Прикладная математика и механика. Том 40, 1976

УДК 62—50

СТАБИЛИЗАЦИЯ В ЛИНЕЙНОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ С ИНТЕГРАЛЬНЫМИ ОГРАНИЧЕНИЯМИ НА РЕСУРСЫ УПРАВЛЕНИЙ ИГРОКОВ

В. М. Решетов, В. Н. Ушаков

(Свердловск)

Рассматривается игровая задача сближения для линейной управляемой системы. Предполагается, что управляющие воздействия первого и второго игроков стеснены интегральными ограничениями. На основе схемы управления с поводырем [1] и методов теории стабилизации [2] строится стабилизированная процедура управления, обеспечивающая устойчивое сближение порождаемых ею движений с заданным целевым множеством. Содержание работы примыкает к исследованиям [1,3–5].

1. Пусть движение управляемой системы описывается векторным дифференциальным уравнением

$$dx / dt = Ax + Bu + Cv, \quad x[t_0] = x_0 \tag{1.1}$$

Здесь x — n-мерный фазовый вектор системы; u, v — управляющие воздействия игроков размерности m; A, B, C — постоянные матрицы соответствующих размерностей.

Реализации управляющих воздействий $u[t]$, $v[t]$ стеснены на $[t_0, \infty]$ соотношениями

$$I_u(t_0, \infty) \leqslant \mu[t_0], \quad I_v(t_0, \infty) \leqslant \nu[t_0] \tag{1.2}$$

$$I_u(a, b) = \left(\int_a^b \|u[\xi]\|^2 d\xi\right)^{1/2}, \quad I_v(a, b) = \left(\int_a^b \|v[\xi]\|^2 d\xi\right)^{1/2}$$

Здесь $\mu[t_0]$, $\nu[t_0]$ — ограничения на ресурсы управляющих воздействий u, v.

Предполагается, что изменение величин $\mu[t]$, $\nu[t]$ определяется расходуемыми ресурсами

$$\mu^2[t+\Delta] = \mu^2[t] - I_u^2(t, t+\Delta), \quad \nu^2[t+\Delta] = \nu^2[t] - I_v^2(t, t+\Delta) \tag{1.3}$$

В евклидовом пространстве R^{n+3} точек $z = (t, \mu, \nu, x)$ задано множество

$$M = \{z = (t, \mu, \nu, x) : (t, x) \in M^*, \ \mu \geqslant 0, \ \nu \geqslant 0\}$$

Здесь M^* — замкнутое множество пространства векторов (t, x), содержащееся в области $D = \{(t, x) : t_0 \leqslant t \leqslant \vartheta, \ x \in R^n\}$ (ϑ — момент времени, ограничивающий продолжительность игры).

Задача, стоящая перед первым игроком, заключается в построении такого позиционного способа действия U^*, который обеспечивал бы первому игроку попадание на M_ε^* точки $(t, x_\Delta[t])$, каково бы ни было аппроксимационное движение $x_\Delta[t]$, порожденное этим способом действия; здесь M_ε^* — ε-окрестность множества M^*.

В работе [5] на основе общей схемы управления с поводырем [1] было построено решение задачи сближения для дифференциальной игры с интегральными ограничениями. Это решение обладало свойством устойчивости относительно малых ошибок измерения фазового вектора. Однако в случае достаточно большой продолжительности игры устойчивость решения обеспечивалась лишь при чрезмерно малых ограничениях на величину шага аппроксимационной схемы и ошибку измерения фазового вектора, поскольку оценка рассогласования между движением $x_\Delta[t]$ реальной системы и движением $w_\Delta[t]$ поводыря носила экспоненциальный характер. В данной работе для улучшения оценки рассогласования искомый способ управления U^* строится в виде стабилизированной процедуры управления с поводырем. Такое решение игровых задач с геометрическими ограничениями на управления игроков было предложено в [3, 4].

Будем предполагать, что выполнены следующие условия.

Условие А. Существует u-стабильный мост W_u^ϑ.

Определение u-стабильного моста W_u^ϑ понимается здесь в смысле работы [5].

Условие Б. Система, описываемая уравнением

$$ds \,/\, dt = As + Br, \quad s \in R^n, \quad r \in R^m \tag{1.4}$$

стабилизируема [2]. Здесь r — вектор управляющего воздействия.

Выполнение условия Б означает существование такой линейной вектор-функции $r_0(s) = R_0 s$, что нулевое решение системы (1.4) будет асимптотически устойчиво по Ляпунову.

Тогда в силу теоремы Ляпунова (см. [2], стр. 453) для любой наперед заданной определенно-отрицательной квадратичной формы $\beta(s)$ найдется определенно-положительная квадратичная форма $\lambda(s)$, полная производная которой в силу уравнения

$$ds \,/\, dt = As + BR_0 s$$

будет удовлетворять равенству

$$d\lambda \,/\, dt = (\partial\lambda/\partial s)' \, (As + BR_0 s) = \beta(s) \tag{1.5}$$

Здесь и ниже штрих означает транспонирование.

2. Дадим определение аппроксимационной процедуры управления с поводырем первого игрока, решающей задачу о сближении. Эта процедура предполагает наличие вспомогательной конструкции, состоящей из u-стабильного моста W_u^ϑ, правила выбора поводыря на этом мосту и правила стабилизированного прицеливания на движение поводыря.

Следуя схеме управления с поводырем [1], наряду с реальной системой, описываемой соотношениями (1.1) и (1.3), рассмотрим вспомогательную идеальную систему, движение $w[t]$ которой описывается уравнением

$$dw \,/\, dt = Aw + Bu_* + Cv_*, \quad w[t_0] = w_0 \tag{2.1}$$

Здесь реализации $u_*[t]$ и $v_*[t]$ управляющих воздействий стеснены на $[t_0, \infty)$ соотношениями

$$I_{u_*}(t_0, \infty) \leqslant \mu_*[t_0], \quad I_{v_*}(t_0, \infty) \leqslant \nu_*[t_0]$$
$$\nu_*[t_0] = \nu[t_0], \quad \mu_*^2[t_0] = \mu^2[t_0] - \eta^2[t_0]$$

где $\eta\,[t_0] > 0$ — ресурс, используемый в процедуре управления для стабилизации движения $s\,[t] = x\,[t] - w\,[t]$ на $[t_0, \vartheta]$.

Введем вспомогательные понятия и обозначения.

Пусть $z_* = (t_*, \mu_*, \nu_*, w)$ — некоторая позиция игры, $v_*\,[t]$ — реализация управления второго игрока, допустимая для этой позиции, т. е. $v_*\,[t]$ — суммируемая функция, удовлетворяющая неравенству $I_{v_*}(t_*, \infty) \leqslant \nu_*$.

Символом $G^{(u)}(z_*, v_*\,[\cdot])$ обозначим замыкание множества точек $z = (t, \mu\,[t], \nu\,[t], w\,[t])$ вида

$$
\begin{gathered}
t_* \leqslant t, \quad 0 \leqslant \mu^2\,[t] \leqslant \mu_*^2 - I_{u*}^2\,(t_*, t) \\
0 \leqslant \nu^2\,[t] \leqslant \nu_*^2 - I_{v*}^2\,(t_*, t) \\
w\,[t] = w + \int_{t_*}^{t} (Aw\,[\sigma] + Bu_*\,[\sigma] + Cv_*\,[\sigma])\,d\sigma
\end{gathered}
$$

где $u_*\,[\sigma]$ $(\sigma \geqslant t_*)$ — всевозможные суммируемые функции, удовлетворяющие неравенству $I_{u_*}(t_*, \infty) \leqslant \mu_*$.

Введем в рассмотрение функции

$$
\begin{aligned}
&u\,[s;\, t,\, \delta;\, u_*\,[\cdot]] = p\,[s;\, t,\, \delta;\, u_*\,[\cdot]] + r\,(s) \\
&v_*\,[s;\, t,\, \delta;\, v\,[\cdot]] = q\,[s;\, t,\, \delta;\, v\,[\cdot]]
\end{aligned}
\tag{2.2}
$$

где величины в правых частях определяются из соотношений (t_0 — начальный момент времени)

$$
\begin{gathered}
p\,[s;\, t,\, \delta;\, u_*\,[\cdot]] = 0, \quad q\,[s;\, t,\, \delta;\, v\,[\cdot]] = 0, \quad t = t_0 \\
(\partial\lambda\,(s)\,/\,\partial s)'\, Bp\,[s;\, t,\, \delta;\, u_*\,[\cdot]] = \min_p\,(d\lambda\,(s)\,/\,ds)'\, Bp \\
\| p \| = \delta^{-1/2} I_{u_*}(t - \delta, t), \quad t \in [t_0, \vartheta], \quad 0 < \delta \leqslant t - t_0 \\
(\partial\lambda\,(s)\,/\,\partial s)'\, Cq\,[s;\, t,\, \delta;\, v\,[\cdot]] = \max_q\,(\partial\lambda\,(s)\,/\,\partial s)'\, Cq \\
\| q \| = \delta^{-1/2} I_v\,(t - \delta, t), \quad t \in (t_0, \vartheta], \quad 0 < \delta \leqslant t - t_0 \\
r\,[s] = R_0 s
\end{gathered}
$$

Здесь $u_*\,[\cdot]\; v\,[\cdot]$ — суммируемые с квадратом функции, заданные на $[t - \delta, t]$, $s \in R^n$.

Пусть заданы исходная позиция $z\,[t_0] = (t_0, \mu\,[t_0], \nu\,[t_0], x\,[t_0])$ и u-стабильный мост W_u^ϑ. Полагая, что мост W_u^ϑ замкнут в пространстве R^{n+3}, выберем точку $z_*\,[t_0] = (t_0, \mu_*\,[t_0], \nu_*\,[t_0], w\,[t_0]) \in W_u^\vartheta$, ближайшую к точке $z\,[t_0]$. Эта точка $z_*\,[t_0]$ — позиция поводыря в исходный момент $t = t_0$.

Пусть $\Gamma_\Delta = \{t_0, t_1, \ldots, t_N = \vartheta\}$ — разбиение отрезка $[t_0, \vartheta]$ моментами t_i на полуинтервалы равной длины Δ.

При определении процедуры управления с поводырем первого игрока будем считать, что система (2.1) находится в распоряжении первого игрока. Этот игрок, распоряжаясь как выбором управления $u\,[t]$ в системе (1.1), (1.3), так и выбором управлений $u_*\,[t]$, $v_*\,[t]$ в системе (2.1), имеет возможность точно вычислять реализующиеся значения фазового вектора $w_\Delta\,[t]$, $t \in [t_0, \vartheta]$ системы (2.1). Однако значения фазового вектора $x_\Delta\,[t]$ и, вместе с тем, вектора $s_\Delta\,[t] = x_\Delta\,[t] - w_\Delta\,[t]$ он вычисляет с погрешностью $\Delta s_\Delta[t]$ и на основании вектора

$$s_\Delta{}^*[t_i] = s_\Delta[t_i] + \Delta s_\Delta[t_i] \quad (\|\Delta s_\Delta[t_i]\| \leqslant \xi)$$

строит управления $u[t]$, $v_*[t]$ на полуинтервале $[t_i, t_{i+1})$.

Опишем подробнее процедуру управления с поводырем. На первом полуинтервале $[t_0, t_1)$ движение $\varkappa(t) \equiv (\mu_\Delta[t], \nu_\Delta[t], x_\Delta[t])$ системы (1.1), (1.3) с начальным условием $\varkappa(t_0) = (\mu[t_0], \nu[t_0], x[t_0])$ порождается постоянным управлением первого игрока

$$u[t] = u[s_\Delta{}^*[t_0]; t_0, \Delta; u_*[\cdot]] \quad (t_0 \leqslant t < t_1)$$

в паре с некоторой допустимой реализацией $v[t]$ $(t \geqslant t_0)$ управления второго игрока. Здесь

$$s_\Delta{}^*[t_0] = s_\Delta[t_0] + \Delta s_\Delta[t_0]$$
$$s_\Delta[t_0] = x_\Delta[t_0] - w_\Delta[t_0], \quad w_\Delta[t_0] = w[t_0]$$

Пусть

$$v_*[t] = v_*[s_\Delta{}^*[t_0]; t_0, \Delta; v[\cdot]] \quad (t_0 \leqslant t < t_1)$$

Выберем позицию поводыря $z_{*\Delta}[t_1]$ в момент t_1 из условия

$$z_{*\Delta}[t_1] \in W_u{}^\vartheta(t_1) \cap G^{(u)}(z_{*\Delta}[t_0], v_*[\cdot])$$

предполагая, что это пересечение не пусто. Здесь

$$z_{*\Delta}[t_0] = z_*[t_0]$$
$$W_u{}^\vartheta(t) = \{z : z = (t, \mu, \nu, x), z \in W_u{}^\vartheta\}$$

Далее, на полуинтервале $[t_1, t_2)$ движение $\varkappa(t)$ системы (1.1), (1.3) с начальным значением $\varkappa(t_1)$ порождается постоянным управлением

$$u[t] = u[s_\Delta{}^*[t_1]; t_1, \Delta; u_*[\cdot]] \quad (t_1 \leqslant t < t_2)$$

где $u_*[\sigma]$ $(t_0 \leqslant \sigma < t_1)$ — управление первого игрока, которое в паре с $v_*[\sigma]$ $(t_0 \leqslant \sigma < t_1)$ порождает позицию $z_{*\Delta}[t_1]$ поводыря.

В результате выбора управления $u[t]$ $(t_1 \leqslant t < t_2)$ и некоторого управления $v[t]$ $(t \geqslant t_1)$ реализуется позиция игры $z_\Delta[t_2]$.

Выберем

$$v_*[t] = v_*[s_\Delta{}^*[t_1]; t_1, \Delta; v[\cdot]] \quad (t_1 \leqslant t < t_2)$$

и позицию поводыря в момент $t = t_2$ определим из условия

$$z_{*\Delta}[t_2] \in W_u{}^\vartheta(t_2) \cap G^{(u)}(z_{*\Delta}[t_1], v_*[\cdot])$$

полагая, что это пересечение не пусто.

Если на последующих полуинтервалах $[t_i, t_{i+1})$ $(i = 2, 3, \ldots, N-1)$ также выполняется условие $W_u{}^\vartheta(t_{i+1}) \cap G^{(u)}(z_{*\Delta}[t_i], v_*[\cdot]) \neq \phi$, то указанная процедура осуществляется до момента $t = \vartheta$.

В случае, когда это условие не выполняется, найдется момент t_j, при котором впервые

$$W_u{}^\vartheta(t_j) \cap G^{(u)}(z_{*\Delta}[t_{j-1}], v_*[\cdot]) = \phi$$

Тогда из условия $z_{*\Delta}[t_{j-1}] \in W_u{}^\vartheta(t_{j-1})$ и определения u-стабильности вытекает, что

$$M_{[t_{j-1}, t_j]} \cap G^{(u)}(z_{*\Delta}[t_{j-1}], v_*[\cdot]) \neq \phi$$

т. е. существует момент времени $t^* \in [t_{j-1}, t_j]$, когда позицию поводыря можно определить из условия

$$z_{*\Delta}[t^*] \in M_{t^*} \cap G^{(u)}(z_{*\Delta}[t_{j-1}], v_*[\cdot])$$

Здесь

$$M_{[t_*,t^*]} = \{z : z = (t, \mu, \nu, x) \in M, \quad t_* \leqslant t \leqslant t^*\}$$
$$M_{t^*} = \{z : z = (t^*, \mu, \nu, x) \in M\}$$

На последующих полуинтервалах $[t_i, t_{i+1})$ $(j \leqslant i \leqslant N-1)$ управления $u[t]$, $v_*[t]$ определяем соотношениями

$$u[t] = u[s_\Delta^*[t_i]; t_i, \Delta; u_*[\cdot]]$$
$$v_*[t] = v_*[s_\Delta^*[t_i]; t_i, \Delta; v[\cdot]]$$

где $u_*[\cdot]$ $[t_i, t_{i+1})$ $(j \leqslant i \leqslant N-1)$ выбирается произвольным образом.

Реальное движение $(\mu_\Delta[t], \nu_\Delta[t], x_\Delta[t])$ и движение поводыря $(\mu_{*\Delta}[t], \nu_{*\Delta}[t], w_\Delta[t])$, которые реализуются на $[t_0, \vartheta]$ в процессе управления с поводырем, будем называть аппроксимационными движениями.

3. Из приведенного в п. 2 определения аппроксимационной процедуры управления с поводырем первого игрока вытекает, что возмущение $s_\Delta[t] = x_\Delta[t] - w_\Delta[t]$ на полуинтервале $[t_i, t_{i+1})$ описывается уравнением

$$ds_\Delta[t]/dt = As_\Delta[t] + B(p[s_\Delta^*[t_i], t_i] + R_0 s_\Delta^*[t_i] - u_*[t]) + C(v[t] - q[s_\Delta^*[t_i], t_i]) \tag{3.1}$$

где

$$p[s_\Delta^*[t_i], t_i] = p[s_\Delta^*[t_i]; t_i, \Delta; u_*[\cdot]]$$
$$q[s_\Delta^*[t_i], t_i] = q[s_\Delta^*[t_i], t_i, \Delta; v[\cdot]]$$

Полная производная функции Ляпунова $\lambda(s)$ в силу уравнения (3.1) имеет вид на $[t_i, t_{i+1})$

$$d\lambda(s_\Delta[t])/dt = (\partial\lambda/\partial s)'_{s=s_\Delta^*[t]}(As_\Delta[t] + B(p[s_\Delta^*[t_i], t_i] + R_0 s_\Delta^*[t_i] - u_*[t]) + C(v[t] - q[s_\Delta^*[t_i], t_i])) \tag{3.2}$$

Из равенства (3.2) вытекает соотношение, справедливое для любых двух моментов $t_{k_1}, t_{k_2} \in \Gamma_\Delta$

$$\int_{t_{k_1}}^{t_{k_2}} \left(d\lambda \frac{(s_\Delta[t])}{dt}\right) dt = \sum_{i=k_1}^{k_2-1} \int_{t_i}^{t_{i+1}} \left(d\lambda \frac{(s_\Delta[t])}{dt}\right) dt = \tag{3.3}$$
$$= \sum_{i=k_1}^{k_2-1} \int_{t_i}^{t_{i+1}} \Bigg\{ \left(\partial\lambda \frac{(s)}{\partial s}\right)'_{s=s_\Delta^*[t_i]} B(p[s_\Delta^*[t_i], t_i] - u_*[t]) +$$
$$+ \left(\partial\lambda \frac{(s)}{\partial s}\right)'_{s=s_\Delta^*[t_i]} C(v[t] - q[s_\Delta^*[t_i], t_i]) +$$
$$+ \left(\partial\lambda \frac{(s)}{\partial s}\right)'_{s=s_\Delta^*[t_i]} A(s_\Delta[t] - s_\Delta^*[t_i]) + \left[\left(\partial\lambda \frac{(s)}{\partial s}\right)_{s=s_\Delta^*[t]} -\right.$$

$$-\left(\partial\lambda\frac{(s)}{\partial s}\right)_{s=s_\Delta^*[t_i]}\Big]'[As_\Delta[t]+B(p[s_\Delta^*[t_i],t_i]+R_0s_\Delta^*[t_i]-$$

$$-u_*[t])+C(v[t]-q[s_\Delta^*[t_i],t_i])]+\beta(s_\Delta^*[t])\Big\}dt$$

Для обоснования того, что предлагаемая процедура управления с поводырем обеспечивает ε-близость аппроксимационных движений $w_\Delta[t]$ и $x_\Delta[t]$, проведем оценку сверху интеграла в левой части первого равенства (3.3). Можно показать, что справедлива оценка

$$\int\limits_{t_{k_1}}^{t_{k_2}}\left(d\lambda\frac{(s_\Delta[t])}{dt}\right)dt\leqslant\sum_{i=k_1}^{k_2-1}\{K(1+\|s_\Delta[t_i]\|)^2(\sqrt{\Delta}+\zeta)+ \tag{3.4}$$

$$+\beta(s_\Delta[t_i])\}\Delta+K(\sqrt{\Delta}+\zeta),\quad K=\mathrm{const}>0$$

Из оценки (3.4), используя рассуждения теории устойчивости [2], приходим к выводу: для любого наперед выбранного $\varepsilon>0$ можно указать числа $\Delta(\varepsilon)>0$ и $\zeta(\varepsilon)>0$, такие, что всякое аппроксимационное движение $s_\Delta[t]$ с начальным условием $\|s_\Delta[t_{k_1}]\|\leqslant\zeta(\varepsilon)$ будет при всех $t\in[t_{k_1},t_{k_2}]$ удовлетворять неравенству

$$\|s_\Delta[t]\|\leqslant\varepsilon \tag{3.5}$$

если только $\Delta\leqslant\Delta(\varepsilon)$, $\|\Delta s_\Delta[t]\|\leqslant\zeta(\varepsilon)$.

Заметим, что все предыдущие рассуждения и вывод будут справедливы только в том случае, когда стабилизирующая часть $r[s_\Delta^*[t]]$ управления $u[t]$ первого игрока удовлетворяет ограничению

$$I\equiv\int\limits_{t_0}^{\vartheta}\|r[s_\Delta^*[t]]\|^2dt\leqslant\mu^2[t_0]-\mu_*^2[t_0]=\eta^2[t_0] \tag{3.6}$$

Из оценки (3.4) видно, что для конечного промежутка времени $[t_0,\vartheta]$ можно подобрать диаметр Δ разбиения Γ_Δ и ограничение ζ на ошибку измерения фазового вектора, такие, что при $\|s_\Delta[t_{k_1}]\|\leqslant\zeta$ управление $r[s_\Delta^*[t]]=R_0s_\Delta^*[t]$ является допустимым, т. е. удовлетворяет соотношению (3.6).

В этом случае неравенство (3.5) будет означать, что описанная процедура управления обеспечивает ε-близость аппроксимационных движений $x_\Delta[t]$ реальной системы и $w_\Delta[t]$ поводыря, если только шаг разбиения Δ и ошибка измерения ζ будут достаточно малыми.

Отметим здесь, что в отличие от оценки рассогласования между движениями реальной системы и поводыря, приведенной в [5], оценка (3.4) уже не содержит экспоненциального множителя.

Если же измерение фазового вектора системы производится с погрешностью, то наличие такой погрешности при достаточно большом значении ϑ приводит к чрезмерному накоплению величины I, т. е. накоплению, нарушающему неравенство (3.6). Однако, если управление первого игрока представимо в виде $u=p+r$, где часть управляющих сил p стеснена интегральным ограничением, а другая часть r — геометрическим ограничением, то небольшое изменение приведенной выше процедуры обеспечит устойчивое сближение всех аппроксимационных движений $x_\Delta[t]$ с ε-окрестностью множества M^* к моменту ϑ, где ϑ — сколько угодно большая величина.

Авторы благодарны А. И. Субботину за постоянное внимание к работе и полезные советы.

Поступила 17 IX 1975

ЛИТЕРАТУРА

1. *Красовский Н. Н.* Дифференциальная игра сближения — уклонения. I, II. Изв. АН СССР. Техн. кибернетика, 1973, № 2—3.
2. *Малкин И. Г.* Теория устойчивости движения. М., «Наука», 1966.
3. *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М., «Наука», 1974.
4. *Решетов В. М.* Об одной линейной дифференциальной игре уклонения. ПММ, 1974, т. 38, вып. 4.
5. *Субботин А. И., Ушаков В. Н.* Альтернатива для дифференциальной игры сближения — уклонения при интегральных ограничениях на управления игроков. ПММ, 1975, т. 39, вып. 3.