i=1}^{r} rang\,(A_{2Q(i)}) = \sum_{\lambda\in\nu(\Lambda)} dim\,Ker\,(T(\lambda)-I). \quad (26)$$

Тогда решение ОЗ единственно и имеет вид (11).

§ 7. Некоторые особенности ДЗУП

В этом параграфе мы займемся применением теоремы 7 для изучения решения ДЗУП. Сформулируем ее в "развернутом" виде.

Теорема 8. Пусть выполнены условия

1. $\| \, |A_1| B_1^{-1} + |A_2| B_2^{-1} \|_1 \le 1$,
2. $(d(-B_1) \cup d(B_2)) \cap d(B_0) = \varnothing$;
3. $\forall \sigma \in d(-B_1) \cup d(B_2) \cup d(B_0) = \varnothing \Rightarrow det(\sigma^2 I + \sigma I + P) \ne 0$,
   $F = A_2 - A_1 + B_1 - B_2$, $P = A_1 B_2 + A_2 B_1 - B_1 B_2$;
4. $\sum_{i=1}^{q} rang\,(A_{1M(i)}) + \sum_{i=1}^{r} rang\,(A_{2Q(i)}) = \sum_{\lambda\in\nu(\Lambda)} dim\,Ker\,(T(\lambda)-I)$

Тогда решение ОЗ единственно и имеет вид

$$x(t) = G_1 \exp(\mathcal{J}_1 t)\xi_1 + G_2 \exp(\mathcal{J}_2 t)\xi_2 + G_0 \exp(B_0 t)\xi_0,$$

где матрицы $G_\omega, \mathcal{J}_\omega$ и вектора ξ_1 и ξ_2 определяются по следующим правилам:

а) j -тый элемент диагональной матрицы $\mathcal{J}_\varkappa = (\lambda_j^\varkappa \delta_{ij})$ является корнем уравнения (по теореме 2)

$$det\,(T(\lambda) - I) = 0; \quad (27)$$

б) $d(\mathcal{J}_1) \cup d(\mathcal{J}_2) = \Lambda$, где Λ есть множество всех корней уравнения (27);

в) каждый элемент базиса пространства $Ker\,(T(\lambda_j^\varkappa) - I)$ есть столбец матрицы G_1 или G_2;

Замечание 1. Правило б) и в) вытекает из следствия леммы 5 и теорем 4–6;

г) j -тый столбец $g_j^\varkappa$ матрицы $G_\varkappa$ удовлетворяет уравнению (по теореме 2)

$$(T(\lambda_j^\varkappa) - I)\, g_j^\varkappa = 0;$$

д) векторы ξ_1 и ξ_2 удовлетворяют систему (18);

е) j -тый столбец g_j° матрицы $G_\circ$ удовлетворяет (по теореме 3)

$$(T(b_j^\circ)-I)g_j^\circ = a_j^\circ ;$$

ж) $\xi_\circ = -1$.

Далее мы будем использовать обозначения, которые были приняти во введении.

Замечание 2. В §1 было показано, что условие 1 теоремы 7 для ДЗУП выполнено всегда.

Замечание 3. Учитывая, что $B_1 = B_2 = B$ условие 2 можно переписать в виде

$$(d(-B_1)\cup d(B))\cap d(B_\circ) = \emptyset$$

7.1. Рассмотрим условие 3. Пусть матрицы E и w есть

$$E = \left\| \begin{matrix} I_{pn+pn} & 0 \\ 0 & -I_{pn+pn} \end{matrix} \right\|, \quad w = w_1 + w_2 .$$

Тогда, учитывая вид матриц A_1 и A_2 для ДЗУП, получим

$$T(s) - I = -(sI - ((A_1 - A_2) - BE))(sI + BE)^{-1}.$$

Но

$$A_1 - A_2 = w_1\Pi - w_2\Pi = (w_1 - w_2)\Pi = (w_1 + w_2)E\Pi = w\Pi E .$$

Отсюда $T(s) - I = -(sI - (w\Pi E - BE)(sI + BE)^{-1}$. (28)

Учитывая это, условие (16) примет вид

$$\forall \sigma \in d(-B)\cup d(B)\cup d(B_\circ) \Rightarrow \det(\sigma I - (w\Pi E - BE)) \neq 0 . \quad (29)$$

7.2. Исследование условия 3. Пусть $h_j = (w_{ij})$ и

$$h_j^1 = e_1^2 \otimes h_{-j} + e_2^2 \otimes h_j , \quad h_j^2 = e_1^2 \otimes h_j + e_2^2 \otimes h_{-j}$$

Тогда, учитывая вид матрицы B, получим

$$\sum_i \operatorname{rang} A_{1M(i)} + \sum_i \operatorname{rang} A_{2Q(i)} = \sum_i (\operatorname{rang} h_i^1 + \operatorname{rang} h_i^2) =$$
$$= 2\sum_i \operatorname{rang} h_i^1 = 2\sum_i \operatorname{rang} h_i^2$$

Так как матрицы $h_j^{\varkappa}$ имеют размерности $2np + p$, получим

$$\sum_i \operatorname{rang} A_{1M(i)} + \sum_i \operatorname{rang} A_{2Q(i)} \leq 2np . \quad (30)$$

Нам понадобятся следующие леммы.

Лемма 6. Пусть A есть симметричная, положительно определенная и C симметричная матрицы в некотором гильбертовом пространстве со скалярным произведением $(\cdot,\cdot)$. Тогда AC будет симметрична в пространстве со скалярным произведением

$$(x,y)_2 = (A^{-1}x, y)_1 ,$$

а CA симметрична в пространстве со скалярным произведением -нием

$$(x,y)_3 = (Ax, y)_1 .$$

Доказательство леммы очевидно.

Лемма 7. Имеет место неравенство

$$((\mathcal{W}B^{-1}\Pi - I)x, x) \le (\lambda - 1)(x,x), \quad 0 < \lambda \le 1,$$

где λ есть параметр, входящий в уравнение (2).

Доказательство. Имеем

$$\Delta = ((\mathcal{W}B^{-1}\Pi - I)x, x) = (\mathcal{W}B^{-1}\Pi x, x) - (x,x) \le (\|\mathcal{W}B^{-1}\Pi\|_2 - 1)(x,x),$$

где $\|\cdot\|_2$ есть норма, определяемая скалярным произведением $(\cdot,\cdot)$. Известно

$$\|\mathcal{W}B^{-1}\Pi\|_2 = \theta ,$$

где θ есть максимальное по абсолютной величине собственное значение матрицы $\mathcal{W}B^{-1}\Pi$. Отсюда

$$\theta \le \|\mathcal{W}B^{-1}\Pi\|_1 = \max_{k,m} \sum_i \sum_l \left(\frac{\lambda \alpha_i \beta_l}{4\pi \mu_i B_i}\right)(g_{kmil} + g_{-kmil}) .$$

Учитывая равенства (4)–(6), получаем $\theta \le \lambda$. Значит

$$\Delta \le (\|\mathcal{W}B^{-1}\Pi\|_2 - 1) \le (\lambda - 1)(x,x).$$

Лемма доказана.

Пусть $\lambda < 1$. Ввиду равенств (4)–(5) матрицы $\mathcal{W}$ симметрична. На основании лемм 6 и 7 заключаем, что матрица $(\mathcal{W}\Pi B - BE)$ будет симметрична в пространстве со скалярным произведением

$$(x,y)_1 = -(B^{-1}\Pi(\mathcal{W}B^{-1}\Pi - I)^{-1}x, y) .$$

Отсюда

$$\sum_{\lambda \in \nu(\sigma(R))} \dim \operatorname{Ker}(\lambda I - R) = 2np, \quad R = (\mathcal{W}B^{-1}\Pi - I)BE .$$

Пусть выполнено условие (27). Используя, что ранг произведения вырожденной и невырожденной матрицы равен рангу вырожденной матрицы, получим

$$\sum_{s\in\nu(\Lambda)} \dim \operatorname{Ker}(T(s)-I) = \sum_{s\in\nu(\Lambda)} \dim \operatorname{Ker}(sI-R).$$

Отсюда, учитывая (30) и теорему 5, получаем, что условие 4 теоремы 7 для ДЗУП выполнено, если $0<\lambda<1$.

§ 8. Вид решения ДЗУП

В этом параграфе мы, на основе правил а) – ж), построим решение ДЗУП.

8.аб. Уравнение (27) имеет не более $2np$ корней. Поэтому, согласно правилам а) и б), можем положить

$$\mathcal{J}_1=\mathcal{J}_2=\mathcal{J}, \quad M(\mathcal{J})=\Lambda.$$

8.вг. Применяя правила в) и г) получим, что $G_1=G_2=G$. Учитывая 8.аб и 8.вг, получим, что решение ДЗУП имеет вид

$$x(t)=G\exp(\mathcal{J}t)\xi+G_0\exp(B_0t)\xi_0. \tag{31}$$

8.гд. Вектор ξ, участвующий в (31), удовлетворяет системе

$$\begin{gathered}(I-P_1)x=0,\ x=G[B,\mathcal{J}]\xi+G_0[B,B_0](-\mathbb{1}),\\ (I-P_2)y=0,\ y=G[-B,\mathcal{J}]\exp(\mathcal{J}H)\xi+G_0[-B,B_0]\exp(B_0H)(-\mathbb{1}).\end{gathered} \tag{32}$$

Пусть λ_j есть j-тый элемент диагонали матрицы $\mathcal{J}$, а g_j есть j-тый столбец матрицы G и z_j есть собственный вектор матрицы $(WB^{-1}\Pi-I)BE$, соответствующий собственному значению λ_j. Тогда, ввиду (28), имеет место

$$g_j=\lambda_j z_j+BEz_j.$$

или, обозначая через Z матрицу, составленную из z_j, получим

$$G=Z\mathcal{J}+BEZ. \tag{33}$$

Пусть $\beta=b\otimes I_{p\times p}$. Разобьем матрицы G, Z, $\mathcal{J}$, G_0 на равные блоки

$$G=\begin{Vmatrix}G_{11} & G_{12}\\ G_{21} & G_{22}\end{Vmatrix},\quad Z=\begin{Vmatrix}Z_{11} & Z_{12}\\ Z_{21} & Z_{22}\end{Vmatrix},\quad \mathcal{J}=\begin{Vmatrix}j^1 & 0\\ 0 & j^2\end{Vmatrix},$$

$$G_0 = \begin{Vmatrix} G_{11}^0 & G_{12}^0 \\ G_{21}^0 & G_{22}^0 \end{Vmatrix}, \quad B_0 = \begin{Vmatrix} B_1^0 & 0 \\ 0 & B_2^0 \end{Vmatrix}.$$

<u>**Лемма**</u> **8.** Система (32) имеет вид

$$\begin{aligned} &G_{11}[\beta, j_1]\xi_1 + G_{12}[\beta, j_2]\xi_2 = \eta_0 = \left(G_{11}^0[\beta, B_1^0] + G_{12}^0[\beta, B_2^0]\right)\mathbb{1} \\ &\sum_{\varkappa} G_{2\varkappa}[-\beta, j_\varkappa]\exp(j_\varkappa H)\xi_\varkappa = \eta_H = \left(\sum_{\varkappa} G_{2\varkappa}^0[-\beta, B_\varkappa^0]\exp(B_\varkappa^0 H)\right)\mathbb{1}, \end{aligned} \tag{34}$$

где $e_1^0 \otimes \xi_1 + e_2^2 \otimes \xi_2 = \xi$, $\xi_\varkappa \in R^{np}$.

<u>Доказательство.</u> Имеют место эквивалентности

$$(I - P_1)x = 0 \sim x \in \bigcap_{k=0}^{\infty} \operatorname{Ker}(A_1 B_1^k) \sim A_1 \exp(B_1 t) \equiv 0.$$

Записывая последнее тождество в координатном виде, получим

$$\sum_{j=1}^{np} a_{ij}^1 \exp(b_j^1 t) x_j \equiv 0, \quad i = 1, 2, \dots, 2np.$$

Учитывая, что для ДЗУП $a_{ij}^1 = 0$, если $j > np$, получим

$$\sum_{j=1}^{np} a_{ij}^1 \exp(b_j^1 t) \equiv 0, \quad i = 1, 2, \dots, 2np.$$

Приравнивая к нулю коэффициенты при разных $\exp(b_j^1 t)$, получим

$$\sum_{j \in M(s)} a_{ij}^1 x_j = 0, \quad i = 1, 2, \dots, 2np, \ s = 1, 2, \dots, n.$$

Из равенства (26) следует, что матрица $A_1 M(s)$ имеет свой максимальный ранг p. Отсюда следует, что $x_j = 0$ при $j = 1, \dots, np$, или, учитывая вид вектора x, получим

$$G_{11}[\beta, j_1]\xi_1 + G_{12}[\beta, j_2]\xi_2 = \left(G_{11}^0[\beta, B_1^0] + G_{12}^0[\beta, B_2]\right)\mathbb{1}.$$

Аналогично доказывается вторая строка системы (34). Лемма доказана.

Учитывая равенство (33), имеем

$$G_{11}[\beta, j_1] = (Z_{11} j_1 + \beta Z_{11})[\beta, j_1] = Z_{11}.$$

Аналогично можно показать справедливость равенств

$$G_{12}[\beta, j_2] = Z_{12}, \ G_{21}[-\beta, j_1] = Z_{21}, \ G_{22}[-\beta, j_2] = Z_{22}.$$

Отсюда, система (34) примет вид

$$\begin{aligned} &Z_{11}\xi_1 + Z_{12}\xi_2 = \eta_0, \\ &Z_{12}\exp(j_1 H)\xi_1 + Z_{22}\exp(j_2 H)\xi_2 = \eta_H. \end{aligned} \tag{35}$$

8.е. Столбцы g_j^0 матрицы G_0 удовлетворяют уравнениям

$$(T(b_j^0) - I)\, g_j^0 = a_j^0, \quad j = 1, \dots, 2np,$$

где a_j^0 есть j-тый столбец матрицы A_0. Пусть S есть матрица, столбцы которой удовлетворяют уравнению

$$(b_j I - R) s_j = a_j^\circ , \quad j = 1, \ldots, 2np. \tag{36}$$

Ввиду того, что имеет место (28), получим

$$g_j^\circ = s_j b_j^\circ + BE s_j , \quad j = 1, 2, \ldots, 2np,$$

или, что тоже самое,

$$G_\circ = S B_\circ + BES. \tag{37}$$

Отсюда

$$\eta_\circ = (S_{11} + S_{12}) \mathbb{1} , \quad \eta_H = (S_{21} \exp(B_1^\circ H) + S_{22} \exp(B_2^\circ H)) \mathbb{1} , \tag{38}$$

где S_{ij} и $B_\varkappa^\circ$ таковы, что

$$S = \begin{Vmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{Vmatrix} , \quad B_\varkappa^\circ = \begin{Vmatrix} B_1^\circ & 0 \\ 0 & B_2^\circ \end{Vmatrix} ,$$

а вектор-функции $G_\circ \exp(B_\circ t) \mathbb{1}$ примет вид $(SB_\circ + BES) e^{B_\circ t} \mathbb{1}$.

Пусть $B_\circ = \sigma I$ Матрицы такого вида возникают в случае, если функция $u(\tau, \varphi, \eta)$ в уравнении (2) описывает излучающую способность среды. Учитывая (36) и (37), получим

$$S = (\sigma I - R)^{-1} A_\circ , \quad G_\circ = -(\sigma I + BE) S ,$$

а из (38) имеем, что вектор $\theta = e_1^2 \otimes \eta_\circ + e_2^2 \otimes \eta_H$ удовлетворяет

$$\theta = -\ell S \mathbb{1} ,$$

где

$$\ell = \begin{Vmatrix} I_{np\times np} & 0 \\ 0 & \exp(\sigma H) I_{np\times np} \end{Vmatrix} .$$

Последнее уравнение можно записать в виде

$$(\sigma I - R) \ell^{-1} \theta = -A_\circ \mathbb{1} . \tag{39}$$

В рассматриваемом случае вектор-функция $G_\circ e^{B_\circ t} \mathbb{1}$ будет иметь вид $e^{\sigma t} (\sigma I + BE) \varphi$, где вектор φ удовлетворяет равенству

$$(\sigma I - R) \varphi = -A_\circ \mathbb{1} . \tag{40}$$

Так как матрица Z составлена из собственных векторов матрицы $R = (WB^{-1}\Pi - I) BE$, то

$$R = Z J Z^{-1} .$$

Отсюда

$$x(t)=G\exp(\mathcal{J}t)\xi+G_0\exp(B_0t)(-\mathbb{1})=(Z\mathcal{J}+BEZ)Z Z^{-1}\exp(\mathcal{J}t)Z^{-1}Z\xi$$
$$+G_0\exp(B_0t)(-\mathbb{1})=(R+BE)\exp(Rt)Z\xi+G_0\exp(B_0t)(-\mathbb{1})=$$
$$=W\Pi E\exp(Rt)Z\xi+G_0\exp(B_0t)(-\mathbb{1}).$$

Положим $v(t)=e^{Rt}Z\xi$. Тогда v удовлетворяет уравнению

$$v'=Rv.$$

Пусть $v_1(t)$ и $v_2(t)$ есть вектор-функции длины np, такие что

$$v(t)=e_1^2\otimes v_1(t)+e_2^2\otimes v_2(t).$$

Тогда, учитывая определение вектор-функции v и (35), получим

$$v_1(0)=\eta_0,\quad v_2(H)=\eta_H.$$

Итогом последних двух параграфов является

<u>Теорема 9</u>. Пусть $0<\lambda<1$, где λ есть параметр, входящий в уравнение (2), и пусть выполнены условия (12) и (29). Тогда решение ДЗУП существует, единственно и имеет вид

$$x(t)=W\Pi E\,v(t)+(SB_0+BES)\exp(B_0t)\mathbb{1},$$

где столбцы матрицы S удовлетворяют (36), а $v(t)$ находится из задачи

$$v'=Rv,$$

$$v_1(0)=\eta_0,\ v_2(H)=\eta_H,\ e_1^2\otimes v_1+e_2^2\otimes v_2=v,$$

векторы η_0 и η_H определяются равенством (38). Если $B_0=\sigma I$, то вектора η_0 и η_H определяются из (39) и решение имеет вид

$$x(t)=W\Pi E\,v(t)-\exp(\sigma t)(\sigma I+BE)\varphi,$$

где вектор φ удовлетворяет системе (40).

Автор выражает благодарность профессору Г.М.Вайникко за оказание неоценимой помощи и внимание при написании данной статьи.

Литература

1. Беллман Р., Введение в теорию матриц. Москва, 1976.
2. Канторович Л.В., Акилов Г.П., Функциональный анализ. Москва, 1977.
3. Крейн С.Г. и др., Функциональный анализ. Серия: "Справочная математическая библиотека". Москва, 1972.

4. С о б о л е в В.В., Перенос лучистой энергии в атмосферах звезд и планет. Москва, 1956.
5. Ч а н д р а с е к а р С., Перенос лучистой энергии. Москва, 1953.

Поступило
7 II 1979

STRUCTURE OF THE SOLUTION OF SYSTEMS OF INTEGRAL EQUATIONS CONNECTED WITH APPROXIMATE SOLUTION OF THE RADIATIVE TRANSFER PROBLEM

J. Knyazikhin

Summary

In the present paper the class of integral equation (8) which can be obtained by solving approximately equation (2) is given. The conditions are given under which equation (1) can be solved exactly and has the form (11). Equation (8) is considered in detail. The main result is presented by theorem 7.

Уч.зап.Тартуск.ун-та,1979,500,92-104

О БЫСТРОТЕ СХОДИМОСТИ МЕТОДА ДИСКРЕТНЫХ ОРДИНАТ В СЛУЧАЕ КВАДРАТУР ГАУССА

А. Маршак

Тартуский государственный университет

Введение

Метод дискретных ординат решения интегро-дифференциального уравнения переноса основывается на замене интеграла какой-то квадратурой и решении получаемой системы обыкновенных дифференциальных уравнений. Этот метод хорошо зарекомендовал себя в вычислительной практике и широко используется в атмосферной оптике (см. [4, 6]).

В работе [2] была подробно изучена быстрота сходимости метода дискретных ординат в случае квадратурной формулы прямоугольников, указана удобная численная характеристика для сравнения различных квадратурных формул, работающих в этом методе, и приведены численные результаты. В настоящей статье по подходу, разработанному в [2], изучается быстрота сходимости в случае квадратур Гаусса (отличных от рекомендованных в [6]) и приводятся оценки, согласующиеся с численными результатами работы [2].

§ 1. Решаемая задача

Рассмотрим одно из основных уравнений теории переноса излучения в однородной плоско-параллельной среде - уравнение Милна (см. [4, 6])

$$u(\tau) = \frac{\lambda}{2} \int_0^H E(|\tau - s|)\, u(s)\, ds + f(\tau), \tag{1}$$

где

$$E(t) = \int_0^1 \exp(-t/\mu)/\mu \, d\mu \quad (t>0). \tag{2}$$

- интегральная экспонента. Задача {(1), (2)} описывает перенос излучения в плоско-параллельной однородной среде оптической толщины H со сферической индикатрисой рассеяния и с вероятностью λ $(0 \le \lambda \le 1)$ выживания кванта в акте столкновения с частицей среды.

Напомним некоторые свойства интегральной экспоненты (см. [4]). Функция $E(t)$ бесконечно раз дифференцируема при $t>0$ и имеет логарифмическую особенность при $t=0$

$$E(t)>0, \quad \int_0^\infty E(t)\,dt = 1 .$$

Последнее соотношение гарантирует однозначную разрешимость в $C[0,H]$ уравнения (1) при любом $f \in C[0,H]$, ибо для оператора

$$(Tu)(\tau) = \frac{\lambda}{2}\int_0^H E(|t-s|)\,u(s)\,ds$$

имеем

$$\|T\|_{C\to C} = \frac{\lambda}{2}\max_{0\le\tau\le H}\int_0^H E(|\tau-s|)\,ds < \lambda \le 1 .$$

Если $f\in C^1[0,H]$, то решение u интегрального уравнения (1) непрерывно на $[0,H]$ непрерывно дифференцируемо в промежутке $(0,H)$ и удовлетворяет неравенству

$$|u'(\tau)| \le c\left(|\ln\tau| + |\ln(H-\tau)|\right), \; 0<\tau<H . \qquad (3)$$

Метод дискретных ординат решения уравнения (1) основывается (см. [1, 6, 9]) на замене интеграла (2) какой-то квадратурой

$$\int_0^1 f(\mu)\,d\mu \approx \sum_{j=1}^n \alpha_j f(\mu_j), \qquad (4)$$

где $\alpha_j>0$, $\mu_j \in (0,1]$ $(j=1,\dots,n)$. Будем дополнительно предполагать, что она точна для постоянной функции, т.е.

$$\sum_{j=1}^n \alpha_j = 1 \quad (n=1,2,\dots) . \qquad (5)$$

Таким образом,

$$E(t)\approx E_n(t) = \sum_{j=1}^n \alpha_j \exp(-t/\mu_j)/\mu_j ,$$

и уравнение (I) заменяется на уравнение

$$u_n(\tau) = \frac{\lambda}{2}\int_0^H E_n(|\tau-s|)\,u_n(s)\,ds + f(\tau) . \qquad (6)$$

Обозначим через T_n интегральный оператор уравнения (6)

$$(T_n u)(\tau) = \frac{\lambda}{2}\int_0^H E_n(|\tau-s|)\,u(s)\,ds .$$

В случае сходящегося квадратурного процесса (4) легко установить сходимость (см. [8])

$$\|T_n - T\|_{C\to C} \le \lambda \int_0^H |E_n(\tau) - E(\tau)|\, d\tau \to 0 \quad \text{при} \quad n\to\infty.$$

Поэтому решения u_n интегральных уравнений (6) при $n\to\infty$ равномерно сходятся к решению интегрального уравнения (1).

Целью данной работы является изучение быстроты сходимости $u_n \to u$ в случае квадратуры Гаусса (4).

Замечание 1. В работе [2] были получены оценки скорости сходимости $u_n \to u$ в случае квадратурной формулы прямоугольников, а именно:

$$c_0 h \le \|u_n - u\|_C \le c_1 h,$$

$$\|u_n - u\|_L \le c_2 h,$$

$$|u_n(\tau) - u(\tau)| \le c_1 \min\{h,\ h^3/\tau^2 + h^2|\ln h|,\ h^3/(H-\tau)^2 + h^2|\ln h|\} \ (0 \le \tau \le H).$$

В работе [2] показано, что для решений u и u_n уравнений (1) и (6) имеет место равенство

$$u_n - u = (I - T_n)^{-1}(T_n u - Tu), \tag{7}$$

причем при достаточно больших n операторы $I - T_n$ обратимы и

$$\|(I-T_n)^{-1}\|_{C\to C} = \|(I-T_n)^{-1}\|_{L\to L} \le c = \mathrm{const},\ n \ge n_0,$$

и вывод оценок погрешности сводится к изучению малости функции $T_n u - Tu$. При этом, с учетом (5),

$$(T_n u - Tu)(\tau) = \frac{\lambda}{2}\{u(0)v_n^0(\tau) + u(H)v_n^1(\tau) + w_n^0(\tau) + w_n^1(\tau)\}, \tag{8}$$

где

$$v_n^0(\tau) = \int_0^1 \exp(-\tau/\mu)\, d\mu - \sum_{j=1}^n \alpha_j \exp(-\tau/\mu_j),\ v_n^1(\tau) = v_n^0(H-\tau), \tag{9}$$

$$w_n^0(\tau) = \int_0^\tau v_n^0(\tau-\sigma)u'(\sigma)\,d\sigma,\ w_n^1(\tau) = -\int_\tau^H v_n^1(\sigma-\tau)u'(\sigma)\,d\sigma. \tag{10}$$

Из (8) - (10) видно, что точность приближения зависит от того, насколько точно квадратурная формула (4) аппроксимирует интегралы от функций

$$\psi_\tau(\mu) = \exp(-\tau/\mu) \qquad (0 \le \tau \le H).$$

Заметим, что $\psi_\tau \in C^\infty[0,1]$, $\psi_\tau^{(k)}(0) = 0$, однако, производные возле $\mu = 0$ не ограничены в совокупности, причем

$$|\psi_\tau^{(k)}(\mu)| \le \gamma_k \tau^{-k} \ (0 \le \mu \le 1,\ 0 < \tau \le H,\ k = 0,1,\dots), \tag{11}$$

$$|\psi_\tau^{(k)}(\mu)| \le \gamma_k' \mu^{-k} \ (0 < \mu \le 1,\ 0 \le \tau \le H,\ k = 0,1,\dots) \tag{12}$$

с некоторыми постоянными γ_k и γ_k'.

Нам понадобятся известные выражения для остаточного члена формулы прямоугольников:

$$\int_a^b f(x)dx = (b-a)f\left(\frac{a+b}{2}\right) + \frac{(b-a)^3}{24} f''(\xi^*), \quad \xi^* \in [a,b], \tag{13}$$

$$\int_a^b f(x)dx = (b-a)f\left(\frac{a+b}{2}\right) + r_0, \quad |r_0| \le (b-a) \bigvee_a^b (f), \tag{14}$$

причем

$$\bigvee_a^b (\psi_\tau) = \exp(-\tau/b) - \exp(-\tau/a) \le 1 \quad (0 \le \tau \le H).$$

Пусть $P_m(\psi_\tau, \mu) = c_0 + \cdots + c_m \mu^m$ - многочлен степени не выше, чем m, наименее уклоняющийся от функции ψ_τ на отрезке $[0,1]$, тогда

$$E_m(\psi_\tau) = \|\psi_\tau - P_m\|_{C[0,1]} = \inf_{P_m^*} \|\psi_\tau - P_m^*\|_{C[0,1]}$$

- есть наилучшее приближение порядка m функции ψ_τ. Имеет место неравенство (см. [5])

$$E_m(\psi_\tau) \le M_k (1/m)^k \omega(\psi_\tau^{(k)}, 1/m), \quad k = 0,1,\ldots, \tag{15}$$

где M_k - константа, зависящая только от k, а

$$\omega(f,\theta) = \sup\{|f(x_1) - f(x_2)| : |x_1 - x_2| \le \theta, \ x_1, x_2 \in [0,1]\}$$

- модуль непрерывности.

§ 2. Замечания о формуле Гаусса

Пусть в процессе (4) квадратурой является формула Гаусса с постоянным весом на $[0,1]$, причем $\mu_j = (1-\mu_j')/2$, а $\alpha_j = \alpha_j'/2$ $(j = 1,\ldots,n)$, где $\{\alpha_j'\}_1^n, \{\mu_j'\}_1^k$ - соответственно веса (коэффициенты Кристофеля) и узлы (корни полинома Лежандра степени n на $[-1,1]$). Очевидно, что условие (5) в этом случае выполнено. Известно (см. [3]), что узлы полинома Лежандра лежат между соответствующей суммой коэффициентов Кристофеля, т.е.

$$-1 + \sum_{i=1}^{j-1} \alpha_i' < \mu_j' < -1 + \sum_{i=1}^{j} \alpha_i' \quad (j = 1,\ldots,n),$$

что наводит на мысль представить формулу Гаусса на каждом отрезке $[a_{j-1}, a_j]$ $(a_j = \alpha_1 + \cdots + \alpha_j, \ j = 1,2,\ldots,n, \ a_0 = 0)$ как формулу средних прямоугольников, для которой остаточные члены определяются из (13) и (14).

В [3] даны асимптотики узлов и весов формулы Гаусса[1]

$$\begin{cases} \alpha_i' = \alpha_{in}' \cong \frac{2\pi}{n} \sin\frac{\theta_i}{2}\cos\frac{\theta_i}{2} = \frac{\pi}{n}\sin\theta_i \,; \; \mu_i' = \cos\theta_i \,, \\ \frac{i-1/2}{n+1/2} < \theta_i < \frac{i}{n+1/2} \,, \quad i = 1,2,\dots,n \,. \end{cases}$$

Отсюда

$$\mu_i = \sin^2\frac{i+\varepsilon_i}{2(n+1/2)}\pi \,; \; \alpha_i \cong \frac{\pi}{2n}\sin\frac{i+\varepsilon_i}{n+1/2}\pi \quad (i=1,\dots,n,\ \varepsilon_i \in (-\tfrac{1}{2},0)), \tag{16}$$

а

$$a_j = \sum_{i=1}^{j} \alpha_i \cong \sin\frac{(j+\varepsilon_j')\pi}{2(n+1/2)} \sin\frac{j}{2(n+1/2)}\pi \quad (j=1,\dots,n,\ \varepsilon_j' \in (0,1)). \tag{17}$$

Выражение (17) вытекает из цепочки равенств[2]

$$\sum_{i=1}^{j} \sin iu = \frac{1}{2\sin\frac{u}{2}} \sum_{i=1}^{j} 2\sin\frac{u}{2}\sin iu = \frac{1}{2\sin\frac{u}{2}} \sum_{i=1}^{j} (\cos(i-1/2)u -$$

$$-\cos(i+1/2)u) = \frac{\cos\frac{u}{2} - \cos(j+1/2)u}{2\sin\frac{u}{2}} = \frac{\sin\frac{j+1}{2}u \,\sin\frac{j}{2}u}{\sin u/2} \quad (j=1,\dots,n).$$

<u>Лемма 1</u>. Пусть μ_j – узлы квадратурной формулы Гаусса, а μ_j^* – узлы формулы прямоугольников на каждом отрезке $[a_{j-1},a_j]$ $(j=1,\dots,n,\ a_0=0)$. Тогда, начиная с некоторого $n \geqslant n_0$,

$$|\mu_j^* - \mu_j| \leqslant \frac{3\pi^2}{16} h^2 \left|\cos\frac{j+3/8}{n+1/2}\right| \quad \left(h=\frac{1}{n},\ j=1,\dots,n\right). \tag{18}$$

<u>Доказательство</u>. Здесь везде $j=1,\dots,n$. По условию леммы имеем

$$\mu_j^* - \mu_j = a_j - \mu_j - \alpha_j/2 \,, \quad \mu_j - \mu_j^* = \mu_j - a_{j-1} - \alpha_j/2 \,.$$

Но из (16) и (17) следует, что

$$a_j - \mu_j \cong \frac{\varepsilon_j' - 2\varepsilon_j}{4(n+1/2)}\pi \sin\frac{2j+\varepsilon_j+\varepsilon_j'/2}{2(n+1/2)}\pi = x(\varepsilon_j',\varepsilon_j)$$

и

$$\mu_j - a_{j-1} \cong \frac{2+2\varepsilon_j - \varepsilon_{j-1}'}{4(n+1/2)}\pi \sin\frac{2j-1+\varepsilon_j+\varepsilon_{j-1}'/2}{2(n+1/2)}\pi = y(\varepsilon_j,\varepsilon_{j-1}'),$$

---

[1] Запись $z_n \cong \omega_n$ означает, что $z_n/\omega_n \to 1$ при $n \to \infty$.

[2] На возможность суммирования асимптотических равенств в данном случае указывает тот факт, что

$$a_n = \sum_{i=1}^{n} \alpha_{in} = \sum_{i=1}^{n} \beta_{in} + o(1/n^2) \,, \quad \text{где } \alpha_{in} \cong \beta_{in}.$$

причем

$$\sup_{\varepsilon_j', \varepsilon_j} x(\varepsilon_j', \varepsilon_j) = x(1, -1/2), \sup_{\varepsilon_j, \varepsilon_{j-1}'} y(\varepsilon_j, \varepsilon_{j-1}') = y(0,0).$$

Однако, для любого $\varepsilon_{j-1}' \in (0,1)$ имеем

$$y(-1/2, \varepsilon_{j-1}') \le \frac{\pi}{4(n+1/2)} \sin \frac{j-1/4}{n+1/2}\pi$$

и для любого $\varepsilon_j' \in (0,1)$ имеем

$$x(\varepsilon_j', 0) \le \frac{\pi}{4(n+1/2)} \sin \frac{j+1/4}{n+1/2}\pi .$$

Отсюда для любых $\varepsilon_{j-1}' \in (0,1)$, $\varepsilon_j \in (-1/2, 0)$ и $\varepsilon_j' \in (0,1)$ начиная с некоторого $n \geqslant n_0$, справедливо неравенство

$$\min\{a_j - \mu_j, \mu_j - a_{j-1}\} \le \frac{\pi}{4(n+1/2)} \sin \frac{j+\varepsilon_j''}{n+1/2}, \ \varepsilon_j'' \in [-1/4, 1/4].$$

Далее, вычитая половину j-го веса из (16), приходим к оценке (18). Лемма доказана.

Нам также понадобится следующий вспомогательный результат. Сформулируем его в виде леммы.

<u>Лемма 2</u>. Если $\{\alpha_j\}_1^n$ и $\{\mu_j\}_1^n$ - соответственно веса и узлы квадратурной формулы Гаусса, определенные в (16), то для любого $k = 0, 1, 2, \cdots$ имеют место равенства

$$\sum_{j=1}^{n} \frac{\alpha_j}{\mu_j^{2k+1}} \cong \begin{cases} 2(1+1/2n)(\ln n + \gamma + c_n^0) + c_n^*, & k=0 \\ 2(1+1/2n) b^{(k)} (n+1/2)^{2k} \left[c_n^k + O\left(\frac{\ln n}{(n+1/2)^2}\right)\right], & k=1,2,\cdots, \end{cases} \tag{19}$$

где γ - постоянная Эйлера, $b^{(k)} \in ((2/\pi)^{2k}, 1)$, $k = 1, 2, \cdots$, $c_n^* \to c^* < \infty$; $c_n^k \to c^k < \infty$, $k = 0, 1, \cdots$, причем

$$c^k = \begin{cases} \sum_{i=1}^{\infty} \frac{\varepsilon_i}{i(i+\varepsilon_i)}, & \varepsilon_i \in (-1/2, 0), k=0, \\ \sum_{i=1}^{\infty} \frac{1}{(i+\varepsilon_i)^{2k+1}}, & \varepsilon_i \in (-1/2, 0), k=1,2,\dots . \end{cases}$$

<u>Доказательство</u>. Из (16) вытекает

$$\sum_{j=1}^{n} \alpha_j / \mu_j^{k+1} \cong \frac{\pi}{n} \sum_{j=1}^{n} \frac{\operatorname{ctg} y_j}{\sin^{2k} y_j}, k=0,\dots, y_j = \frac{j+\varepsilon_j}{2(n+1/2)}\pi, \ \varepsilon_j \in (-1/2, 0).$$

Раскладывая $\operatorname{ctg} y$ в ряд и представляя $\sin^{2k} y$ как $y^{2k}(\sin^{2k} y / y^{2k})$, получим

$$\sum_{j=1}^{n} \alpha_j / \mu_j^{k+1} \cong \frac{\pi}{n} \left\{ \sum_{j=1}^{n} \frac{1}{y_j^{2k+1}} A(y_j^{(k)}) - \sum_{j=1}^{n} \sum_{i=1}^{\infty} \frac{2^{2i} B_i}{(2i)!} y_j^{2i-2k-1} A(y_j^{(k)}) \right\}, \tag{20}$$

где $A(y_j^{(k)}) = \frac{y_j^{2k}}{\sin^{2k} y_j}$ $(j = 1, 2, \cdots, n, k = 0, 1, \cdots; y_j \in (0, \pi/2))$,

13

причем

$$1 < A(y^{(k)}) < (\pi/2)^{2k} \quad (k=0,1,\dots;\ y^{(k)} \in (0,\pi/2)). \quad (21)$$

Здесь B_i $(i=1,2,\dots)$ – числа Бернулли, которые выражаются через дзета-функцию Римана ζ (см. [7]) как

$$B_i = \frac{(2i)!}{(2\pi)^{2i}} \zeta(2i) \quad (i=1,2,\dots) \qquad (22)$$

Обозначим через r_n – сумму отрицательных членов асимптотического равенства (20), а за r_n^1 – сумму положительных. Тогда, с учетом (22), можно записать

$$r_n = \frac{2}{n}\pi^{-2k} \sum_{j=1}^{n} A(y_j^{(k)}) \sum_{i=1}^{\infty} \zeta(2i)\left[\frac{j+\varepsilon_j}{2(n+1/2)}\right]^{2i-2k-1} \quad (k=0,1,\dots).$$

По формуле суммы бесконечно-убывающей геометрической прогрессии получим

$$\sum_{i=1}^{\infty}\left[\frac{j+\varepsilon_j}{2(n+1/2)}\right]^{2i-2k-1} < \frac{1}{3}2^{2k+1}\left[\frac{n+1/2}{j+\varepsilon_j}\right]^{2k-1} \quad (j=1,\dots,n;\ k=0,1,\dots).$$

Теперь, учитывая неравенство (см. [7])

$$\zeta(2i) < \zeta(2) = \pi^2/6 \quad (i=1,2,\dots)$$

и неравенства (21), приходим к оценке

$$r_n < \frac{2}{9}\pi^2(1+1/n2)(n+1/2)^{2k-2}\sum_{j=1}^{n}\left[\frac{1}{j+\varepsilon_j}\right]^{2k-1} \quad (k=0,1,\dots).$$

Применяя также неравенства (21), легко увидеть, что

$$r_n^1 = 2(1+1/2n)\,b^{(k)}(n+1/2)C_n^k,\ k=0,1,2,\dots,$$

где $b^{(k)} \in ((2/\pi)^{2k}, 1)$, а

$$C_n^k \longrightarrow \sum_{j=1}^{\infty}[1/(j+\varepsilon_j)]^{2k+1} \quad (k=0,1,\dots)$$

Далее, сравнивая неравенство для r_n и равенство для r_n', с учетом того, что

$$\sum_{j=1}^{n}\frac{1}{j+\varepsilon_j} = \sum_{j=1}^{n} 1/j - \sum_{j=1}^{n}\frac{\varepsilon_j}{j(j+\varepsilon_j)} = \ln n + \gamma + C_n^0 \quad (n=1,2,\dots),$$

где γ – постоянная Эйлера, а $C_n^0 \to C^0 < \infty$, приходим для любого $k=0,1,2,\dots$ к формуле (19). Лемма доказана.

Справедлива

<u>Лемма 3</u>. Для квадратурного процесса Гаусса имеет место неравенство

$$\left|\int_0^1 f(\mu)d\mu-\sum_{j=1}^{n}\alpha_j f(\mu_j)\right|\le 2M_k\left(1/(2n-1)\right)^k\omega\left(f^{(k)},1/(2n-1)\right), \quad (23)$$

где M_k и $\omega^d(f,\theta)$ определенны неравенством (15).

Доказательство следует из точности квадратур Гаусса для полиномов степени $2n-1$ и неравенства (15).

§ 3. Точность аппроксимации (4) в случае квадратур Гаусса

Представим квадратурный процесс Гаусса (4),(5) в виде

$$\int_0^1 \psi_\tau(\mu)d\mu-\sum_{j=1}^{n}\alpha_j\psi_\tau(\mu_j)=\int_0^{a_1}[\psi_\tau(\mu)-\psi_\tau(\mu_1)]d\mu+\sum_{j=2}^{n}\int_{a_{j-1}}^{a_j}[\psi_\tau(\mu)-\psi_\tau(\mu_j)]d\mu$$

где α_j,μ_j и a_j $(j=1,2,\dots,n)$ определенны соответственно равенствам (16) и (17). Первый интеграл оценивается на основании (14) как ch^2 $(c=const,\ h=1/n)$, а остальные $n-1$ интегралов представим в виде:

$$\sum_{j=2}^{n}\int_{a_{j-1}}^{a_j}[\psi_\tau(\mu)-\psi_\tau(\mu_j)]d\mu=\sum_{j=2}^{n}\int_{a_{j-1}}^{a_j}[\psi_\tau(\mu)-\psi_\tau(\mu_j^*)]d\mu+\sum_{j=2}^{n}\alpha_j[\psi_\tau(\mu_j^*)-\psi_\tau(\mu_j)] \quad (24)$$

где μ_j^* – середины отрезков $[a_{j-1},a_j]$ $(j=2,3,\dots,n)$. Интеграл в (24) оценивается на основании (13) как $\alpha_j^3\psi_\tau^{(2)}(\xi_j)/24$, $\xi_j\in[a_{j-1},a_j]$, $j=2,3,\dots,n$. Просуммировав это выражение от 2 до n, учитывая (12) при $k=2$ и (19) при $k=0$, получим

$$\left|\frac{1}{24}\sum_{j=2}^{n}\alpha_j^3\psi_\tau^{(2)}(\xi_j)\right|\le\frac{\gamma_2}{24}\sum_{j=2}^{n}\alpha_j^3/a_j^2=ch^2\sum_{j=2}^{n}\alpha_j/a_j\le ch^2|\ln h|,\ c=const.$$

Сумма в (24) оценивается аналогично, с учетом (18), (12) при $k=1$ и (19) при $k=0$.

Таким образом

$$\left|\int_0^1\psi_\tau(\mu)-\sum_{j=1}^{n}\alpha_j\psi_\tau(\mu_j)\right|\le ch^2|\ln h|,\ c=const,\ h=1/n. \quad (25)$$

Теперь мы в состоянии доказать следующий результат.

Лемма 4. Для определенных в (9), (10) функций v_n^0, v_n^1, w_n^0, w_n^1 в случае квадратурной формулы Гаусса (4),(5) справедливы оценки:

$$|v_n^0(\tau)|\le ch^2|\ln h|,\quad |v_n^1(\tau)|\le ch^2|\ln h|,\quad (0\le\tau\le H) \quad (26)$$

$$\int_0^H|v_n^0(\tau)|d\tau=\sigma(h^2),\quad \int_0^H|v_n^1(\tau)|d\tau=\sigma(h^2), \quad (27)$$

$$\max_{0\le\tau\le H}|w_n^0(\tau)|=\sigma(h^2),\quad \max_{0\le\tau\le H}|w_n^1(\tau)|=\sigma(h^2), \quad (28)$$

$$\int_0^H |w_n^0(\tau)|\,d\tau = o(h^2), \quad \int_0^H |w_n^1(\tau)|\,d\tau = o(h^2). \tag{29}$$

Доказательство. Оценки (26) немедленно следуют из (25). Для получения оценки (27) оценим в равенстве

$$\|v_n^0\|_{L[0,H]} = \int_0^{h^{1/p}} |v_n^0(\tau)|\,d\tau + \int_{h^{1/p}}^H |v_n^0(\tau)|\,d\tau \quad (p>1)$$

первый интеграл на основании (26), а второй на основании (23) и (11). Тогда при $p = 1+\varepsilon$ $(\varepsilon>0)$ получим

$$\|v_n^0\|_{L[0,H]} \le \inf_{\varepsilon>0} \max\left\{ h^{(3+2\varepsilon)/(1+\varepsilon)} |\ln h|, \frac{M_k \gamma_{k+1}}{2^k \cdot k} h^{k\varepsilon/(1+\varepsilon)+1} \right\}, \tag{30}$$

где M_k, γ_k – константы неравенств (23) и (11), зависящие только от k. Отсюда вытекает оценка (27) для v_n^0; для v_n^1 – она получается аналогично.

При $0 \le \tau \le h^{1/p}$ $(p>1)$ на основании (3) и (26) находим

$$|w_n^0(\tau)| \le \int_0^\tau |v_n^0(\tau-\sigma)|\,|u'(\sigma)|\,d\sigma \le c h^{2+1/p} \ln^2 h .$$

При $\tau > h^{1/p}$ разобьем интеграл на три части от 0 до $h^{1/p}$, от $h^{1/p}$ до $\tau - h^{1/p}$ и от $\tau - h^{1/p}$ до τ, затем оценим первый и третий – на основании (3) и (26), а второй – на основании (27). В результате получим оценку (28) для w_n^0, для w_n^1 она устанавливается аналогично. Оценки (29) очевидны. Лемма доказана.

Замечание 2. Из неравенства (30) следует, что интегральная оценка (27), а следовательно и оценки (28) – (29), могут иметь порядок близкий к третьему, т.е. $h^{3-\varepsilon}$ $(0<\varepsilon\ll 1)$. Но при этом необходимо использовать большие значения k, а константы γ_k $(k=0,1,\dots)$ неравенства (11) имеют вид

$$\gamma_k = \exp(-1/d_k) \sum_{j=0}^{k-1} c_j^{(k-1)} / d_k^{2k-j},$$

где $c_j^{(k)}$ – константы, зависящие от k, а d_k есть наименьший корень уравнения

$$\sum_{j=0}^{k} c_j^{(k)} d_j = 0 ,$$

что указывает на быстрое возрастание γ_k по мере увеличения k $(k=0,1,\dots)$. Например, для v_n^0 имеют место интегральные оценки

$$\|v_n^0\|_{L[0,H]} \le 2{,}7 \cdot h^{2,5}, \quad \|v_n^0\|_{L[0,H]} \le 56{,}9\, h^{2,75}.$$

§ 4. Оценки сходимости

В этом параграфе мы сформулируем и докажем основной результат быстроты сходимости $u_n \longrightarrow u$ $(n \to \infty)$ при применении квадратурной формулы Гаусса.

Теорема. Если $f \in C^1[0,H]$ и процесс (4),(5) есть квадратура Гаусса, тогда при выполнении хотя бы одного из неравенств

$$u(0) \neq 0, \; u(H) \neq 0 \tag{31}$$

справедливы оценки

$$c_1 h^2 \leq \|u_n - u\|_{C[0,H]} \leq c_2 h^2 |\ln h|. \tag{32}$$

(Оценка сверху верна и без предположения (31), c_1, c_2 – постоянные.)

Доказательство. Оценка сверху немедленно следует из (7), с учетом равномерной ограниченности операторов, обратных к $I - T_n$, $n \geq n_0$ равенства (8) и оценок (26), (28). Для доказательства нижней оценки (32) обозначим

$$g(\tau) = \int_0^1 exp(-\tau/\mu)\,d\mu, \quad g_n(\tau) = \sum_{j=1}^{n} \alpha_j \, exp(-\tau/\mu_j).$$

Поскольку $g(0) = 1$, $g'(\tau) = -E(\tau)$, то, используя известное разложение (см. [4])

$$E(\tau) = -\ln\tau - \gamma + \tau - \frac{\tau^2}{2 \cdot 2!} + \frac{\tau^3}{3 \cdot 3!} - \cdots \quad (\tau > 0)$$

(γ = 0,5572 – постоянная Эйлера), получаем для малых τ

$$g(\tau) = g(0) + \int_0^\tau g'(\tau)\,d\tau = 1 + \tau(\ln\tau - 1 - \gamma) + O(\tau^2).$$

С другой стороны, воспользуясь соотношением (5), получим

$$g_n(\tau) = 1 - \tau \sum_{j=1}^{n} \left\{ \frac{\alpha_j}{\mu_j} - \frac{\tau}{2!}\frac{\alpha_j}{\mu_j^2} + \frac{\tau^2}{3!}\frac{\alpha_j}{\mu_j^3} - \cdots \right\}.$$

Теперь, обозначив $s = \tau(n+1/2)^2$, учитывая соотношение (19) и соотношение $1 + 1/2 + \cdots + 1/n - \ln n + \gamma \longrightarrow 0$, имеем

$$g_n(\tau) = 1 - \tau \Big\{ 2(1 + 1/2n) \Big[\Big(\ln n + \gamma + \sum_{j=1}^{n} \Big(\frac{1}{j+\varepsilon_j} - \frac{1}{j} \Big) + c - \frac{\tau}{2!} \times$$

$$\times (n+1/2)^2 b^{(1)} \Big(\sum_{j=1}^{n} \frac{1}{(j+\varepsilon_j)^3} + O\Big(\frac{\ln n}{(n+1/2)^2} \Big) \Big) - \frac{\tau^2}{3!}(n+1/2)^4 b^{(2)} \Big(\sum_{j=1}^{n} \frac{1}{(j+\varepsilon_j)^5} +$$

$$+ O\Big(\frac{\ln n}{(n+1/2)^2} \Big) + \cdots \Big] \Big\} = 1 - \tau \Big\{ 2(1 + 1/2n) \big[\ln n + \gamma + d_n(s) \big] \Big\},$$

причем при $n \to \infty$ имеет место сходимость

$$d_n(s) \to d(s) \equiv c + \sum_{j=1}^{\infty}\left(\frac{1}{j+\varepsilon_j} - \frac{1}{j}\right) - \frac{s}{2!} b^{(1)} \sum_{j=1}^{\infty} \frac{1}{(j+\varepsilon_j)^3} - \frac{s^2}{3!} b^{(2)} \sum_{j=1}^{\infty} \frac{1}{(j+\varepsilon_j)^5} + \ldots$$

где $b^{(k)} \in [(2\pi)^{2k}, 1]$. Сопостовляя асимптотические разложения g и g_n, получаем

$$g\left(\frac{s}{(n+1/2)^2}\right) - g_n\left(\frac{s}{(n+1/2)^2}\right) = \frac{s}{(n+1/2)^2}\left\{\ln s - \gamma - 1 - d_n^0(s)\right\} + O\left(\frac{s^2}{(n+1/2)^4}\right),$$

где

$$d_n^0(s) \equiv 2(1+1/2n)d_n(s) - 2\ln(1+1/2n) + \frac{\gamma}{n} + \frac{\ln n}{2n} = 2(1+1/2n)d_n(s) + \varepsilon_n \to 2d(s).$$

Зафиксируем $s \in (0,1]$, так, что $\ln s - \gamma - 1 - 2d(s) \neq 0$. На возможность этого шага указывает то, что функция

$$r(s) = \ln s - \gamma - 1 - 2d(s)$$

суть аналитическая, неравная тождественно нулю. Тогда

$$\left|v_n^0\left(\frac{s}{(n+1/2)^2}\right)\right| = \left|g\left(\frac{s}{(n+1/2)^2}\right) - g_n\left(\frac{s}{(n+1/2)^2}\right)\right| \geq a h^2$$

Поскольку на основании (28)

$$\left|w_n^0\left(\frac{s}{(n+1/2)^2}\right)\right| = o(h^2), \quad \left|w_n^1\left(\frac{s}{(n+1/2)^2}\right)\right| = o(h^2),$$

а $v_n^1(\tau)$ при малых τ есть величина, заведомо меньшая, чем ah^2, то из (8) в случае $u(0) \neq 0$ получаем

$$\left|(T_n u - Tu)\left(\frac{s}{(n+1/2)^2}\right)\right| \geq a_0 h^2 \quad (n \geq n_0', \ a_0 = const > 0)$$

и по (7) для $n \geq n_0'$ имеем

$$\|u_n - u\|_{C[0,H]} \geq \frac{1}{1+\|T\|} \|T_n u - Tu\|_{C[0,H]} \geq c_1 h^2,$$

$$c_1 = const > 0.$$

Мы установили нижнюю оценку (32) в случае $u(0) \neq 0$, случай $u(H) \neq 0$ аналогичен. Доказательство теоремы завершено.

Замечание 3. Согласно неравенству (30) и замечанию 2 для интегральной оценки скорости сходимости $u_n \to u$ $(n \to \infty)$ имеет место неравенство аналогичное (30).

Замечание 4. Оценки (32) удачно согласуются с численными результатами статьи [2] для формулы Гаусса с постоянным весом на $[0,1]$.

§ 5. Связь с интегродифференциальным уравнением переноса

Рассмотрим следующую краевую задачу атмосферной оптики (см. [4], [6])

$$\mu \frac{\partial \varphi(\tau,\mu)}{\partial \tau} + \varphi(\tau,\mu) = \frac{\lambda}{2} \int_{-1}^{1} \varphi(\tau,\mu')d\mu' + f(\tau) \qquad (33)$$

$$(0 \le \tau \le H,\ -1 \le \mu \le 1,\ \lambda \in [0,1])$$

$\varphi(0,\mu) = 0$ при $\mu > 0$, $\varphi(H,\mu) = 0$ при $\mu < 0$.

С помощью замены

$$u(\tau) = \frac{\lambda}{2} \int_{-1}^{1} \varphi(\tau,\mu')d\mu' + f(\tau) \qquad (34)$$

краевая задача (33) легко сводится (см. [4]) к интегральному уравнению (1). Решив уравнение (34), решение краевой задачи востанавливается по формуле

$$\varphi(\tau,\mu) = \begin{cases} \frac{1}{\mu} \int_0^{\tau} u(\tau') \exp(-(\tau-\tau')/\mu)\, d\tau' & \text{при } \mu > 0, \\ u(\tau), & \text{при } \mu = 0, \\ \frac{1}{|\mu|} \int_{\tau}^{H} u(\tau') \exp(-|\tau-\tau'|/\mu)\, d\tau' & \text{при } \mu < 0. \end{cases} \qquad (35)$$

Решая задачу (33) методом дискретных ординат, мы получим, что решения соответствующего дискретного аналога уравнения (33) φ_n равномерно сходятся к φ – решению краевой задачи (33), причем из (35) и его дискретного аналога следует, что (см. [2])

$$|\varphi_n(\tau,\mu) - \varphi(\tau,\mu)| \le \|u_n - u\|_{C[0,H]} \quad (0 \le \tau \le H,\ -1 \le \mu \le 1)$$

Отсюда следует равномерная оценка

$$c_1 h^2 \le \max_{-1 \le \mu \le 1} \|\varphi_n - \varphi\|_{C[0,H]} \le c_2 h^2 |\ln h|.$$

Замечание 5. Аналогично замечанию I в работе [2] были полученны оценки скорости сходимости $\varphi_n \to \varphi$ $(n \to \infty)$ в случае квадратурной формулы прямоугольников, а именно:

$$\max_{-1 \le \mu \le 1} \|\varphi_n - \varphi\|_{L[0,H]} \le c_2 h^2,$$

$$c_0 h_1 \le \max_{-1 \le \mu \le 1} \|\varphi_n - \varphi\|_{C[0,H]} \le c_1 h,$$

$$|\varphi_n(\tau,\mu) - \varphi(\tau,\mu)| \le \min\left\{c_1 h, c_2 \frac{h^2}{|\mu|}\right\} \quad (0 \le \tau \le H,\ -1 \le \mu \le 1).$$

Автор выражает глубокую благодарность профессору Г.М. Вайникко за неоценимую помощь и внимание к работе.

Литература

1. Вайникко Г., Карпенко Л., Шильман А., Решение интегральных уравнений с экспоненциальными ядрами. Изв. АН Эст ССР. Физ. мат., 1976, 25, № 2, 118-123.
2. Вайникко Г.М., Маршак А.Л., О быстроте сходимости метода дискретных ординат в задаче переноса излучения. Изв. ВУЗ. Математика, 1978, № 11, 11-22.
3. Сеге Г., Ортогональные многочлены. Москва, 1962.
4. Соболев В.В., Перенос лучистой энергии в атмосферах звезд и планет. Москва, 1956.
5. Тиман А.Ф., Теория приближения функций действительного переменного. Москва, 1960.
6. Чандрасекар С., Перенос лучистой энергии. Москва, 1953.
7. Янке Е., Эмде Ф., Леш Ф., Специальные функции (формулы, графики, таблицы). Москва, 1977.
8. Anselon, P.M., Collectively compact operator approximation theory. Prentice-Hall, New Jersey, 1971.
9. Viik, T., On the solution of the equation for the source function of the albedo problem. Публ. Тартуской астрофиз. обс., Тарту, т. 44, 1976, 72-90.

Поступило
15 III 1979

ON THE CONVERGENCE SPEED OF THE DISCRETE ORDINATES METHOD IN THE CASE OF GAUSSIAN QUADRATURE

A. Marshak

Summary

In the present paper the convergence speed of the discrete ordinates method for the radiation transfer in the case of Gaussian quadrature formula normalized to the interval $[0,1]$ is considered. By the method elaborated in [2] uniform estimations (31) are derived. These estimations agree with the numerical results obtained in [2].

Уч.зап.Тартуск.ун-та,1979,500,105-118

ОБ ОПРЕДЕЛЕНИИ ВЕРШИН ТРАНСПОРТНЫХ МНОГОГРАННИКОВ

К. Рийвес
Лаборатория математического программирования
ЭО ЦЭМИ АН СССР

В работах [2-5, 7-10, 12] изучаются транспортные многогранники, т.е. такие выпуклые многогранники, которые задаются ограничениями транспортной задачи линейного программирования. В них получены оценки для минимального и максимального числа вершин и описаны некоторые специальные классы этих многогранников. Настоящая работа посвящена исследованию произвольных транспортных многогранников. Именно, в ней предлагается методика нахождения полного множества вершин любого транспортного многогранника, исходя из системы ограничений, задающих его, причем указывается способ вычисления координат его вершин. С этой целью в п. 2 даны аналитические признаки вершины транспортного многогранника (каждая вершина определяется подмножеством множества ограничений, задающих многогранник). Эти признаки вытекают из аналитических признаков вершины общего выпуклого многогранника [11] с учетом специфики исследуемого класса многогранников. В п. 3 указывается более узкий класс подмножеств множества ограничений, для которых нужно проверять признаки вершины. Дана оценка числа подмножеств из этого класса. Перебрав все подмножества из данного класса и проверив для них выполнение признака вершины, мы получим все вершины транспортного многогранника.

1. <u>Вводные замечания</u>. Назовем выпуклым многогранником $\mathcal{U}$ в N-мерном евклидовом пространстве $\mathcal{R}_N$ пересечение конечного числа замкнутых полупространств H_i $(i=1,\ldots,M)$ пространства $\mathcal{R}_N$. Заметим, что в литературе такое пересечение часто называется многогранным множеством. Если в пространстве $\mathcal{R}_N$ фиксирована некоторая система координат, то любое замкнутое полупространство H_i задается в $\mathcal{R}_N$ некоторым линейным неравенством. Поэтому выпуклый многогранник $\mathcal{U}$ в про-

14

странстве $\mathcal{R}_N$ задается конечной системой линейных неравенств

$$\sum_{\nu=1}^{N} a_{i\nu} x^{\nu} \leq a_{i0}, \ i=1,\dots,M, \tag{1}$$

о которой предполагается, что ни один из векторов $A_i=(a_{i1},\dots,a_{iN})$ не является нулевым. Пусть $\rho = rang\|a_{i\nu}\|$. Тогда $1 \leq \rho \leq min(M,N)$

Пусть во множестве индексов $\{1,\dots,M\}$ системы (1) фиксировано подмножество $\{i(1),\dots,i(N-\sigma)\}$ так, чтобы ранг матрицы коэффициентов системы уравнений

$$\sum_{\nu=1}^{N} a_{i\nu} x^{\nu} = a_{i0}, \ i \in \{i(1),\dots,i(N-\sigma)\}, \tag{2}$$

равнялся $N-\sigma$. Каждое из этих уравнений определяет граничную гиперплоскость Γ_i замкнутого полупространства H_i $(i=i(1),\dots,i(N-\sigma))$ и, в силу сделанного предположения, эти Γ_i пересекаются по σ-мерной плоскости $\mathcal{R}_\sigma \subset \mathcal{R}_N$ (см. [1] , стр.360). Непустое пересечение многогранника $\mathcal{U}$ с этой плоскостью $\mathcal{R}_\sigma$ называется σ-мерной гранью многогранника $\mathcal{U}$. Грани размерности $\sigma=0$ называются вершинами многогранника $\mathcal{U}$. Так как $1 \leq N-\sigma \leq \rho$, то $\sigma=0$ возможно только при $\rho=N$. Предположим, что последнее равенство выполняется. Тогда множество вершин непустого многогранника $\mathcal{U}$ - непустое. Определим понятия, с помощью которых задаются аналитические признаки вершины многогранника (1) при предположении $\rho=N$ (см. [11] , стр. 199).

Рассмотрим векторы $A_i=(a_{i1},\dots,a_{iN})$ и $'A_i=(a_{i1},\dots,a_{iN},a_{i0})$ - строчные векторы матрицы $\mathcal{O}\!\mathcal{L}=\|a_{i\nu}\|$ и расширенной матрицы $'\mathcal{O}\!\mathcal{L}=\|a_{i\nu}a_{i0}\|$ коэффициентов системы (1), соответственно. Пусть во множестве индексов $\{1,\dots,M\}$ неравенств этой системы произвольным образом фиксированы непустые неупорядоченные подмножества $\mathcal{J}=\{i(1),\dots,i(N)\}$ и $\mathcal{J}(i)=\{i\}\cup\mathcal{J}=\{i, i(1),\dots,i(N)\}$. Обозначим через $d_\mathcal{J}$ определитель $(N\times N)$-матрицы, составленной из строчных векторов $A_{i(1)},\dots,A_{i(N)}$ и через $\mathcal{D}_{\mathcal{J}(i)}$, соответственно, определитель матрицы, составленной из строчных векторов $'A_i, 'A_{i(1)},\dots,'A_{i(N)}$.

<u>Определение 1</u>. Неупорядоченная последовательность $(i(1),\dots,i(N))$ называется p-вырожденной относительно множества $\{1,\dots,M\}$, если $d_\mathcal{J}\neq 0$ и во множестве $\mathcal{J}'=\{1,\dots,M\}\setminus\mathcal{J}$ существует точно p индексов i таких, что $\mathcal{D}_{\mathcal{J}(i)}=0$. Число p, которое может принимать значения $0,\dots,M-N$, называется порядком вырожденности последовательности $(i(1),\dots,i(N))$ и обозначается через $p_\mathcal{J}$. Если $M=N$, то для единственной соответствующей последовательности $(i(1),\dots,i(N))$ считается $p_\mathcal{J}=0$.

Пусть последовательность $(i(1),\ldots,i(N))$ фиксирована так, что $d_{\mathcal{J}}\neq 0$ и пусть

$$\mathcal{P}_{\mathcal{J}}=\{i\in\{1,\ldots,M\}\mid \mathcal{D}_{\mathcal{J}(i)}=0\}. \qquad (3)$$

Очевидно, $\mathcal{P}_{\mathcal{J}}\supseteq\mathcal{J}$ и $|\mathcal{P}_{\mathcal{J}}|=N+p_{\mathcal{J}}$.

Определение 2. Пусть неупорядоченная последовательность $(i(1),\ldots,i(N))$ является p-вырожденной относительно множества $\{1,\ldots,M\}$. Такая последовательность $(i(1),\ldots,i(N))$ называется q-допустимой относительно множества $\{1,\ldots,M\}$, если во множестве $\mathcal{J}''=\{1,\ldots,M\}\setminus\mathcal{P}_{\mathcal{J}}$ существует точно q индексов i таких, что

$$\operatorname{sgn}\mathcal{D}_{\mathcal{J}(i)}=(-1)^{N}\operatorname{sgn}d_{\mathcal{J}}. \qquad (4)$$

Число q, которое может принимать значения $0,\ldots,M-N-p$ называется порядком допустимости последовательности $(i(1),\ldots,i(N))$ и обозначается через $q_{\mathcal{J}}$. При $p_{\mathcal{J}}=M-N$ считается $q_{\mathcal{J}}=0$.

Пусть для фиксированной последовательности $(i(1),\ldots,i(N))$ справедливо $d_{\mathcal{J}}\neq 0$. Тогда для соответствующей системы (2) имеет место $\sigma=0$ и она определяет однозначно точку $\mathfrak{X}_{\mathcal{J}}$ пересечения гиперплоскостей $\Gamma_{i(1)},\ldots,\Gamma_{i(N)}$. Координаты этой точки можно вычислить по формуле

$$x^{\nu}_{\mathcal{J}}=(-1)^{N-\nu}({}^{\nu}\mathcal{D}_{\mathcal{J}}/d_{\mathcal{J}}),\ \nu=1,\ldots,N, \qquad (5)$$

где ${}^{\nu}\mathcal{D}_{\mathcal{J}}$ обозначает определитель $(N\times N)$-матрицы, составленной из строчных векторов ${}'A_{i(1)},\ldots,{}'A_{i(N)}$ с опущенным ν-ым столбцом. Тогда при любом $i=1,\ldots,M$ имеет место

$$\sum_{\nu=1}^{N}a_{i\nu}x^{\nu}_{\mathcal{J}}-a_{i0}=(-1)^{N-1}\mathcal{D}_{\mathcal{J}(i)}/d_{\mathcal{J}}. \qquad (6)$$

Если $\mathfrak{X}_{\mathcal{J}}\in\mathcal{U}$, то из системы (1) следует

$$(-1)^{N-1}\mathcal{D}_{\mathcal{J}(i)}/d_{\mathcal{J}}\leq 0,\ i=1,\ldots,M. \qquad (7)$$

Из определения вершин и из геометрического истолкования определений 1 и 2 (см. [11], стр. 196–198), а также из условий (7) вытекает следующее утверждение.

Предложение 1. Точка $\mathfrak{X}_{\mathcal{J}}$ с координатами (5) будет вершиной многогранника $\mathcal{U}$, заданного системой (1) при $\rho=N$ тогда и только тогда, когда порядок вырожденности $p_{\mathcal{J}}$ и порядок допустимости $q_{\mathcal{J}}$ последовательности $(i(1),\ldots,i(N))$ удовлетворяют условию

$$p_\jmath + q_\jmath = M - N. \tag{8}$$

Предложение 2. Пусть различные последовательности $(i(1),\ldots,i(N))$, $(j(1),\ldots,j(N))$ определяют вершины $\mathfrak{X}_\jmath$, $\mathfrak{X}_\gamma$ для многогранника $\mathcal{U}$. Если при этом $\mathcal{P}_\jmath = \mathcal{P}_\gamma$, то $\mathfrak{X}_\jmath = \mathfrak{X}_\gamma$.

Это утверждение следует по определению вершины из предложений 1.3 и 1.5 работы [11].

Перейдем к исследованию транспортных многогранников, модифицируя сперва подходящим образом условия (8). Этим мы получим аналитические признаки вершин для рассматриваемого специального класса многогранников.

2. Аналитические признаки вершины транспортного многогранника. Пусть векторы $A=(a_1,\ldots,a_m)$, $B=(b_1,\ldots,b_n)$ со строго положительными координатами $(a_k, b_\ell > 0,\ k=1,\ldots,m;\ \ell=1,\ldots,n)$ даны так, что выполняется равенство

$$\sum_{k=1}^{m} a_k = \sum_{\ell=1}^{n} b_\ell . \tag{9}$$

Тогда система линейных неравенств

$$\begin{gathered} -x_{k\ell} \le 0, \\ \sum_{\ell=1}^{n} x_{k\ell} \le a_k,\ k=1,\ldots,m, \\ \sum_{k=1}^{m} x_{k\ell} \le b_\ell,\ \ell=1,\ldots,n, \\ -\sum_{\ell=1}^{n} x_{k\ell} \le -a_k,\ k=1,\ldots,m, \\ -\sum_{k=1}^{m} x_{k\ell} \le -b_\ell,\ \ell=1,\ldots,n, \end{gathered} \tag{10}$$

задает непустой транспортный многогранник $M(A,B)$ порядка $m \times n$. Переобозначим переменные $x_{k\ell}$, записанные с помощью двух нижних индексов $k=1,\ldots,m;\ \ell=1,\ldots,n$, используя один верхний индекс $\nu=1,\ldots,m\cdot n$ по формуле

$$x^\nu \equiv x_{k\ell}, \text{ где } \nu=(k-1)n+\ell. \tag{11}$$

Тогда очевидно, что в данном случае $M = mn + 2(m+n)$, $N = m\cdot n$ и $M > N$. Так как матрица $\mathfrak{A} = \|a_{i\nu}\|$ $(i=1,\ldots,M;\ \nu=1,\ldots,N)$ коэффициентов системы (10) содержит матрицу $-E$, где E — единичная матрица порядка N, то $\rho = \operatorname{rang} \mathfrak{A} = N$. Следовательно, по п.1 множество вершин многогранника $M(A,B)$ - непустое. По предложению 1 любая вершина для $M(A,B)$ определяется некоторой N-элементной последовательностью индексов $(i(1),\ldots,i(N))$ при $J=\{i(1),\ldots,i(N)\}\subset\{1,\ldots,mn+2(m+n)\}$, удовлетворяющей условию (8). Благодаря специфике исследуемых многогранников при определении полного множества их вершин не надо проверять условия (8) для всевозможных C_M^N последовательностей, а число проверяемых последовательностей будет гораздо меньше. Для установления этого заметим, что строки расширенной матрицы $'\mathfrak{A}$ коэффициентов системы (10) линейно зависимы. Если обозначить, как и раньше, i-ый строчный вектор матрицы $'\mathfrak{A}$ через $'A_i$, можно проверить, что кроме векторов $'A_1,\ldots,'A_{mn}$ и векторы $'A_{mn+1},\ldots,'A_{mn+m+n-1}$ между собой линейно независимы. При этом справедливы следующие соотношения

$$'A_{mn+(m+k)+n}=-'A_{mn+k},\ k=1,\ldots,m,$$
$$'A_{mn+2m+(n+\ell)}=-'A_{mn+m+\ell},\ \ell=1,\ldots,n; \tag{12A}$$
$$'A_{mn+m+n}=\sum_{k=1}^{m}{'A_{mn+k}}-\sum_{\ell=1}^{n-1}{'A_{mn+n+\ell}}. \tag{12Б}$$

В дальнейшем нам удобно пользоваться следующим разбиением множества $\{1,\ldots,M\}=\{1,\ldots,mn+2(m+n)\}$ строчных индексов матрицы $'\mathfrak{A}$:

$$\begin{aligned}
&\mathcal{N}(0)=\{i\mid i=1,\ldots,mn\}\equiv\{\nu\mid \nu=1,\ldots,mn\},\\
&\mathcal{N}(1)=\{i=mn+k\mid k=1,\ldots,m\},\\
&\mathcal{N}(2)=\{i=mn+m+\ell\mid \ell=1,\ldots,n\},\\
&\mathcal{N}(3)=\{i=mn+(m+k)+n\mid k=1,\ldots,m\},\\
&\mathcal{N}(4)=\{i=mn+2m+(n+\ell)\mid \ell=1,\ldots,n\}.
\end{aligned} \tag{13}$$

Из условий (12А) получается следующее утверждение, доказательство которого опускается для краткости изложения.

<u>Предложение 3</u>. При нахождении всех вершин транспортного многогранника $M(A,B)$ порядка $m\times n$ можно ограничиться рассмотрением лишь таких последовательностей $(i(1),\ldots,i(N))$, элементы которых принадлежат множеству $\mathcal{N}(0,1,2)=\mathcal{N}(0)\cup\mathcal{N}(1)\cup\mathcal{N}(2)$.

Будем в дальнейшем рассматривать только последовательности $(i(1),\dots,i(N))$ с $\{i(1),\dots,i(N)\}\subset\mathcal{H}(0,1,2)$ не оговаривая это специально. Введем для любой фиксированной $(i(1),\dots$ $\dots,i(N))$ обозначения

$$\mathcal{J}=\{i(1),\dots,i(N)\},\ \mathcal{J}(0)=\mathcal{H}(0)\cap\mathcal{J},\ \mathcal{J}(1,2)=\mathcal{H}(1,2)\cap\mathcal{J},$$

$$(\mathcal{H}(1,2)=\mathcal{H}(1)\cup\mathcal{H}(2)),\ \mathcal{J}'(0)=\mathcal{H}(0)\setminus\mathcal{J}(0),\ \mathcal{J}'(1,2)=\mathcal{H}(1,2)\setminus\mathcal{J}(1,2),$$

$$\mathcal{J}'(0;0)=\{i\in\mathcal{J}'(0)\mid \mathcal{D}_{\mathcal{J}(i)}=0\},\ \mathcal{J}'(0,1)=\mathcal{J}'(0)\setminus\mathcal{J}'(0;0) \quad (14)$$

и

$$m(0)=|\mathcal{J}(0)|,\ m(1,2)=|\mathcal{J}(1,2)|,\ m'(0)=|\mathcal{J}'(0)|,\ m'(1,2)=$$

$$=|\mathcal{J}'(1,2)|,\ m'(0;0)=|\mathcal{J}'(0;0)|,\ m'(0;1)=|\mathcal{J}'(0,1)|. \quad (15)$$

Кроме того, пусть для фиксированных i и $\mathcal{J}$ при $\mathcal{J}(i)=\{i\}\cup\mathcal{J}$ определены множества $\mathcal{J}(i;0)=\{i\}\cup\mathcal{J}(0)$, если $i\in\mathcal{H}(0)$ и $\mathcal{J}(i;1,2)=\{i\}\cup\mathcal{J}(1,2)$, если $i\in\mathcal{H}(1,2)$. Упорядочим элементы

последовательности $(i(1),\dots,i(N))$ так, чтобы ввиду $m(0)+$ $+m(1,2)=mn=N$, имело место

$$\mathcal{J}(0)=\{i(1),\dots,i(m(0))\},\ \mathcal{J}(1,2)=\{i(m(0)+1),\dots,i(N)\}$$

и кроме того, выполнялись неравенства

$$i(1)<\dots<i(m(0)). \quad (16)$$

Так как по определению $m(1,2)\le m+n<N$, то $m(0)=N-m(1,2)\ne 0$ и $\mathcal{J}(0)\ne\varnothing$.

Из первой группы неравенств системы (10) в обозначениях (11) видно, что у строчных векторов $'A_{i(1)},\dots,'A_{i(m(0))}$ матрицы $'\mathcal{O}\mathcal{C}$ имеется единственная отличная от нуля координата -1. Она будет, соответственно $i(1)$-ой, $\dots$, $i(m(0))$-ой. Поэтому легко разложить определитель $d_{\mathcal{J}}$ порядка N по первым $m(0)$ строкам, определители $\mathcal{D}_{\mathcal{J}(i)}(i\in\mathcal{H}(0))$ порядка $N+1$ по первым $m(0)+1$ строкам, а определители $\mathcal{D}_{\mathcal{J}(i)}(i\in\mathcal{H}(1,2))$ порядка $N+1$ по строкам с номерами $2,\dots,m(0)+1$. Для этого предположим, что $\mathcal{V}=\{v_1,\dots,v_s\}\subset\mathcal{H}(0)$ и $\mathcal{W}=\{w_1,\dots,$ $w_t\}\subset\mathcal{H}(1,2)$. Обозначим через $^{\mathcal{V}}d_{\mathcal{W}}$ определитель матрицы, составленной из w_1-ой, $\dots$, w_t-ой строк матрицы $\mathcal{O}\mathcal{C}$ с опущенными v_1-ым, $\dots$, v_s-ым столбцами и через $^{\mathcal{V}}\mathcal{D}_{\mathcal{W}}$ – определитель матрицы, составленной таким же образом по матрице $'\mathcal{O}\mathcal{C}$.

Введем для фиксированной последовательности $(i(1),\dots,i(N))$

с определенным по (15) значением $m(o)$ функцию

$$h(m(o))=\begin{cases}0, & \text{если } m(o)(m(o)+2) \text{ — четное число},\\ 1, & \text{в противном случае},\end{cases} \tag{17А}$$

и пусть при каждом $i\in \mathcal{N}(o)$

$$f(i)=\begin{cases}0, & \text{если } \left|\{i(\tau)\in \mathfrak{J}(o)\mid i(\tau)<i\}\right| \text{ четное число},\\ 1, & \text{в противном случае}.\end{cases} \tag{17Б}$$

Тогда имеют место равенства

$$\begin{aligned} &sgn\, d_{\mathfrak{J}} = (-1)^{x}\, sgn\ {}^{\mathfrak{J}(o)}d_{\mathfrak{J}(1,2)}\,,\\ &sgn\, \mathcal{D}_{\mathfrak{J}(i)} = (-1)^{y}\, sgn\ {}^{\mathfrak{J}(i;o)}\mathcal{D}_{\mathfrak{J}(1,2)}\,,\ i\in \mathcal{N}(o),\\ &sgn\, \mathcal{D}_{\mathfrak{J}(i)} = (-1)^{z}\, sgn\ {}^{\mathfrak{J}(o)}\mathcal{D}_{\mathfrak{J}(i;1,2)}\,,\ i\in \mathcal{N}(1,2),\end{aligned} \tag{18А}$$

где

$$\begin{aligned} &x = m(o)+0{,}5\,m(o)(m(o)+1)+\sum_{\tau=1}^{m(o)} i(\tau),\\ &y \sim 0.5\,m(o)(m(o)+1)+\sum_{\tau=1}^{m(o)} i(\tau)+i+f(i),\\ &z \sim 0.5\,m(o)(m(o)+1)+\sum_{\tau=1}^{m(o)} i(\tau).\end{aligned} \tag{18Б}$$

Здесь знак $\sim$ между целыми числами означает, что эти числа имеют одинаковую четность. Из соотношений (18) следует также, что соответствующие определители обращаются одновременно в нуль.

<u>Предложение 4</u>. Пусть последовательность $(i(1),\dots,i(N))$ с $\mathfrak{J}=\mathfrak{J}(o)\cup\mathfrak{J}(1,2)$ фиксирована так, что $d_{\mathfrak{J}}\neq 0$ и $i(1)<\dots<i(m(o))$. Определенная однозначно этой последовательностью точка $\mathfrak{X}_{\mathfrak{J}}$ пересечения гиперплоскостей $\Gamma_{i(1)},\dots,\Gamma_{i(N)}$ будет вершиной транспортного многогранника $M(A,B)$, заданного системой (10), тогда и только тогда, когда при всех $i\in \mathfrak{J}'(o;1)$ имеет место

$$sgn\ {}^{\mathfrak{J}(i;o)}\mathcal{D}_{\mathfrak{J}(1,2)} = (-1)^{N(i)}\, sgn\ {}^{\mathfrak{J}(o)}d_{\mathfrak{J}(1,2)}, \tag{19А}$$

где

$$N(i) = N+i+f(i)+h(m(o)), \tag{19Б}$$

и для всех $i\in \mathfrak{J}'(1,2)$ справедливо

$${}^{\mathfrak{J}(o)}\mathcal{D}_{\mathfrak{J}(i;1,2)} = 0. \tag{20}$$

Доказательство. Достаточность. Покажем, что при условиях (19) и (20) выполняются условия (8) предложения 1, необходимые и достаточные для того, чтобы точка $\mathfrak{X}_{\mathfrak{I}}$ являлась вершиной многогранника $M(A,B)$. Для этого вычислим порядок вырожденности $p_{\mathfrak{I}}$ и порядок допустимости $q_{\mathfrak{I}}$ рассматриваемой последовательности $(i(1),\dots,i(N))$, если условия (19) и (20) выполнены. По определениям 1 и 2 нам надо знать значения всех определителей $\mathcal{D}_{\mathfrak{I}(i)}$ при всевозможных $i\in\mathfrak{I}(0)\cup\mathfrak{I}(1,2)\cup\mathfrak{I}'(0;0)\cup\mathfrak{I}'(0,1)\cup\mathfrak{I}'(1,2)\cup\mathcal{H}(3)\cup\mathcal{H}(4)$.

Если $i\in\mathfrak{I}(0)\cup\mathfrak{I}(1,2)$, то определитель $\mathcal{D}_{\mathfrak{I}(i)}$ имеет две одинаковые строки и поэтому равен нулю. Если $i\in\mathfrak{I}'(0;0)$, то по формуле (14), определяющей множество $\mathfrak{I}'(0;0)$, также $\mathcal{D}_{\mathfrak{I}(i)}=0$. В случае $i\in\mathfrak{I}'(1,2)$ по условию (20) справедливо ${}^{\mathfrak{I}(0)}\mathcal{D}_{\mathfrak{I}(i;1,2)}=0$ и одновременно с ним обращается в нуль также соответствующий определитель $\mathcal{D}_{\mathfrak{I}(i)}$. Если же $i\in\mathcal{H}(3)\cup\mathcal{H}(4)$, то $i-(m+n)\in\mathcal{H}(1,2)=\mathfrak{I}(1,2)\cup\mathfrak{I}'(1,2)$, и поэтому $\mathcal{D}_{\mathfrak{I}(i-(m+n))}=0$. В силу (12А), справедливо ${}'A_{i-(m+n)}=-{}'A_i$, а тогда $\mathcal{D}_{\mathfrak{I}(i)}=0$. По определению множества $\mathfrak{I}'(0;1)$ (см. формулу (14)), имеем $\mathcal{D}_{\mathfrak{I}(i)}\neq0$ при $i\in\mathfrak{I}'(0;1)$, следовательно,

$$p_{\mathfrak{I}}=M-(N+m'(0;1)).$$

Чтобы определить значение $q_{\mathfrak{I}}$, заметим, что в силу (18А) и (19А) при всех $i\in\mathfrak{I}'(0;1)$ справедлива цепочка равенств

$$\operatorname{sgn}\mathcal{D}_{\mathfrak{I}(i)}=(-1)^{y}\operatorname{sgn}{}^{\mathfrak{I}(i;0)}\mathcal{D}_{\mathfrak{I}(1,2)}=(-1)^{y+N(i)}\operatorname{sgn}{}^{\mathfrak{I}(0)}d_{\mathfrak{I}(1,2)}=$$
$$=(-1)^{y+N(i)+x}\operatorname{sgn}d_{\mathfrak{I}},$$

где с учетом (18Б), (19Б) и (17А) имеет место

$$y+N(i)+x=0.5\,m(0)(m(0)+1)+\sum_{\tau=1}^{m(0)}i(\tau)+i+f(i)+N+$$
$$+i+f(i)+h(m(0))+m(0)+0.5\,m(0)(m(0)+1)+\sum_{\tau=1}^{m(0)}i(\tau)\sim N,$$

т.е. $\operatorname{sgn}\mathcal{D}_{\mathfrak{I}(i)}=(-1)^{N}\operatorname{sgn}d_{\mathfrak{I}}$. По условию (4) определения 2 последние равенства являются условиями допустимости для всех индексов $i\in\mathfrak{I}'(0;1)$. Поэтому

$$q_{\mathfrak{I}}=m'(0;1)$$

и, таким образом,

$$p_{\mathfrak{I}}+q_{\mathfrak{I}}=M-N,$$

что и будет аналитическим признаком вершины из предложения

Необходимость доказывается рассуждениями, проведенными в обратном порядке. Для краткости изложения мы их опустим, и закончим на этом доказательство.

Следует отметить, что для каждой фиксированной последовательности $(i(1),\ldots,i(N))$ при проверке аналитических признаков (19) и (20) вершины транспортного многогранника нужно кроме ${}^{\mathfrak{J}(0)}d_{\mathfrak{J}(1,2)}$ вычислить не более $m'(0)=N-m(0)<N$ определителей ${}^{\mathfrak{J}(i,0)}\mathcal{D}_{\mathfrak{J}(1,2)}$ порядка $m(1,2)=N-m(0)=m'(0)$ и не более $m'(1,2)=m+n-m(1,2)=m+n-m'(0)$ определителей ${}^{\mathfrak{J}(0)}\mathcal{D}_{\mathfrak{J}(i;1,2)}$ порядка $m(1,2)+1<N+1$ (здесь учтены определения (14)). Так как $m'(0)+m'(1,2)=m+n$, то число вычисляемых определителей не превышает $m+n+1$.

3. Последовательности, определяющие вершины транспортных многогранников. Как уже отмечалось, число последовательностей, перебираемых в ходе определения полного множества вершин транспортного многогранника $M(A,B)$, сильно сокращается в силу некоторых общих соображений. Мы опишем класс последовательностей, который определяет все вершины многогранника $M(A,B)$, и укажем также некоторые подклассы таких последовательностей, которым никогда не соответствуют вершины многогранника $M(A,B)$. Основные результаты формулируются в виде предложений, подробные доказательства которых опускаются или описываются лишь в общих чертах.

Известно (см. [6], стр. 334–335), что матрица транспортной задачи унимодулярна, т.е. если $d_{\mathfrak{J}}\neq 0$, то $|d_{\mathfrak{J}}|=1$. Рассмотрим последовательность $(i(1),\ldots,i(N))$ с $|d_{\mathfrak{J}}|=1$ такую, что $\mathfrak{J}=\mathfrak{J}(0)\cup\mathfrak{J}(1,2)$ при условии (16). Тогда в силу применяемых обозначений имеют место равенства

$$|{}^{i}\mathcal{D}_{\mathfrak{J}}|=|{}^{\mathfrak{J}(i;0)}\mathcal{D}_{\mathfrak{J}(1,2)}|,\quad i\in\mathcal{H}(0).$$

Поэтому из соотношений (5) получается следующее утверждение.

Предложение 5. Если для последовательности $(i(1),\ldots,i(N))$ выполняются условия (19) и (20), то координаты $x^i_{\mathfrak{J}}\geq 0$ ($i\in\mathcal{H}(0)$) соответствующей вершины $\mathfrak{X}_{\mathfrak{J}}$ многогранника $M(A,B)$ вычисляются по формуле

$$x^i_{\mathfrak{J}}=\begin{cases}|{}^{\mathfrak{J}(i;0)}\mathcal{D}_{\mathfrak{J}(1,2)}|, & \text{если } i\in\mathfrak{J}'(0;1),\\ 0, & \text{если } i\in\mathfrak{J}(0)\cup\mathfrak{J}'(0;0).\end{cases}$$

15

Отсюда следует, что координаты x^i_j вершины транспортного многогранника $M(A,B)$ отличны от нуля только при $i \in \mathfrak{I}'(0;1)$. Если учесть обозначения (11), (15), из системы (10) и соотношения (12Б), очевидно, вытекает следующий результат.

<u>Предложение 6</u>. Фиксированная последовательность $(i(1),\dots,i(N))$ может определять вершину для $M(A,B)$ только тогда, когда имеют место неравенства

$$\max(m,n) \leq m'(0;1) \leq m'(0) \leq m+n-1.$$

<u>Определение 3</u>. Вершина $\mathfrak{X}_\mathfrak{I}$ транспортного многогранника $M(A,B)$ порядка $m \times n$, определенная последовательностью $(i(1),\dots,i(N))$, называется невырожденной, если $m'(0;1) = m+n-1$ и вырожденной в противном случае [2].

Пусть последовательность $(i(1),\dots,i(N))$ определяет некоторую вершину для многогранника $M(A,B)$. Рассмотрим последовательность $(j(1),\dots,j(N))$ такую, что $\mathcal{J} = \{j(1),\dots,j(N)\} \neq \mathfrak{I}$ и определим аналогично формулам (14), (15) множества $\mathcal{J}(0) = \mathfrak{N}(0) \cap \mathcal{J}$, $\mathcal{J}(1,2) = \mathfrak{N}(1,2) \cap \mathcal{J}$, и обозначим $|\mathcal{J}(0)| = \ell(0)$, $|\mathcal{J}(1,2)| = \ell(1,2)$ и т.д.

<u>Предложение 7</u>. Любая вершина $\mathfrak{X}_\mathfrak{I}$ транспортного многогранника $M(A,B)$ порядка $m \times n$ определяется кроме $(i(1),\dots,i(N))$ также некоторой такой последовательностью $(j(1),\dots,j(N))$, для которой $\mathcal{J}(0) \subseteq \mathfrak{I}(0)$ и $\mathcal{J}(1,2) = \mathfrak{N}(1,2) \setminus \{mn+m+n\}$ (в частности, она может совпасть с $(i(1),\dots,i(N))$. Если при этом $\ell'(0;0) = 0$, то вершина $\mathfrak{X}_\mathfrak{I}$ является невырожденной; если же $\ell'(0;0) \neq 0$, то она – вырождена.

Для <u>доказательства</u> предложения предположим, что последовательность $(i(1),\dots,i(N))$ с $\mathfrak{I} = \mathfrak{I}(0) \cup \mathfrak{I}(1,2)$ удовлетворяет условиям предложения 4, т.е. определяет вершину $\mathfrak{X}_\mathfrak{I}$. Если при этом $\mathfrak{I}(1,2) = \mathfrak{N}(1,2) \setminus \{mn+m+n\}$, то рассматриваемая последовательность $(i(1),\dots,i(N))$ уже имеет описанную в предложении структуру. В противном случае следует показать, что последовательности $(i(1),\dots,i(N))$ можно всегда сопоставить отличную от нее последовательность $(j(1),\dots,j(N))$, имеющую описанную в формулировке предложения структуру и определяющую ту же вершину. Для этого сперва фиксируем $\mathcal{J}^*(1,2) = \{j(\ell(0)+1),\dots,j(N)\} \supseteq \mathfrak{I}(1,2)$ так, что $\ell^*(1,2) = m+n-1$. Оказывается, что во множестве $\mathfrak{I}(0)$ всегда найдется по крайней мере одно $\ell(0)$-элементное ($\ell(0) = mn+1-m-n$) подмножество $\mathcal{J}(0) = \{j(1),\dots,j(\ell(0))\}$ так, чтобы для соответствующей матрицы $d_{\mathcal{J}^*} \neq 0$. Если $mn+m+$ $+n \in \mathcal{J}^*(1,2)$, то с учетом (12Б) можно перейти от $\mathcal{J}^*(1,2)$ к $\mathcal{J}(1,2) = \mathcal{N}(1,2) \setminus \{mn+m+n\}$ и при этом также $d_{\mathcal{J}} = 0$. Для построенной таким образом последовательности $(j(1), \dots, j(N))$ получается равенство $\mathcal{P}_{\mathcal{J}} = \mathcal{P}_{J}$, которое по предложению 2 означает совпадение точек $\mathfrak{X}_J$ и $\mathfrak{X}_{\mathcal{J}}$. Тем самым $\mathfrak{X}_{\mathcal{J}}$ является вершиной для $M(A,B)$, которая определяется уже последовательностью специальной структуры.

Второе утверждение предложения следует из соотношения $\ell'(0;1) = \ell(1,2) - \ell'(0;0) = m+n-1 - \ell'(0;0)$ и предложения 5 с учетом определения 3.

Предложение 7 дает описание специального класса последовательностей, определяющих полное множество вершин транспортных многогранников. Все последовательности этого класса содержат однозначно определенное $(m+n-1)$-элементное множество $\mathcal{J}(1,2)$ и поэтому число последовательностей, подлежащих исследованию при нахождении полного множества вершин для $M(A,B)$, зависит только от возможностей выбора множества $\mathcal{J}(0)$. Число последних не превышает $C_N^{\ell(0)} =$ $= C_{mn}^{m+n-1}$. Нижеследующее предложение дает возможность исключить из рассмотрения еще некоторые классы последовательностей со специальными множествами $\mathcal{J}(0)$.

Пусть при фиксированном $\ell \in \{1, \dots, n\}$ через $\mathcal{D}^\ell$ обозначено m-элементное множество $\mathcal{D}^\ell = \{\ell, n+\ell, \dots, (m-1)n+\ell\} \subset \mathcal{N}(0)$ и при фиксированном $k \in \{1, \dots, m\}$ – через $\mathcal{E}^k$, соответственно, n-элементное множество $\mathcal{E}^k = \{(k-1)n+1, (k-1)n+2, \dots, kn\} \subset \mathcal{N}(0)$

<u>Предложение 8</u>. Если последовательность $(i(1), \dots, i(N))$ содержит по крайней мере одно из множеств $\mathcal{D}^\ell$ $(\ell = 1, \dots, n)$ или одно из множеств $\mathcal{E}^k$ $(k = 1, \dots, m)$, то она не определяет вершины для транспортного многогранника $M(A,B)$ порядка $m \times n$.

Для <u>доказательства</u> можно отметить, что по предложению 5 при сделанных предположениях $x^i_J = 0$ для всех $i \in \mathcal{D}^\ell$ (или $\mathcal{E}^k$), которые с учетом обозначений (11) приводят к противоречивым равенствам $b_\ell = 0$ (или $a_k = 0$) в системе (10).

<u>Определение 4</u>. Последовательность $(i(1), \dots, i(N))$ с $J =$ $= \{i(1), \dots, i(N)\}$ называется допустимой, если $J(0) = \mathcal{N}(0) \cap J$ не содержит ни одного из множеств $\mathcal{D}^\ell$ и $\mathcal{E}^k$.

Основные результаты статьи подытожены в следующем утверждении.

<u>Предложение 9</u>. Все вершины транспортного многогранника $M(A,B)$ порядка $m \times n$ определяются такими допустимыми N-элементными ($N = mn$) последовательностями индексов $(i(1), \ldots, i(N))$, что

1. $J = J(0) \cup J(1,2) \subset \mathcal{H}(0,1,2) \setminus \{mn+m+n\}$ и $m(1,2) = n+m-1$
2. при $J(0) = \{i(1), \ldots, i(m(0))\}$, $i(1) < \ldots < i(m(0))$ выполняются условия (19) и (20).

Для любой вершины множество $J(0) \cup J'(0;0)$ определено однозначно. Если $J'(0;0) = \emptyset$, то вершина не вырождена, в противном случае она вырождена.

Если обозначить через P число допустимых последовательностей и принять $S = [m(0)/n]$, $T = [m(0)/m]$, то, в силу $m(0) = N - m(1,2) < N = mn$, справедливы $0 \le S < m$, $0 \le T < n$ и можно показать, что $P \le P'$, где

$$P' = C_N^{m(0)} + \sum_{\tau=1}^{T} (-1)^{\tau} C_n^{\tau} C_{N-\tau m}^{m(0)-\tau m} + \sum_{\sigma=1}^{S} (-1)^{\sigma} C_m^{\sigma} C_{N-\sigma n}^{m(0)-\sigma n}.$$

По предложению 9 для определения полного множества вершин транспортного многогранника $M(A,B)$ надо перебрать не более P' последовательностей $(i(1), \ldots, i(N))$. Эта оценка является в общем случае более грубой, чем

$$\bar{M}(A,B) \le m^{n-2} n^{m-1}, \quad 2 \le m \le n, \qquad (21)$$

полученная Лихачевым [9]. Например, в случае $m=2$, $n=5$ имеет место $m^{n-2} n^{m-1} = 40$, $P' = 80$. Поэтому при нахождении полного множества вершин транспортного многогранника $M(A,B)$ подходяще упорядоченным перебором всевозможных P допустимых последовательностей $J(0)$ целесообразно пользоваться оценкой (21) с тем, чтобы кончить перебор сразу после нахождения максимального числа вершин. Если число вершин меньше максимального, то надо провести полный перебор.

Литература

1. А л е к с а н д р о в П.С., Лекции по аналитической геометрии. Москва, 1968.
2. Е м е л и ч е в В.А., К о н о н е н к о А.М., Об одном классе транспортных многогранников. Изв. АН БССР. Сер. физ.-мат. н., 1971, № 3, 21-25.

3. Емеличев В.А., Кравцов М.К., К перечислительным задачам на транспортных многогранниках. Изв. АН БССР. Сер. физ.-мат. н., 1976, №1, 9-13.

4. Емеличев В.А., Кравцов М.К., Авербух Н.Д., О максимальном числе вершин транспортного многогранника. Докл. АН БССР, 1976, 20, № 6, 509-512.

5. Емеличев В.А., Кравцов М.К., Крачковский А.П., К гипотезе о максимальном числе вершин транспортного многогранника. Изв. АН БССР. Сер. физ.-мат. н., 1977, № 6, 114-118.

6. Корбут А.А., Финкельштейн Ю.Ю., Дискретное программирование. Москва, 1969.

7. Кравцов М.К., Перечислительные задачи на вырожденных транспортных многогранниках. (Редколлегия ж. "Изв. АН БССР. Сер. физ.-мат. н.",) Минск, 1976, 21 с., Библиогр. 3 назв. (Рукопись деп. в ВИНИТИ 18 марта 1976 г., № 832-76 Деп).

8. Кравцов М.К., К оценке сверху числа вершин вырожденного транспортного многогранника. Изв. АН БССР. Сер. физ.-мат. н., 1976, № 6, 124.

9. Лихачев В.М., Об одной оценке числа вершин транспортного многогранника. Изв. АН БССР. Сер. физ.-мат. н., 1975, № 2, 99.

10. Лихачев В.М., Емеличев В.А., К оценке сверху числа вершин транспортного многогранника. Изв. АН БССР. Сер. физ.-мат. н., 1974, № 3, 121-123.

11. Рийвес К., Об аффинной классификации и признаках выпуклых многогранников в евклидовом пространстве R_n III. Уч. зап. Тартуск. ун-та, 1975, 366, 187-216.

12. Jemelitschew, W.A., Krawzow, M.K., Transportpolyeder. 21. Int. Wiss. Kolloq. Techn. Hochsch. Ilmenau, 1976, Ht. 3, S 1., s.a., 97-99.

Поступило
4 У 1978

A METHOD OF COMPUTING THE VERTICES OF TRANSPORTPOLYTOPES

K. Riives

Summary

In the paper convex transportpolytopes are considered. A method for determining all vertices of transportpolytopes using their analytic character is described. For the special class of polytopes discussed here the characters follow from the results of [11] (see Proposition 4). Each vertex of a transportpolytope is given by a subset of constraints presenting the polytope. A complete characterization of the subsets determining all vertices of transportpolytopes is presented in Proposition 9.

Уч.зап.Тартуск.ун-та,1979,500,119-129

ВЫРАЗИМОСТЬ В ЭЛЕМЕНТАРНОЙ ТЕОРИИ РЕКУРСИВНЫХ МНОЖЕСТВ С ЛОГИКОЙ РЕАЛИЗУЕМОСТИ

Р. Пранк

Московский государственный университет

§1. Введение

Обозначим через $\mathcal{R}$ булеву алгебру рекурсивных подмножеств множества натуральных чисел N.

Определение 1.1. Называем языком $\mathcal{L}(\mathcal{R})$ язык первого порядка с переменными $\mathfrak{X}_1, \mathfrak{X}_2, \dots$ для рекурсивных множеств, с константными символами Φ и N и функциональными символами $\cup$, $\cap$ и $'$.

Пусть $\{\varphi_x\}$ – геделева нумерация частично-рекурсивных функций (ЧРФ) одной переменной. Будем считать, что функция φ_x вычисляется машиной Тьюринга P_x и проводить доказательства подобно [2] в терминах шагов работы P_x. Через $|M|$ обозначим мощность множества M, а через $\|M\|$ – пару μ/ν, где $\mu = |M|$ и $\nu = |M'|$. Ясно, что всегда $\|M\| \in \{0/\omega, 1/\omega, \dots, n/\omega, \dots, \omega/\omega, \dots, \omega/n, \dots, \omega/1, \omega/0\}$. Через $j_i(x)$ обозначим функцию, значением которой является степень i-того простого числа в разложении x на простые множители.

Если φ_i является характеристической функцией (ХФ) некоторого подмножества множества N, то обозначим это подмножество через $\mathcal{R}_i$, т.е. для таких i имеет место $\mathcal{R}_i = \{x \mid \varphi_i(x) = 1\}$. Для остальных i символы $\mathcal{R}_i$ не вводятся.

Определим понятие реализуемости по Клини для замкнутых формул языка $\mathcal{L}(\mathcal{R})$, расширенного константными символами $\mathcal{R}_i$.

Определение 1.2. 1. Если S и T – постоянные термы расширенного языка, то

$$e \circledR (S = T) \rightleftarrows e = 0 \ \& \ (S = T) \text{ истинно.}$$

2. Для пропозициональных связок повторяется определение Клини ([1]) для арифметических формул.

3. Пусть $A(\mathfrak{X})$ – формула расширенного языка с одной свободной переменной $\mathfrak{X}$. Тогда

1) e ⓡ $\exists \mathfrak{X} A(\mathfrak{X}) \rightleftharpoons e = 2^a \cdot 3^b \,\&\, \varphi_b - X\Phi \,\&\, a$ ⓡ $A(\mathcal{R}_b)$,

2) e ⓡ $\forall \mathfrak{X} A(\mathfrak{X}) \rightleftharpoons (\forall x)[\varphi_x - X\Phi \Rightarrow \varphi_e(x)$ ⓡ $A(\mathcal{R}_x)]$.

Называем формулу A реализуемой и пишем ⓡ A, если существует число, реализующее A. Теорией $\mathcal{R}_r$ называем множество всех реализуемых формул языка $\mathcal{L}(\mathcal{R})$. Ясно, что в конечном итоге формулы с вхождениями символов $\mathcal{R}_i$ особого интереса не представляют. Они описывают не рекурсивно инвариантные понятия, осмысленность символа $\mathcal{R}_i$ зависит от выбора нумерации $\{\varphi_x\}$ и не является эффективной и т. д.

Из определения видно, что реализуемость формулы расширенного языка, полученной из $A(\mathfrak{X}_1, \dots, \mathfrak{X}_n)$ подстановкой $\mathcal{R}_{i_1}, \dots, \mathcal{R}_{i_n}$ вместо переменных, зависит только от выбора множеств $\mathcal{R}_{i_k}$, а не от выбора их характеристических индексов. Точнее, имеет место следующая

Лемма 1.1. Пусть $A(\mathfrak{X}_1, \dots, \mathfrak{X}_n)$ – формула расширенного языка, а $\mathcal{R}_{i_1} = \mathcal{R}_{j_1}, \dots, \mathcal{R}_{i_n} = \mathcal{R}_{j_n}$. Тогда

$$e \text{ ⓡ } A(\mathcal{R}_{i_1}, \dots, \mathcal{R}_{i_n}) \Longleftrightarrow e \text{ ⓡ } A(\mathcal{R}_{j_1}, \dots, \mathcal{R}_{j_n}).$$

В частности это означает, что в теории $\mathcal{R}_r$ формулы выражают свойства множеств, а не свойства их индексов.

Определение 1.3. Называем предикат $P(\mathfrak{X}_1, \dots, \mathfrak{X}_n)$ определенный на рекурсивных множествах, выразимым в теории $\mathcal{R}_r$, если существует такая формула $A(\mathfrak{X}_1, \dots, \mathfrak{X}_n)$ языка $\mathcal{L}(\mathcal{R})$, что

$$\text{ⓡ } A(\mathcal{R}_{i_1}, \dots, \mathcal{R}_{i_n}) \Longleftrightarrow P(\mathcal{R}_{i_1}, \dots, \mathcal{R}_{i_n}) = t.$$

Иногда в дальнейшем пишем в простых случаях в формулах вместо подформул выражаемые ими предикаты, например $\mathfrak{X} \neq \mathfrak{Y}$ вместо формулы $\neg(\mathfrak{X} = \mathfrak{Y})$, $\mathfrak{X} \subseteq \mathfrak{Y}$ вместо формулы $\mathfrak{X} \cap \mathfrak{Y}' = \phi$.

В настоящей статье дается описание выразимых в теории $\mathcal{R}_r$ предикатов. Для сравнения приводим сначала описание выразимых в классической элементарной теории $\mathcal{R}$ предикатов.

§ 2. Выразимость в классической элементарной теории $\mathcal{R}$.

Рассмотрим сначала формулы с одним параметром.

Теорема 2.1. Пусть $A(\mathfrak{X})$ – формула языка $\mathcal{L}(\mathcal{R})$, а множества M_1 и M_2 оба

а) конечные множества мощности $n \geq 0$,

в) бесконечные кобесконечные рекурсивные множества.

Тогда $\quad A(M_1)=t \Longleftrightarrow A(M_2)=t$.

Доказательство. Можно построить автоморфизм булевой алгебры $\mathfrak{R}$, отображающий M_1 на M_2.

Из теоремы 2.1. следует, что формула может характеризовать множество $\mathfrak{X}$ только с точностью до $\|\mathfrak{X}\|$ и имеет смысл выражение $A(\|\mathfrak{X}\|)$.

Пусть $y_1,\dots,y_k$ – попарно различные переменные. Упорядочим совершенные элементарные пересечения вида $y_1^*\cap\dots\cap y_k^*$, где y_i^* означает y_i или y_i' некоторым фиксированным способом и обозначим их через

$$\pi_1(y_1,\dots,y_k),\dots,\pi_{2^k}(y_1,\dots,y_k).$$

Ясно, что для любых множеств $M_1,\dots,M_k$ при $i\neq j$

$$\pi_i(M_1,\dots,M_k)\cap\pi_j(M_1,\dots,M_k)=\phi.$$

Определение 2.1. Обозначим через $[|\mathfrak{X}|=0]$ формулу $\mathfrak{X}=\phi$, а для $k\geqslant 1$ и $2^{k-1}\leqslant n<2^k$ через $[|\mathfrak{X}|=n]$ формулу

$$\begin{aligned}&(\exists y_1\dots y_k)[\mathfrak{X}\cap\pi_1(y_1,\dots,y_k)\neq\phi\ \&\dots\\ &\qquad\dots\&\ \mathfrak{X}\cap\pi_n(y_1,\dots,y_k)\neq\phi]\ \&\\ &\&\ \neg(\exists y_1\dots y_k)[\mathfrak{X}\cap\pi_1(y_1,\dots,y_k)\neq\phi\ \&\dots\\ &\qquad\dots\&\ \mathfrak{X}\cap\pi_{n+1}(y_1,\dots,y_k)\neq\phi]\end{aligned}\qquad(1)$$

Следующая теорема очевидна.

Теорема 2.2. Для любого $n\geqslant 0$ формула $[|\mathfrak{X}|=n]$ выражает предикат "$|\mathfrak{X}|=n$".

Оказывается, что пропозициональные комбинации из предикатов типа "$|\mathfrak{X}|=n$" и "$|\mathfrak{X}'|=n$" являются единственными выразимыми предикатами, а их бесконечные комбинации не выразимы.

Теорема 2.3. Пусть $B(\mathfrak{Z})\rightleftharpoons Q_1\mathfrak{X}_1\dots Q_k\mathfrak{X}_k A(\mathfrak{X}_1,\dots,\mathfrak{X}_k,\mathfrak{Z})$ – формула языка $\mathcal{L}(\mathfrak{R})$ в предваренной форме, где A не содержит кванторов, а $|M_1|,|M_1'|,|M_2|,|M_2'|\geqslant 2^k$.

Тогда $\quad B(M_1)=t \Longleftrightarrow B(M_2)=t$.

Для доказательства теоремы достаточно показать, как получить из функций Сколема для $B(M_1)$ функции Сколема для $B(M_2)$. Перестройка функций определяется индукцией по $\exists$ – группам кванторов, но из-за аналогичности случаю 5) в доказательстве теоремы 3.4 мы ее здесь не приводим.

Следствие 2.3. . Предикат "$\mathfrak{X}$ – конечно" не выразим в классической элементарной теории $\mathfrak{R}$.

Следствие 2.3.2. Каждая формула $A(\mathfrak{X})$ с одним параметром имеет в классической элементарной теории $\mathfrak{R}$ среди возможных $\|\mathfrak{X}\|$ конечную или коконечную область истинности, причем область истинности A бесконечна $\Leftrightarrow A(\omega/\omega)=t$.

Аналогично можно доказать, что формула параметрами $\mathfrak{X}_1,\dots,\mathfrak{X}_n$ может выразить только предикат, являющийся булевой комбинацией предикатов от мощности термов из $\mathfrak{X}_1,\dots,\mathfrak{X}_n$, удовлетворяющих условию следствия 2.3.2. Это дает процедуру элиминирования кванторов для классической элементарной теории $\mathfrak{R}$ (разрешимость этой теории следует из доказательства разрешимости элементарной теории булевых алгебр в [3]).

§ 3. Выразимость в теории $\mathfrak{R}_n$

Изучение выразимости начинаем опять с однопараметрических формул. Теоремы 3.1 и 3.2 являются аналогами теорем 2.1 и 2.2.

Теорема 3.1. Для каждой формулы $A(\mathfrak{X})$ можно эффективно найти такую ОРФ $f(x,y)$, что если $\mathfrak{R}_x$ и $\mathfrak{R}_y$

а) конечные множества одинаковой мощности,

в) бесконечные кобесконечные рекурсивные множества, то

$$e \circledR A(\mathfrak{R}_x) \Leftrightarrow \varphi_{f(x,y)}(e) \circledR A(\mathfrak{R}_y).$$

Доказательство. Равномерно по x и y можно построить следующую ЧРФ g:

$g(z)$ = элементу множества $\mathfrak{R}_y$ ($\mathfrak{R}_{y'}$) с номером n в порядке возрастания, если такой элемент существует и z является элементом с номером n в множестве $\mathfrak{R}_x$($\mathfrak{R}_{x'}$)

Ясно, что для $\mathfrak{R}_x$ и $\mathfrak{R}_y$ удовлетворяющих условию а) или в), g – рекурсивная перестановка и $g(\mathfrak{R}_x)=\mathfrak{R}_y$. Функция g индуцирует автоморфизм алгебры $\mathfrak{R}$, отображающий $\mathfrak{R}_x$ в $\mathfrak{R}_y$, и дает нужные соответствия реализаций. Теорема доказана.

В следующей теореме $[|\mathfrak{X}|=n]$ – формула из опр. 2.1.

Теорема 3.2. (i). $|\mathfrak{R}_x|\neq n \Rightarrow$ формула $[|\mathfrak{R}_x|=n]$ не реализуема.

(ii). Существует такая ЧРФ f, что для любых x и n имеет место: $|\mathfrak{R}_x|=n \Rightarrow f(x,n) \circledR [|\mathfrak{R}_x|=n]$.

(iii). Существует такая ЧРФ g, что для любых x, n и имеет место: e ⓡ $[|\mathcal{R}_x|=n] \Rightarrow D_{g(e,x)}=\mathcal{R}_x$.

Доказательство. Утверждение (i) очевидно.

(ii). Зная n, можно по x перебором натуральных чисел найти n элементов множества $\mathcal{R}_x$, а по ним построить множества $Y_1,\dots,Y_k$ из первой части формулы (1).

(iii). По реализующему числу e можно эффективно найти число конъюнктивных членов n в первой части формулы (1). По (i) тогда истинно $|\mathcal{R}_x|=n$. Элементы $\mathcal{R}_x$ можно найти перебором. Теорема доказана.

Обозначим через $\mathcal{F}in(\mathfrak{X})$ формулу $(\forall y)[\mathfrak{X}\subseteq y \vee \neg(\mathfrak{X}\subseteq y)]$

Теорема 3.3. (i). Если множество $\mathcal{R}_x$ – бесконечно, то формула $\mathcal{F}in(\mathcal{R}_x)$ не реализуема.

(ii). Существует такая ОРФ f, что для любых x и u

$$\mathcal{R}_x=D_u \Rightarrow f(u) \text{ ⓡ } \mathcal{F}in(\mathcal{R}_x).$$

(iii). Существует такая ЧРФ g, что для любых x и e

$$e \text{ ⓡ } \mathcal{F}in(\mathcal{R}_x) \Rightarrow D_{g(e)}=\mathcal{R}_x.$$

Утверждения (ii) и (iii) теоремы показывают, что по каноническому индексу множества $\mathcal{R}_x$ можно эффективно вычислить реализацию формулы $\mathcal{F}in(\mathcal{R}_x)$ и наоборот.

Доказательство. (i). Пусть множество $\mathcal{R}_x$ бесконечно, а $K=\{z \mid\, !\varphi_z(z)\}$. Сопоставим каждому $z\in N$ рекурсивное множество $\mathcal{R}_{h(z)}$, где

$$\varphi_{h(z)}(y)=\begin{cases}1 & \text{если } \varphi_x(y)=1 \text{ и машина Тьюринга } P_z \text{ работает над } z \text{ больше чем } y \text{ шагов,}\\ 0 & \text{иначе.}\end{cases}$$

Имеем $\mathcal{R}_{h(z)}\supseteq\mathcal{R}_x \Leftrightarrow z\in K$ и отношение $\lambda y\,.\,y\supseteq\mathcal{R}_x$ не может быть разрешимым.

(ii). По u можно найти $a_1,\dots,a_n$, где $\mathcal{R}_x=\{a_1,\dots,a_n\}$. Тогда $\mathcal{R}_y\supseteq\mathcal{R}_x \Leftrightarrow \varphi_y(a_1)=\dots=\varphi_y(a_n)=1$.

(iii). Пусть e ⓡ $\mathcal{F}in(\mathcal{R}_x)$, т.е. e ⓡ $(\forall y)[\mathcal{R}_x\subseteq y\vee\neg(\mathcal{R}_x\subseteq y)]$. Пусть d – ОРФ, вычисляющая по каноническому индексу конечного множества его характеристический индекс, т.е. $\mathcal{R}_{d(u)}=D_u$ для любого u. Тогда для каждого u имеем

$$\varphi_e(d(u)) \text{ ⓡ } [\mathcal{R}_x\subseteq\mathcal{R}_{d(u)}\vee\neg(\mathcal{R}_x\subseteq\mathcal{R}_{d(u)})].$$

Определим $g(e)=(\mu u)[j_0\varphi_e(d(u))=0]$. Покажем, что при наших предположениях $!\, g(e)$ и $D_{g(e)} = \mathcal{R}_x$. По утверждению (i) $\mathcal{R}_x$ – конечное множество. Тогда из определения канонического индекса видно, что $\mathcal{R}_x = D_v$, где $v = (\mu u)[\mathcal{R}_x \subseteq \mathcal{D}_u]$ так как из $D_u \subseteq D_v$ следует $u \leq v$. Но такое значение v и вычисляется функцией g. Теорема доказана.

<u>Следствие 3.3.1</u>. Предикаты "x – конечное множество" и "x – бесконечное множество" выразимы в теории $\mathcal{R}_r$.

В § 2 мы показали, что в классической элементарной теории $\mathcal{R}$ эти предикаты не выразимы (сл. 2.3.1).

<u>Следствие 3.3.2</u>. Теорема Райса не верна для рекурсивных множеств при кодировке их характеристическими индексами.

Называем <u>кванторной глубиной</u> формулы A максимальное число кванторов с упорядоченными линейно по включению областями действия. Следующая теорема соответствует теореме 2.3 для классической элементарной теории $\mathcal{R}$.

<u>Теорема 3.4</u>. По любой однопараметрической формуле B языка $\mathcal{L}(\mathcal{R})$ можно эффективно найти ЧРФ $f_B(e,u,v)$ со следующим свойством:

если B имеет кванторную глубину k и конечные множества $\mathcal{R}_y = D_u$ и $\mathcal{R}_z = D_v$ обладают мощностями $\geq 2^k$, то для любого e выполняется

$$e \circledR B(\mathcal{R}_y) \Rightarrow f_B(e,u,v) \circledR B(\mathcal{R}_z).$$

<u>Доказательство</u>. Не ограничивая общности можем предполагать, что в формулу B входят связанно только переменные $x_1, \ldots, x_k$. Пусть $A(x, x_1, \ldots, x_\ell)$ – подформула формулы B, не содержащая других параметров, кроме явно указанных, и требуется построить функцию $f_A^{(2\ell+3)}$ со следующим сойством:

если $\mathcal{R}_y = D_u$ и $\mathcal{R}_z = D_v$ обладают мощностями $\geq 2^k$ и если $\mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell}, \mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell}$ – такие рекурсивные множества, что для любого совершенного элементарного пересечения $\pi(x_1, \ldots, x_\ell)$ имеет место

$$|\mathcal{R}_y' \cap \pi(\mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell})| = |\mathcal{R}_z' \cap \pi(\mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell})|, \quad (2)$$

$$|\mathcal{R}_y \cap \pi(\mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell})| \geq 2^{k-\ell} \Leftrightarrow |\mathcal{R}_z \cap \pi(\mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell})| \geq 2^{k-\ell}, \quad (3)$$

$$|\mathcal{R}_y \cap \pi(\mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell})| < 2^{k-\ell} \Rightarrow$$
$$\Rightarrow |\mathcal{R}_y \cap \pi(\mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell})| = |\mathcal{R}_z \cap \pi(\mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell})|, \quad (4)$$

то для всякого e имеем

$$e \ ⓡ \ A(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}) \Longrightarrow f_A(e,u,v,\tilde{m}(\ell),\tilde{n}(\ell)) \ ⓡ \ A(\mathcal{R}_z, \tilde{\mathcal{R}}_{\tilde{n}(\ell)}),$$

где $\tilde{m}(\ell)$ означает $m_1,\dots,m_\ell$; $\tilde{\mathcal{R}}_{\tilde{m}(\ell)}$ означает $\mathcal{R}_{m_1},\dots,\mathcal{R}_{m_\ell}$ и аналогично для n вместо m.

Отметим, что задача построения f_B имеет указанный вид, причем $\ell = 0$. Покажем, как построение каждой f_A сводится к построению f для подформул A, причем вид задачи сохраняется. Рассмотрим главную связку формулы A.

1) Пусть $A(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell) \mathrel{\overline{\overline{\circ}}} C(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell) \ \&$ $\& \ D(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell)$, а функции f_C и f_D уже построены. Если $e \ ⓡ \ [C(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}) \ \& \ D(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)})]$, то

$$j_0(e) \ ⓡ \ C(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}) \text{ и } j_1(e) \ ⓡ \ D(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}).$$

Тогда $f_C(j_0(e),u,v,\tilde{m}(\ell),\tilde{n}(\ell)) \ ⓡ \ C(\mathcal{R}_z, \tilde{\mathcal{R}}_{\tilde{n}(\ell)})$,

$$f_D(j_1(e),u,v,\tilde{m}(\ell),\tilde{n}(\ell)) \ ⓡ \ D(\mathcal{R}_z, \tilde{\mathcal{R}}_{\tilde{n}(\ell)})$$

и можно взять $f_A(e,u,v,m_1,\dots,m_\ell,n_1,\dots,n_\ell) =$ $= 2^{f_C(j_0(e),u,v,\tilde{m}(\ell),\tilde{n}(\ell))} \cdot 3^{f_D(j_1(e),u,v,\tilde{m}(\ell),\tilde{n}(\ell))}$.

2) Случай дизъюнкции аналогичен.

3) Пусть $A(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell) \mathrel{\overline{\overline{\circ}}} C(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell) \supset D(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell)$ и $e \ ⓡ \ [C(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}) \supset D(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)})]$, т.е.

$$(\forall x)[x \ ⓡ \ C(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}) \Rightarrow \varphi_e(x) \ ⓡ \ D(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)})].$$

Пусть $x_1 \ ⓡ \ C(\mathcal{R}_z, \tilde{\mathcal{R}}_{\tilde{n}(\ell)})$. Тогда

$$x_2 = f_C(x_1,v,u,\tilde{n}(\ell),\tilde{m}(\ell)) \ ⓡ \ C(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}),$$

$$\varphi_e(x_2) \ ⓡ \ D(\mathcal{R}_y, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}),$$

$$f_D(\varphi_e(x_2),u,v,\tilde{m}(\ell),\tilde{n}(\ell)) \ ⓡ \ D(\mathcal{R}_z, \tilde{\mathcal{R}}_{\tilde{n}(\ell)}).$$

Положим $f_A(x_1,u,v,\tilde{m}(\ell),\tilde{n}(\ell)) =$

$$= f_D(\varphi_e(f_C(x_1,u,v,\tilde{n}(\ell),\tilde{m}(\ell)),u,v,\tilde{m}(\ell),\tilde{n}(\ell)).$$

4) Случай отрицания следует из 3).

5) Пусть $A(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell) \mathrel{\overline{\overline{\circ}}} \exists \mathfrak{X}_{\ell+1} C(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_{\ell+1})$ и

$$e \ ⓡ \ \exists \mathfrak{X}_{\ell+1} C(\mathcal{R}_y, \mathcal{R}_{m_1},\dots,\mathcal{R}_{m_\ell}, \mathfrak{X}_{\ell+1}), \text{ т.е.}$$

$j_0(e) \ ⓡ \ C(\mathcal{R}_y, \mathcal{R}_{m_1},\dots,\mathcal{R}_{m_\ell}, \mathcal{R}_{j_1(e)})$, а для множеств

$$\mathcal{R}_y, \mathcal{R}_z, \tilde{\mathcal{R}}_{\tilde{m}(\ell)}, \tilde{\mathcal{R}}_{\tilde{n}(\ell)} \tag{5}$$

17

выполняются условия (2)-(4). Покажем, как найти эффективно такое натуральное число w, что для множеств

$$\mathcal{R}_y, \mathcal{R}_{m_1}, \ldots \mathcal{R}_{m_\ell}, \mathcal{R}_{j_1(e)}, \mathcal{R}_z, \mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell}, \mathcal{R}_w \qquad (6)$$

выполняются условия (2)-(4) с $\ell+1$ вместо ℓ. Для этого пусть $g(x)$ – ЧРФ, отображающая каждое элементарное пересечение $\mathcal{R}_z' \cap \pi(\mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell})$ взаимно-однозначно на $\mathcal{R}_y' \cap \pi(\mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell})$ и не определенная на $\mathcal{R}_z$. Ввиду выполнения (2) для множеств (5) такая функция существует и ее можно построить аналогично функции g из доказательства теоремы 3.1 равномерно по y, z (или u, v), $\widetilde{m}(\ell)$ и $\widetilde{n}(\ell)$.

Если $\varphi_z(x) = 0$, то положим $\varphi_w(x) = \varphi_{j_1(e)}(g(x))$. Этим гарантируется выполнение (2) для множеств (6). Остается доопределить φ_w на конечном множестве $\mathcal{R}_z$. Элементы из каждого пересечения $\mathcal{R}_z \cap \pi(\mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell})$ делятся между $\mathcal{R}_w$ и $\mathcal{R}_w'$ именно так, чтобы выполнялись (3) и (4) для (6).

Следовательно, по индуктивному предположению, при

$$e_1 = f_C(j_0(e), u, v, \widetilde{m}(\ell), j_1(e), \widetilde{n}(\ell), w)$$

$$e_1 \circledR C(\mathcal{R}_z, \mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell}, \mathcal{R}_w) \quad \text{и}$$

$$2^{e_1} \cdot 3^{w} \circledR \exists x_{\ell+1} C(\mathcal{R}_z, \mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell}, x_{\ell+1}).$$

6) Пусть $A(x, x_1, \ldots, x_\ell) \rightleftharpoons \forall x_{\ell+1} C(x, x_1, \ldots, x_{\ell+1})$ и

$$e \circledR \forall x_{\ell+1} C(\mathcal{R}_y, \mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell}, x_{\ell+1}), \text{ т.е.}$$

$$(\forall w)[\varphi_w - \text{ХФ} \Rightarrow \varphi_e(w) \circledR C(\mathcal{R}_y, \widetilde{\mathcal{R}}_{\widetilde{m}(\ell)}, \mathcal{R}_w)] \qquad (7)$$

Покажем, как по x найти число, реализующее формулу $C(\mathcal{R}_z, \widetilde{\mathcal{R}}_{\widetilde{n}(\ell)}, \mathcal{R}_x)$, если множества $\mathcal{R}_y, \widetilde{\mathcal{R}}_{\widetilde{m}(\ell)}, \mathcal{R}_z, \widetilde{\mathcal{R}}_{\widetilde{n}(\ell)}$ удовлетворяют условиям (2)-(4). Для этого строится подобно конструкции случая 5) множество $\mathcal{R}_w$ так, что условия (2)-(4) выполняются (с $\ell+1$ вместо ℓ) для

$$\mathcal{R}_y, \mathcal{R}_{m_1}, \ldots, \mathcal{R}_{m_\ell}, \mathcal{R}_w, \mathcal{R}_z, \mathcal{R}_{n_1}, \ldots, \mathcal{R}_{n_\ell}, \mathcal{R}_x.$$

Учитывая, что по (7) имеем

$$\varphi_e(w) \circledR C(\mathcal{R}_y, \widetilde{\mathcal{R}}_{\widetilde{m}(\ell)}, \mathcal{R}_w),$$

получим по индуктивному предположению

$$f_C(\varphi_e(w), u, v, \widetilde{m}(\ell), w, \widetilde{n}(\ell), x) \circledR C(\mathcal{R}_z, \widetilde{\mathcal{R}}_{\widetilde{n}(\ell)}, \mathcal{R}_x),$$

что и требовалось.

Остается еще рассмотреть построение функции f_A для элементарной формулы A. Можем предполагать, что A имеет вид $T=\phi$, так как равенство вида $S=T$ можно заменить на $S'\cap T\cup S\cap T'=\phi$. Если подформула A находится в области действия ℓ кванторов, например по $\mathfrak{X}_1,\dots,\mathfrak{X}_\ell$, то функция f_A должна иметь вид $f_A(e,u,v,m_1,\dots,m_\ell,n_1,\dots,n_\ell)$ и терм T не может содержать переменных, кроме $\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell$. Равенство $T(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell)=\phi$ выполняется тогда и только тогда, когда все элементарные пересечения $\pi_i(\mathfrak{X},\mathfrak{X}_1,\dots,\mathfrak{X}_\ell)$ из совершенной дизъюнктивной нормальной формы терма T. Но из (2) и (4) получим, что для каждого такого пересечения имеет место

$$\pi_i(\mathcal{R}_y,\mathcal{R}_{m_1},\dots,\mathcal{R}_{m_\ell})=\phi \Leftrightarrow \pi_i(\mathcal{R}_z,\mathcal{R}_{n_1},\dots,\mathcal{R}_{n_\ell})=\phi.$$

Следовательно,

$$T(\mathcal{R}_y,\mathcal{R}_{m_1},\dots,\mathcal{R}_{m_\ell})=\phi \Leftrightarrow T(\mathcal{R}_z,\mathcal{R}_{n_1},\dots,\mathcal{R}_{n_\ell})=\phi$$

и для единственного возможного реализующего элементарную формулу числа 0 имеем

$$0 \; ⓡ \; A(\mathcal{R}_y,\tilde{\mathcal{R}}_{\tilde{m}(\ell)}) \Leftrightarrow 0 \; ⓡ \; A(\mathcal{R}_z,\tilde{\mathcal{R}}_{\tilde{n}(\ell)}).$$

Можем выбрать тождественно $f_A(e,u,v,\tilde{m}(\ell),\tilde{n}(\ell))=0$

Теорема доказана.

Из доказательства теоремы 3.4 видно, что для формул с параметрами имеет место следующая

Теорема 3.5. По любой n-параметрической формуле B языка $\mathcal{L}(\mathcal{R})$ для любого собственного подмножества $I\subset\{1,\dots,2^n\}$ можно эффективно найти ЧРФ $f^{(2n+3)}_{B,I}(e,u,v,\tilde{y}(n),\tilde{z}(n))$ со следующим свойством:

если B имеет кванторную глубину k и для множеств $\mathcal{R}_{y_1},\dots,\mathcal{R}_{y_n},\mathcal{R}_{z_1},\dots,\mathcal{R}_{z_n}$ удовлетворены условия

1) при $i\notin I$ $|\pi_i(\mathcal{R}_{y_1},\dots,\mathcal{R}_{y_n})|=|\pi_i(\mathcal{R}_{z_1},\dots,\mathcal{R}_{z_n})|=\omega$,

2) при $i\in I$ эти пересечения конечны и

$$|\pi_i(\tilde{\mathcal{R}}_{\tilde{y}(n)})|\geq 2^k \Leftrightarrow |\pi_i(\tilde{\mathcal{R}}_{\tilde{z}(n)})|\geq 2^k$$

$$|\pi_i(\tilde{\mathcal{R}}_{\tilde{y}(n)})|<2^k \Rightarrow |\pi_i(\tilde{\mathcal{R}}_{\tilde{y}(n)})|=|\pi_i(\tilde{\mathcal{R}}_{\tilde{z}(n)})|,$$

то для любого e и для канонических индексов u и v множеств $\bigcup_{i\in I}\pi_i(\tilde{\mathcal{R}}_{\tilde{y}(n)})$ и $\bigcup_{i\in I}\pi_i(\tilde{\mathcal{R}}_{\tilde{z}(n)})$ выполняется

$$e \; ⓡ \; B(\tilde{\mathcal{R}}_{\tilde{y}(n)}) \Rightarrow f_{B,I}(e,u,v,\tilde{y}(n),\tilde{z}(n)) \; ⓡ \; B(\tilde{\mathcal{R}}_{\tilde{z}(n)})$$

Следствие 3.51. Предикат "$\|x\| = \|y\|$" не выразим в теории $\mathcal{R}_r$.

Доказательство. Допустим от противного, что этот предикат выразим формулой $B(y_1, y_2)$ кванторной глубины k. Пусть $|\mathcal{R}_{y_1}| = |\mathcal{R}_{y_2}| = |\mathcal{R}_{z_1}| = 2^k, |\mathcal{R}_{z_2}| = 2^{k+1}$ и $\mathcal{R}_{y_1} \cap \mathcal{R}_{y_2} = \mathcal{R}_{z_1} \cap \mathcal{R}_{z_2} = \emptyset$, а $I \subset \{1,2,3,4\}$ – множество номеров для $y_1 \cap y_2$, $y_1 \cap y_2'$ и $y_1' \cap y_2$ в нумерации элементарных пересечений из § 2. Пусть $\mathcal{R}_{y_1} \cup \mathcal{R}_{y_2} = D_u$ и $\mathcal{R}_{z_1} \cup \mathcal{R}_{z_2} = D_v$. Тогда по теореме 3.5 имеем

$$e \; r \; B(\mathcal{R}_{y_1}, \mathcal{R}_{y_2}) \Rightarrow f_{B,I}(e, u, v, y_1, y_2, z_1, z_2) \; r \; B(\mathcal{R}_{z_1}, \mathcal{R}_{z_2})$$

Теоремы 3. –3.5 дают следующее описание выразимых в теории $\mathcal{R}_r$ предикатов.

Теорема 3.6. Определенный на рекурсивных множествах предикат $P(x_1, \dots, x_n)$ выразим в теории $\mathcal{R}_r$ тогда и только тогда, когда P является булевой комбинацией предикатов $P_i(T_i(x_1, \dots, x_n))$, где T_i – термы, а истинность $P_i(y)$ зависит только от мощности y, причем каждый из предикатов P_i имеет среди возможных $|y|$ конечную или коконечную область истинности.

Результаты §§ 2 и 3 показывают, что по сравнению с классической теорией в теории $\mathcal{R}_r$ класс выразимых предикатов расширился по существу только предикатом "x конечно". Но это не означает, что теория $\mathcal{R}_r$ равносильна классической теории с добавленным предикатным символом. Формулу теории $\mathcal{R}_r$ характеризует кроме области реализуемости еще необходимая информация, содержащаяся в реализации. И в этом смысле формулы, выражающие один и тот же "классический предикат", могут оказаться неравносильными.

Литература

1. К л и н и С.К., Введение в метаматематику. Москва, 1957.
2. Р о д ж е р с Х., Теория рекурсивных функций и эффективная вычислимость. Москва, 1972.
3. Е р ш о в Ю.П., Разрешимость элементарной теории дистрибутивных структур с относительными дополнениями и теории фильтров. Алгебра и логика, 1964, 3, № 3, 17–38.

Поступило

4 III 979

EXPRESSIBILTY IN ELEMENTARY THEORY OF RECURSIVE SETS WITH REALIZABILITY LOGIC

R. Prank

Summary

Let $\mathfrak{R}$ be the Boolean algebra of recursive sets, and $\mathcal{L}(\mathfrak{R})$ denote the first order language in signature $\langle \emptyset, N, \cup, \cap, ' \rangle$ We define the nonclassical elementary theory $\mathfrak{R}_r$ consisting of all recursively realizable (def. 1.2) formulas of $\mathcal{L}(\mathfrak{R})$. We write $\|\mathfrak{X}\| = \mu/\nu$ $(\mu, \nu \in N \cup \{\omega\})$ if $|\mathfrak{X}| = \mu$ and $|\mathfrak{X}'| = \nu$.

The predicates $P(\mathfrak{X})$ defined on $\mathfrak{R}$, and expressible in the classical elementary theory of $\mathfrak{R}$ are characterized as predicates depending only on $\|\mathfrak{X}\|$, and being true on the infinite (if $P(\omega/\omega) = f$) or coinfinite (if $P(\omega/\omega) = t$)set of $\|\mathfrak{X}\|$.

The predicate "$\mathfrak{X}$ is finite" is expressible in $\mathfrak{R}_r$ but not in classical theory. The predicates $P(\mathfrak{X})$ expressible in $\mathfrak{R}_r$ are those being true on the finite or cofinite set of $\|\mathfrak{X}\|$ in both typed μ/ω and ω/ν.

СОДЕРЖАНИЕ

CONTENTS INHALT

Ученые записки Тартуского государственного университета. Выпуск 500. ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ РЕШЕНИЯ ДИФФЕРЕНЦИАЛЬНЫХ И ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ. Труды по математике и механике ХХУ. На русском языке. Резюме на английском и немецком языках. Тартуский государственный университет. ЭССР, г. Тарту, ул. Юликооли, 18. Ответственный редактор Э.Тамм. Корректоры С.Пискарев, Г.Лийв, К.Уусталу. Сдано в печать 20/06 1979. Бумага печатная 30х45 1/4. Печ.листов 8,25. Учетно-издат. листов 6,59. Тираж 400. МВ 02771. Типография ТГУ, ЭССР, г. Тарту, ул. Пялсона, 14. Зак. № 967. Цена 1 руб.