Управляемые системы, ИМ,ИК СО АН СССР I97I г. выпуск 9.
УДК.519.3.

О ДИФФЕРЕНЦИАЛЬНЫХ ИГРАХ КАЧЕСТВА С НЕСКОЛЬКИМИ ТЕРМИНАЛЬНЫМИ ПОВЕРХНОСТЯМИ

Е.П.Волокитин

В подавляющем большинстве работ, посвященных решению дифференциальных игр качества двух игроков, рассматривается традиционная постановка задачи, которая была предложена Р.Айзексом в [I] и состоит в следующем: игра протекает в некоторой области $Z \subset R_n$, ограниченной кусочно-гладкими поверхностями, и описывается уравнениями движения.

$$\dot{z} = f(z, \varphi, \psi), \qquad /I/$$

где $z \in Z$ - состояние игры, $\varphi \in \Phi$ - уравнение одного из игроков (P); $\psi \in \Psi$ - управление его противника (E) ./Φ и Ψ - выпуклые компактные множества в R_{m_1} и R_{m_2}, соответственно, на функции f, φ, ψ накладываются необходимые ограничения: обычно f предполагают непрерывно дифференцируемой по всем ее аргументам, а φ и ψ считают кусочно-непрерывными функциями/. Имеется некоторая поверхность Ω, называемая терминальной поверхностью и представляющая собой часть границы пространства игры Z. Игра считается оконченной, если z достигает Z. Предполагается, что один игрок /обычно это P / хочет окончить игру, а его противник препятствует этому. Выяснение возможных исходов этой конфликтной ситуации в зависимости от начального состояния игры и поведения обоих игроков и составляет решение игры качества. В рассматриваемой проблеме предполагается наличие полной информации в смысле [I]: считается, что обоим игрокам всегда известны текущие значения всех фазовых координат.

Однако иногда при рассмотрении конфликтных ситуаций возникает необходимость в решении дифференциальных игр качества, формулировка которых несколько отличается от приведенной выше, а именно: игра, как и раньше, протекает в области $Z \subset R_n$ - пространстве игры и описывается системой /I/, однако терминальная поверхность Ω, представляющая собой часть границы Z, разбивается на части Ω_P, Ω_E, Ω_D, игра по-прежнему считается оконченной в момент времени $\tilde{t} < +\infty$, если $z(\tilde{t}) \in \Omega$, при этом выигравшим считается игрок P, если $z(\tilde{t}) \in \Omega_P$, и игрок E, если $z(\tilde{t}) \in \Omega_E$, наконец, окончание траектории $z(t)$ на Ω_D означает ничейный исход. Каждый из обоих игроков теперь, по-прежнему располагая полной информацией, стремится окончить игру, приведя точку $z(t)$ на соответствующую часть терминальной поверхности Ω.

Требуется найти множества точек, начиная из которых каждый игрок при соответсвующем поведении в состоянии добиться желаемого для него исхода независимо от поведения противника.

Основная цель настоящей заметки, публикуемой, главным образом, в порядке постановки вопроса, - привлечь внимание к рассмотрению подобного рода задач и нахождению методов их решения.

Поверхность, разделяющая области выигрыша игроков /если она существует/ будем называть, придерживаясь терминологии [I] , барьером. Очевидно, этот барьер, будет полупроницаемой поверхностью. Поэтому иногда рассматриваемые задачи удается решить с помощью построения таких поверхностей, построение которых, впрочем, не всегда удается осуществить методами, предположенными в [I] , /см. [3] /.

Однако даже после построения барьера, как правило, остается открытым вопрос о нахождении стратегий, дающих игроку возможность добиться желаемого для него исхода независимо от действий противника. Зачастую здесь приходится довольствоваться "методом явной политики", т.е. нахождением и изучением конкретной стратегии игрока, приводящей его к цели. Недостатки такого подхода очевидны.

Метод желательных направлений [2] , предназначенный для решения подобных задач, к сожалению, пока еще мало разработан: даже при рассмотрении некоторых простых примеров он наталкивается на трудности аналитического характера, не поддающиеся качественному исследованию.

Некоторые из подобного рода задач удается решить, если рассматривать их как игры качества двух игроков с одной терминальной поверхностью, связывая с рассматриваемой игрой качества, как это предлагается в [I] , игру степени, платой в которой служит время окончания игры. К числу примеров можно отнести задачу о конфликтном взаимодействии двух экономических моделей, решаемую в качестве примера в [2] . Неправомерность такого подхода в общем случае становится понятной, если заметить, что, сводя задачу к традиционной постановке /т.е. отбрасывая часть терминальной поверхности, например $\Omega_E \cup \Omega_D$ /, и решая игру степени, соответствующую полученной игре качества, мы должны, вообще говоря, следить еще и за выполнением фазовых ограничений; ясно, что игнорирование этих ограничений может весьма существенно изменить характер искомого решения.

Ниже рассмотрены два примера, на которых иллюстрируются некоторые моменты сделанных замечаний.

П р и м е р I. Пусть каждый из игроков P и E обладает на плоскости Oxy простым движением, т.е. уравнения движения имеют вид:

$$\dot{x} = u\cos\varphi + v\cos\psi, \qquad /2/$$
$$\dot{y} = u\sin\varphi + v\sin\psi.$$

Пространством игры является неотрицательный квадрант фазовой плоскости, R_2^+. Игрок P, в распоряжении которого находится направление φ $(0 \le \varphi \le 2\pi)$ скорости, по величине равной u, стремится привести фазовую точку на положительную часть оси Ox, цель игрока E, управляющего направлением ψ $(0 \le \psi \le 2\pi)$ скорости, по величине равной v, - привести фазовую точку на положительную часть оси Oy, $u > v$, окончание траектории в начале координат означает ничейный исход партии в данной игре. Требуется решить сформулированную игру качества.

Полупроницаемые кривые в данном случае провести не удается по той причине, что их не существует вовсе /любая гладкая кривая не будет полупроницаемой, так как $u > v$ /.

Если мы попытаемся решить нашу игру качества, сопоставляя ей игру степени, в которой в качестве платы берется время окончания игры, и считая положительную часть оси Ox единственной терминальной поверхностью, то выясняется, что игрок P в состоянии выиграть партию, начинающуюся в любой внутренней точке пространства R_2^+, а оптимальные в смысле выбранного критерия траектории имеют вид прямых перпендикулярных оси Ox /рис.I/. При этом игрок P должен придерживаться стратегии $\bar\varphi = \frac{3}{2}\pi$ $(\cos\bar\varphi = 0, \sin\bar\varphi = -1)$, игрок E - стратегии $\bar\psi = \frac{\pi}{2}$ $(\cos\bar\psi = 0,\ \sin\bar\psi = 1)$.

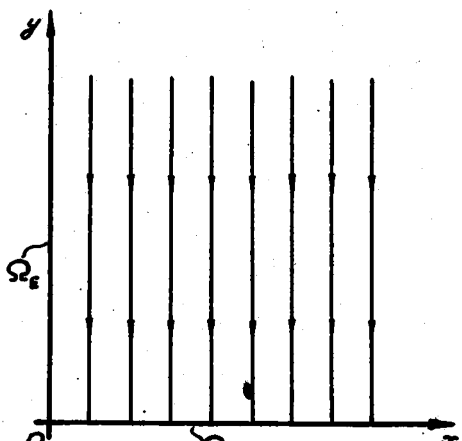

Рис. I.

Однако, очевидно, что если фазовая точка находится достаточно близко к оси Oy, а игрок P слепо придерживается указанной "оптимальной" стратегии, то E в состоянии помешать противнику добиться выигрыша, придерживаясь какой-либо стратегии ψ, отличной от $\bar\psi$. Точнее, можно показать, что стратегия $\bar\varphi = \frac{3}{2}\pi$ является эффективной с точки зрения игрока P лишь в той части пространства игры, где $y < \sqrt{\frac{u^2}{v^2} - 1}\,x$, если же партия начинается вне этой области, т.е. $y_0 > \sqrt{\frac{u^2}{v^2} - 1}\,x_0$ /(x_0, y_0) - начальное состояние игры/, а P избирает стратегию $\bar\varphi = \frac{3}{2}\pi$, то его противник может помешать ему привести траекторию на Ω_P, для чего E достаточно придерживаться, например стратегии $\tilde\psi$,
$$\cos\tilde\psi = -\sqrt{1 - \frac{v^2}{u^2}},\ \sin\tilde\varphi = \frac{u}{v}.$$

С другой стороны, ясно, что P, на самом деле, в силу своего превосходства в скорости в состоянии одержать победу в любой партии, начинающейся из внутренней точки пространства R_2^+. Проблема заключается лишь в указании соответствующего способа поведения P.

Рассматриваемая задача может быть решена, если применить метод желательных направлений в том виде, в каком он изложен в [2]. Однако решение выглядит проще, если мы воспользуемся следующим утверждением, справедливость которого очевидна: если в рассматриваемой игре траектория $z(t)$ такова, что $\dot{z}(t)$ ни в один из моментов времени не направлена ни на $\mathcal{P}_E$, ни на $\mathcal{P}_D$, то партия, описываемая этой траекторией, не может окончиться победой игрока E или вничью. /Использованы обозначения, введение в начале статьи/.

Применим это утверждение к нашему случаю. Заметим, что если игрок P выбирает какое-либо направление φ своего движения, то все возможные направления траекторий, получающихся в результате выбора игроком E своего управления $0 \leq \psi \leq 2\pi$, отклоняются от направления φ не более чем на угол γ_0 ($\cos\gamma_0 = \sqrt{1-\frac{v^2}{u^2}}$, $\sin\gamma_0 = \frac{u}{v}$), $0<\gamma_0<\frac{\pi}{2}$.

Пусть теперь игрок $\hat{P}$ в каждый момент времени выбирает направление своего движения, перпендикулярное радиус-вектору состояния игры в этот момент времени и направленное к оси Ox, т.е. если радиус-вектор состояния игры составляет с осью Ox угол α, то игрок P выбирает направление $\hat{\varphi}=\alpha-\frac{\pi}{2}$. Тогда, в силу вышесделанного замечания, условия утверждения будут выполнены, и, следовательно, нам достаточно показать, что при таком выборе управления игроком P любая партия оканчивается за конечное время. Сделаем это.

В полярных координатах r, α ($x=r\cos\alpha$, $y=r\sin\alpha$) уравнения движения /2/ при $\varphi=\hat{\varphi}$ имеют вид:

$$\dot{r} = v\cos(\alpha-\psi),$$
$$\dot{\alpha} = -\frac{1}{r}\left[u+v\sin(\alpha-\psi)\right]<0. \qquad /3/$$

Нам надо показать, что угол $\alpha(t)$ за конечное время станет равным нулю. Из /3/ следует, что $r(t)$ ограничен:

$$r(t) \leq r_0 e^{\frac{v}{\sqrt{u^2-v^2}}\alpha_0}.$$

Но тогда

$$\dot{\alpha} \leq -\frac{1}{r_0 e^{\frac{v}{\sqrt{u^2-v^2}}\alpha_0}}\left[u+v\sin(\alpha-\psi)\right] \leq -\frac{u-v}{r_0 e^{\frac{v}{\sqrt{u^2-v^2}}\alpha_0}}.$$

Таким образом, при любых начальных состояниях игры $\alpha(t)$ убывает с ненулевой скоростью, а значит, за конечное время

$$\tilde{t} \leq \frac{\alpha_0 r_0 e^{\frac{v}{\sqrt{u^2-v^2}}\alpha_0}}{u-v}$$

обратится в нуль, что и требовалось показать.

Заметим, что, сводя игру качества к указанной игре степени, мы получили формально правильный ответ на вопрос о том, какова область выигрыша игрока P. Однако, если бы пространство игры имело, например, вид тупого угла /Рис. 2/, подобное обстоятельство не имело бы места, в то время, как метод, с помощью которого была найдена выигрышная стратегия игрока P, оказался бы применимым и в этом случае.

П р и м е р 2. Уравнения движения имеют вид:

$$\dot{x} = u\cos\varphi + yv\cos\psi, \qquad /4/$$
$$\dot{y} = u\sin\varphi + yv\sin\psi,$$

$0 \leq \varphi,\ \varphi \leq 2\pi,\ u > v$, т.е. игроки P и E по-прежнему обладают простым движением, но теперь величина скорости игрока E зависит от состояния игры. Пространством игры вновь является неотрицательный квадрант фазовой плоскости Oxy, условия окончания игры те же, что в предыдущем примере.

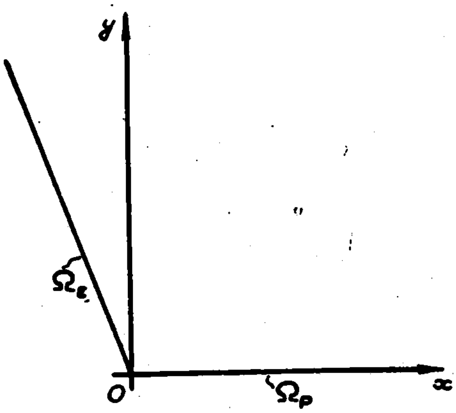

Рис. 2.

В данном случае существует единственная полупроницаемая поверхность, прямая $y = \frac{u}{v}$, на которой скорости обоих игроков равны по величине. Она, очевидно, будет статическим барьером: партии начинающейся из точек, расположенных над этой прямой, выигрывает игрок E, начинающиеся из точек под ней - игрок P. Таким образом, остается найти выигрышные стратегии игроков.

Оказывается, что попытка найти эти стратегии, рассматривая предложенную игру как игру степени с временем окончания игры в качестве платы так же, как и в предыдущем случае, не дает нужного результата: получаемые "оптимальные" стратегии могут оказаться даже неприемлемыми при некоторых начальных состояниях игры.

Аналогично тому, как это сделано выше, можно показать, что игрок E за конечное время может одержать победу над своим противником, если в каждой точке, расположенной выше барьера, он будет выбирать направление своего движения, перпендикулярное прямой, проходящей через эту точку и точку $(0, \frac{u}{v})$, и направленное в сторону оси Oy.

Чтобы найти стратегию игрока P, приводящую его к выигрышу в партиях, начинающихся из точек, расположенных ниже барьера, поступим следующим образом.

Рассмотрим функцию $F(x,y) = x(\frac{u}{v} - y)$. Линии уровня этой функции $x(\frac{u}{v} - y) = C$ суть гиперболы с асимптотами $x = 0$, $y = \frac{u}{v}$, сплошь заполняющие пространство игры под барьером, когда C принимает все возможные значения из промежутка $[0,+\infty)$ /рис. 3, стрелкой показано направление возрастания F /.

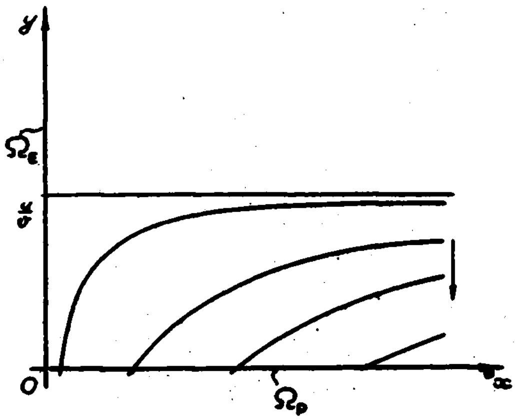

Рис. 3.

Под барьером скорость игрока P превосходит по величине скорость игрока E. Поэтому если игрок P в каждой точке пространства игры, расположенной ниже барьера, выбирает направление своего движения, совпадающее с направлением градиента ∇F функции F $(\nabla F = (-\frac{u}{v}+y,\ x))$, то проекция направления любой траектории, получающейся выбором управления ψ игроком E, на этот градиент строго положительна, т.е. функция F возрастает вдоль траекторий, получающихся при таком выборе управления φ. Это означает, что фазовая точка, описывающая состояние игры, переходит на линии уровня функции F, все более удаленные от оси Oy, что в свою очередь показывает невозможность победы игрока E. Если игрок P выбирает указанное управление $\bar{\varphi}$, то

$$\cos\bar{\varphi} = \frac{\frac{u}{v}-y}{\sqrt{x^2+(\frac{u}{v}-y)^2}}, \quad \sin\bar{\varphi} = \frac{-x}{\sqrt{x^2+(\frac{u}{v}-y)^2}},$$

и уравнения /4/ принимают вид

$$\dot{x} = u\frac{\frac{u}{v}-y}{\sqrt{x^2+(\frac{u}{v}-y)^2}} + yv\cos\psi, \qquad /4'/$$

$$\dot{y} = -u\frac{x}{\sqrt{x^2+(\frac{u}{v}-y)^2}} + yv\sin\psi .$$

Далее, при $\varphi = \bar{\varphi}$

$$\frac{d}{dt}F(x(t),y(t)) = (\frac{u}{v}-y)\left[u\frac{\frac{u}{v}-y}{\sqrt{x^2+(\frac{u}{v}-y)^2}} + yv\cos\psi\right] -$$

$$-x\left[-u\frac{x}{\sqrt{x^2+(\frac{u}{v}-y)^2}} + yv\sin\psi\right] > vx(\frac{u}{v}-y),$$

т.е.

$$\frac{d}{dt}F(x(t), y(t)) > vF(x(t), y(t)),$$

вследствие чего

$$x(t)\left(\frac{u}{v} - y(t)\right) > x_0\left(\frac{u}{v} - y_0\right)e^{v(t-t_0)} \qquad /5/$$

Рассмотрим функцию

$$f(t) = x(t) - \frac{x_0\left(\frac{u}{v} - y_0\right)}{\frac{u}{v}} e^{v(t-t_0)},$$

где $x(t)$ подчиняется системе /4'/. В силу /4'/,

$$\dot f(t) = \dot x(t) - v\frac{x_0\left(\frac{u}{v} - y_0\right)}{\frac{u}{v}} e^{v(t-t_0)} = u\frac{\frac{u}{v} - y}{\sqrt{x^2 + \left(\frac{u}{v} - y\right)^2}} + yv\cos\varphi -$$

$$- v\frac{x_0\left(\frac{u}{v} - y_0\right)}{\frac{u}{v}} e^{v(t-t_0)} < 2u - v\frac{x_0\left(\frac{u}{v} - y_0\right)}{\frac{u}{v}} e^{v(t-t_0)}.$$

Начиная с некоторого $\bar t$, будем иметь $\dot f(t) \leq -\varepsilon < 0$ при всех $t \geq \bar t$, и, следовательно, наступит момент времени t_1 такой, что $f(t_1) = 0$. С другой стороны, если допустить, что $y(t) > 0$ при всех t, то $\frac{u}{v} - y(t) < \frac{u}{v}$ при всех t, а тогда из неравенства /5/ следует, что при всех t

$$x(t) > \frac{x_0\left(\frac{u}{v} - y_0\right)}{\frac{u}{v}} e^{v(t-t_0)}$$

т.е. функция $f(t)$ всегда строго положительна, что неверно.

Таким образом, при выборе игроком P управления $\bar\varphi$ величина $y(t)$ за конечное время обратиться в ноль, что и требовалось показать.

В заключение сделаем следующее замечание.

Назовем гладкую поверхность S P-проницаемой, если в каждой точке $x \in S$ среди управлений $\varphi(x)$ игрока P найдется управление $\bar\varphi(x)$, обладающее тем свойством, что при $\varphi = \bar\varphi$ ψ-вектограмма для E во всех точках поверхности S лежит по одну и ту же ее сторону /аналогично вводится определение E -проницаемой поверхности/.

Мы видим, что в обоих разобранных примерах задача нахождения выигрышных стратегий, по сути дела, сводилась к отысканию P - или E -проницаемых в нужном направлении поверхностей и соответствующих управлений игроков.

Поступила в редакцию 7.6.1971 г.

Литература

1. Р.Айзекс. Дифференциальные игры. М., "Мир", 1967.

2. Е.П.Волокитин, И.А.Красс. Один метод исследования дифференциальных игр качества.- Управляемые системы. Новосибирск, 1970, Вып. 6, стр. 3-16.

3. И.А.Красс, А.М.Мухсинов. Некоторые методы исследования игр качества с многими терминальными поверхностями.- Управляемые системы, в печати.