Доклады Академии наук СССР
1972. Том 207, № 1

УДК 518.731.343.1 МАТЕМАТИКА

В. Д. БАТУХТИН

# ЭКСТРЕМАЛЬНОЕ ПРИЦЕЛИВАНИЕ В НЕЛИНЕЙНОЙ ИГРЕ СБЛИЖЕНИЯ

*(Представлено академиком Н. Н. Красовским 14 X 1971)*

Рассматривается дифференциальная игра сближения фазовой точки $x[t]$ с замкнутым множеством $\mathcal{M}$. В предположении единственности программного максиминного оптимального управления доказывается дифференцируемость цены игры. Устанавливается возможность успешного завершения позиционной игры выведения $x[t]$ на $\mathcal{M}$.

Пусть движение конфликтно управляемой системы в пространстве $E^n$ описывается дифференциальным уравнением

$$dx\,/\,dt = f^{(1)}(t, x, u) + f^{(2)}(t, x, v), \quad x(t_0) = x_0, \tag{1}$$

где $x$ — фазовый вектор системы; $u$ и $v$ — векторы управляющих воздействий игроков, стесненные ограничениями

$$u[t] \in P \subset E^r, \qquad v[t] \in Q \subset E^k; \tag{2}$$

$P$ и $Q$ — компакты; функции $f^{(1)}(t, x, u)$ и $f^{(2)}(t, x, v)$ непрерывны по всем аргументам и непрерывно дифференцируемы по $x$. Кроме того, производные $\partial f^{(1)}(t, x, u)\,/\,\partial x$ и $\partial f^{(2)}(t, x, v)\,/\,\partial x$ равномерно ограничены в $E^n$.

Допустимой стратегией первого игрока $U \div U(t, x)$ будем называть функцию, которая вектору $x \in E^n$ ставит в соответствие некоторое непустое замкнутое множество $U(t, x) \subset P$, причем множество $U(t, x)$ полунепрерывно сверху относительно включения по $\{t, x\}$. (Символ $\div$ означает соответствие между стратегией $U$ и функцией $U(t, x)$, которая задает эту стратегию.)

Обозначим через $F^{(2)}(t, x)$ выпуклую оболочку множества, пробегаемого вектором $f^{(2)}(t, x, v)$, когда $v$ пробегает множество $Q$. Совокупность возможных действий второго игрока опишем множеством $F^{(2)}(t, x)$. Будем называть движением системы (1), порожденным стратегией $U \div U(t, x)$, всякую абсолютно непрерывную векторфункцию $x[t] = x[t;\ t_0, x_0, U]$ ($t_0 \leqslant t \leqslant \vartheta$, где $\vartheta$ — некоторый конечный момент времени), которая при почти всех $t \geqslant t_0$, удовлетворяет контингенции

$$dx\,[t]/dt \in F_U^{(1)}(t, x\,[t]) + F^{(2)}(t, x\,[t]). \tag{3}$$

$F_U^{(1)}(t, x[t])$ — выпуклая оболочка множества, пробегаемого вектором $f^{(1)}(t, x[t], u)$ в позиции $\{t, x[t]\}$, когда $u$ пробегает множество $U(t, x)$. Из определения способа действий игроков вытекает, что совокупность движений $x[t;\ t_0, x_0\ U]$ содержит всякое движение, порожденное стратегией $U \div U(t, x)$ в паре с любым способом формирования управления $v[t]$, порождающим измеримые реализации $v[t] \in Q, t \geqslant t_0$.

Проблема сближения с заданным замкнутым множеством $\mathcal{M}$ для первого игрока формализуется в следующих двух задачах ([1]).

Задача 1. Среди допустимых стратегий $U(t, x)$ требуется найти оптимальную минимаксную стратегию $U^0(t, x)$, которая удовлетворяет равенству

$$\min_U \max_{x[t]} \rho\,(x_U\,[\vartheta], \mathcal{M}) = \max_{x[t]} \rho\,(x_{U^0}\,[\vartheta], \mathcal{M}), \tag{4}$$

где $\rho(x_U[\vartheta], \mathcal{M})$ — расстояние от точки $x_U[\vartheta] = x[\vartheta;\ t_0, x_0, U]$ до множества $\mathcal{M}$.

Задача 2. Среди допустимых стратегий $U(t, x)$ требуется найти стратегию $U^*(t, x)$, которая гарантирует выведение точки $x_U[t]$ на множество $\mathscr{M}$ не позже некоторого момента $\vartheta^0$, т. е. обеспечивает выполнение равенства

$$\rho(x_{U^*}[\vartheta], \mathscr{M}) = 0, \quad \vartheta \leqslant \vartheta^0. \tag{5}$$

Для решения этих задач введем вспомогательную конструкцию. Назовем программными управлениями борелевские меры $\mu(dt, du)$, $\nu(dt, dv)$, заданные на множествах $T \times P$, $T \times Q$ и удовлетворяющие условиям

$$\int_{T^*}\int_P \mu(dt, du) = \lambda(T^*), \quad \int_{T^*}\int_Q \nu(dt, dv) = \lambda(T^*), \tag{6}$$

где $\lambda(T^*)$ — лебеговская мера множества $T^* \subseteq T = [t_0, \vartheta]$. Определим программное движение $x(t)$ как абсолютно непрерывную вектор-функцию $x(t) = x(t;\ t_0, x_0, \mu, \nu)$, $t_0 \leqslant t \leqslant \vartheta$, удовлетворяющую при всех $t \in [t_0, \vartheta]$ уравнению

$$x(t) = x_0 + \int_{[t_0, t]}\int_P f^{(1)}(\tau, x(\tau), u)\,\mu(d\tau, du) + \int_{[t_0, t]}\int_Q f^{(2)}(\tau, x(\tau), v)\,\nu(d\tau, dv). \tag{7}$$

Такие движения существуют при всяком выборе $\mu(dt, du)$, $\nu(dt, dv)$ (см., например, ([2], [3])).

Обозначим через $G(t, x, \vartheta;\ \nu)$ область достижимости ([1]) для движений $x(t)$ (7) к моменту $t = \vartheta$ из позиции $\{t, x\}$ при некотором выбранном $\nu$ и при всевозможных $\mu$, а через $G_{\mathscr{P}}(t, x, \vartheta;\ \nu)$ — $\mathscr{P}$-окрестность множества $G(t, x, \vartheta;\ \nu)$. (Условие $p \in \mathscr{P}$ равносильно $-p \in \mathscr{M}$.) Расстояние от области $G_{\mathscr{P}}(t, x, \vartheta;\ \nu)$ до начала координат $x = 0$ обозначим через $\varepsilon = \varepsilon(t, x;\ \nu)$, т. е.

$$\varepsilon = \min_{x_\nu(\vartheta)} \rho(x_\nu(\vartheta), \mathscr{M}), \tag{8}$$

где $x_\nu(\vartheta) = x(\vartheta;\ t, x, \mu, \nu)$ при выбранном $\nu$. По аналогии с ([1]), назовем величину $\varepsilon^0 = \varepsilon^0(t, x) \geqslant 0$, определяемую равенством

$$\varepsilon^0 = \max_\nu \min_{x_\nu(\vartheta)} \rho(x_\nu(\vartheta), \mathscr{M}) = \min_{x_{\nu^0}(\vartheta)} (x_{\nu^0}(\vartheta), \mathscr{M}), \tag{9}$$

гипотетическим рассогласованием. Другими словами, $\varepsilon^0$ есть расстояние наиболее удаленной от $\mathscr{M}$ области $G(t, x, \vartheta;\ \nu^0)$ до этого множества $\mathscr{M}$.

Далее будем полагать, что выполнено следующее

Условие А. При $0 < \varepsilon^0(t, x) < N$ равенство (9) обеспечивается на единственной паре программных управлений $\{\mu^0, \nu^0\}$ и точка $x^*$ из $\mathscr{M}$, ближайшая к $x_{\nu}^0(\vartheta)$, единственна.

Лемма. *При выполнении условия* А *функция* $\varepsilon^0(t, x)$ *в области, где* $0 < \varepsilon^0(t, x) < N$, $t < \vartheta$, *есть функция дифференцируемая.*

Обозначим через $U^{(e)}$ экстремальную стратегию первого игрока, которая определяется множеством $U^{(e)}(t, x)$ следующим образом. Если $\varepsilon^0(t, x) = 0$, то $U^{(e)}(t, x) = P$. Если же $\varepsilon^0(t, x) > 0$, то множество $U^{(e)}(t, x)$ складывается из всех векторов $u^{(e)}$, удовлетворяющих условию

$$s'(t) \cdot f^{(1)}(t, x, u^{(e)}) = \max_{u \in P} [s'(t) \cdot f^{(1)}(t, x, u)], \tag{10}$$

где $s(t) = S(\vartheta, t, x^0(\cdot), \mu^0, \nu^0) \cdot l^0$, $S(t, t_*, x^0(\cdot), \mu^0, \nu^0)$ — фундаментальная матрица решений

$$\delta x(t) = \delta x(t_*) + \int_{[t_*, t]}\int_P \frac{\partial f^{(1)}(\tau, x(\tau), u)}{\partial x} \cdot \delta x(\tau)\,\mu(d\tau, du) + \int_{[t_*, t]}\int_Q \frac{\partial f^{(2)}(\tau, x(\tau), v)}{\partial x} \cdot \delta x(\tau)\,\nu(d\tau, dv) \tag{11}$$

уравнения первого приближения в вариациях для уравнения (1), вычисленная на движении $x^0(\tau) = x^0(\tau;\ t,\ x,\ \mu^0,\ \nu^0)$, $l^0$ — единичный вектор внутренней нормали к сфере $\|x\| \leqslant \varepsilon^0$, касающейся области $G_{\mathscr{P}}(t,\ x,\ \vartheta;\ \nu^0)$ в точке $x^0(\vartheta)$, символ $\|x\|$ — эвклидова норма вектора $x$, штрих означает транспонирование.

Теорема 1. *При выполнении условия* А *экстремальная стратегия первого игрока* $U^{(e)}(t,\ x)$, *определяемая соотношением* (10), *является оптимальной минимаксной стратегией* $U^0$, *которая разрешает задачу* 1, *причем*

$$\max_{x[t]} \rho\,(x_{U^0}[\vartheta],\ \mathscr{M}) = \varepsilon^0\,(t,\ x),$$

*какова бы ни была исходная позиция* $\{t_0,\ x_0\}$ *из области* $\varepsilon^0(t_0,\ x_0) < N$.

Теорема 2. *Пусть выполнено условие* А *и начальная позиция игры* $\{t_0,\ x_0\}$ *такова, что в некоторый момент времени* $\vartheta = \vartheta^0$ *(который, следуя* ([1]), *назовем моментом программного поглощения) точка* $x = 0$ *принадлежит области* $G_{\mathscr{P}}(t_0,\ x_0,\ \vartheta^0;\ \nu^0)$.

*Тогда экстремальная стратегия* $U^{(e)}(t,\ x)$, *определяемая соотношением* (10), *гарантирует выведение точки* $x[t]$ *на множество* $\mathscr{M}$ *к моменту* $\vartheta^0$ *при любой допустимой реализации* $v[t]$, $t_0 \leqslant t \leqslant \vartheta^0$, *управления* $v$.

Доказательство леммы проводится по следующей схеме.

а) Множество мер $\mu(dt,\ du)$ $(\nu(dt,\ dv))$, заданное на $T \times P$ $(T \times Q)$, слабо компактно в себе. При этом, если $\mu^{(k)} \xrightarrow{\text{сл.}} \mu$, $\nu^{(k)} \xrightarrow{\text{сл.}} \nu$, то соответствующие решения $x^{(k)}(t)$ (7) сходятся равномерно на $[t_0,\ \vartheta]$ к решению $x(t) = x(t;\ t_0,\ x_0\ \mu,\ \nu)$.

б) Фундаментальная матрица $S(t,\ t_*,\ x(\cdot),\ \mu,\ \nu)$ удовлетворяет следующим двум условиям. Для всякого $\alpha > 0$ найдется $\beta > 0$ такое, что

1°) если $\|x(\cdot\ /\ t_*,\ x,\ \mu,\ \nu^0) - x^0(\cdot\ /\ t_*,\ x,\ \mu^0,\ \nu^0)\| < \beta$, то

$$\|S(t,\ t_*,\ x(\cdot),\ \mu,\ \nu^0) - S(t,\ t_*,\ x^0(\cdot),\ \mu^0,\ \nu^0)\| \leqslant \alpha;$$

2°) если $\|x^0_*(\cdot/t_*,\ x_*,\ \mu^0_*,\ \nu^0_*) - x^{0*}(\cdot/t_*,\ x^*,\ \mu^{0*},\ \nu^{0*})\| < \beta$, то

$$\|S(t,\ t_*,\ x^0_*(\cdot),\ \mu^0_*,\ \nu^0_*) - S(t,\ t_*.\ x^{0*}(\cdot),\ \mu^{0*},\ \nu^{0*})\| \leqslant \alpha$$

при всех $t \in [t_0,\ \vartheta]$ и выбор $\beta$ по $\alpha$ равномерен для всех $x$ из ограниченной замкнутой области в $E^n$, где выполнено условие А. Символ $\|S(t,\ t_*,\ x(\cdot),\ \mu,\ \nu)\|$ означает норму матрицы $S(t,\ t_*,\ x(\cdot),\ \mu,\ \nu)$, соответствующую $\|x\|$.

в) Для всякого $\beta > 0$ найдется $\delta > 0$ такое, что при $\|x(\vartheta) - x^0(\vartheta)\| < \delta$ имеет место $\|x(\cdot) - x^0(\cdot)\| < \beta$ и выбор $\delta$ по $\beta$ равномерен для всех $x$ из ограниченной замкнутой области в $E^n$, где выполнено условие А.

г) Обозначим через $\Gamma$ замкнутую $\delta$-окрестность точки $x^0(\vartheta)$, принадлежащей области $G_{\mathscr{P}}(t,\ x,\ \vartheta;\ \nu^0)$, а через $G_{\mathscr{P}}(t,\ x + \Delta x,\ \vartheta;\ \nu_x{}^0)$ — область достижимости к моменту $t = \vartheta$ из позиции $\{t,\ x + \Delta x\}$ при оптимальном управлении $\nu_x{}^0 = \nu^0$, соответствующем позиции $\{t,\ x\}$. При изменении позиции $\{t,\ x\}$ на $\{t,\ x + \Delta x\}$ точки $x(\vartheta) \in \Gamma$ перейдут в точки $x_*(\vartheta)$ из некоторой области $\Gamma_*$. Справедливо равенство

$$\Delta x_*(\vartheta) = x_*(\vartheta) - x(\vartheta) = S(\vartheta,\ t,\ x(\cdot),\ \mu,\ \nu^0)\Delta x(t) + o(\Delta x). \quad (12)$$

Из (12) вытекает, что с точностью до $\alpha(\delta)\Delta x$ область $\Gamma$ при изменении $x$ перемещается поступательно, причем $\alpha(\delta) \to 0$ при $\delta \to 0$.

д) При достаточно малых $\Delta x$ точка $x^0_{x+\Delta x}(\vartheta) \in G_{\mathscr{P}}(t,\ x + \Delta x,\ \vartheta;\ \nu_x{}^0)$, ближайшая к $x = 0$, лежит в $\Gamma_*$. Это утверждение вытекает из единственности точки $x^0(\vartheta)$.

е) Обозначим расстояние от $x = 0$ до области $G_{\mathscr{P}}(t,\ x + \Delta x,\ \vartheta;\ \nu_x{}^0)$ через $\varepsilon_*(t,\ x + \Delta x)$. Из г), д) следует справедливость равенства

$$\varepsilon_*(t,\ x+\Delta x)-\varepsilon^0(t,\ x)=-l^{0\prime}(t,\ x)\Delta x^0(\vartheta)+o(\Delta x). \tag{13}$$

Из б), непрерывной зависимости $l^0(t,x)$ от позиции $\{t,x\}$ ([1]), неравенства $\varepsilon^0(t,x+\Delta x)\geqslant\varepsilon_*(t,x+\Delta x)$ вытекает равенство

$$\varepsilon^0(t,\ x+\Delta x)=\varepsilon_*(t,\ x+\Delta x)+o(\Delta x). \tag{14}$$

Из (13), (14) следует, что при изменении $x$ изменение $\varepsilon^0(t,\ x)$ изображается соотношением

$$\Delta\varepsilon^0(t,\ x)=-l^{0\prime}(t,\ x)\cdot S(\vartheta,\ t,\ x^0(\cdot),\ \mu^0,\ \nu^0)\cdot\Delta x(t)+o(\Delta x). \tag{15}$$

ж) Из того факта, что $x^0(\vartheta)$ в области $G_{\mathscr{P}}(t,\ x,\ \vartheta;\ \nu^0)$ является ближайшей к точке $x=0$, вытекает утверждение: мера $\mu^0(dt,\ du)$ сосредоточена на множестве тех $\{\tau,\ u\}$ $(t_0\leqslant\tau\leqslant\vartheta,\ u\in P)$, для которых выполняется условие принципа максимума ([4])

$$\begin{gathered} l^{0\prime}(t,x)\cdot S(\vartheta,\tau,x^0(\cdot),\mu^0,\nu^0)\cdot f^{(1)}(\tau,x(\tau),u^0)= \\ =\max_{u\in P}\,[l^{0\prime}(\tau,x)\cdot S(\vartheta,\tau,x^0(\cdot),\mu^0,\nu^0)\cdot f^{(1)}(\tau,x(\tau),u)]. \end{gathered} \tag{16}$$

з) Поскольку управление $\nu^0$ является компонентой единственного максиминного управления $\{\mu^0,\ \nu^0\}$, мера $\nu^0(dt,\ dv)$ сосредоточена на множестве тех $\{\tau,\ v\}$, $t_0\leqslant\tau\leqslant\vartheta$, $v\in Q$, для которых выполняется условие минимума

$$\begin{gathered} l^{0\prime}(t,x)\cdot S(\vartheta,\tau,x^0(\cdot),\mu^0,\nu^0)\cdot f^{(2)}(\tau,x(\tau),v^0)= \\ =\min_{v\in Q}\,[l^{0\prime}(\tau,x)\cdot S(\vartheta,\tau,x^0(\cdot),\mu^0,\nu^0)\cdot f^{(2)}(\tau,x(\tau),v)]. \end{gathered} \tag{17}$$

Из а) — з) следуют выражения для частных производных $\partial\varepsilon^0/\partial t$ и $\partial\varepsilon^0/\partial x$ (в позиции $\{t,x\}$):

$$\partial\varepsilon^0/\partial t=l^{0\prime}(t,\ x)\cdot S(\vartheta,\ t,\ x^0(\cdot),\ \mu^0,\ \nu^0)\cdot(f^{(1)}(t,\ x)-f^{(2)}(t,\ x)), \tag{18}$$

$$\partial\varepsilon^0/\partial x=\{\partial\varepsilon^0/\partial x_i\}=-l^{0\prime}(t,\ x)\cdot S(\vartheta,\ t,\ x^0(\cdot),\ \mu^0,\ \nu^0), \tag{19}$$

где $f^{(1)}(t,\ x)\in F^{(1)}(t,\ x)$, $f^{(2)}(t,\ x)\in F^{(2)}(t,\ x)$ удовлетворяют соответственно условиям максимума и минимума (при $\tau=t$).

Исходя из свойства дифференцируемости функции $\varepsilon^0(t,\ x)$, доказываются аналогично ([1]) теоремы 1 и 2.

Автор выражает благодарность акад. Н. Н. Красовскому за обсуждение работы и ценные советы.

Институт математики и механики
Уральского научного центра Академии наук СССР
Свердловск

Поступило
28 IX 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Н. Н. Красовский, Игровые задачи о встрече движения, «Наука», 1970. [2] А. Д. Иоффе, Дифференциальные уравнения, 5, № 6, 1010 (1969). [3] А. И. Сотсков, Кибернетика, № 4, 110 (1969). [4] Л. С. Понтрягин, В. Г. Болтянский и др., Математическая теория оптимальных процессов, М., 1961.