Теорема об альтернативе

В основу исследований по теории дифференциальных игр в книге „Позиционные дифференциальные игры" положена концепция максимальных стабильных мостов.

В дифференциальных играх могут быть различные задачи с разными показателями оптимальности. Говоря о максимальных стабильных мостах, мы выделяем только один, казалось бы очень частный вид дифференциальных игр, в котором первый игрок заинтересован в приведении фазового вектора на заданное терминальное (целевое) множество, а второй препятствует этому. Оказалось, однако, что занимаясь дифференциальной игрой весьма произвольного вида, мы можем, правильно задавая некоторую систему множеств и объявляя их целевыми, свести изучение нашей дифференциальной игры к рассмотрению системы максимальных стабильных мостов, каждый из которых подсоединён к своему целевому множеству из взятой нами системы.

Т.е. конструкции, связанные с максимальными стабильными мостами, — это ядро в исследовании дифференциальных игр.

Опишем общую идею.

Теорема об альтернативе указывает возможности игроков в задаче с терминальным множеством

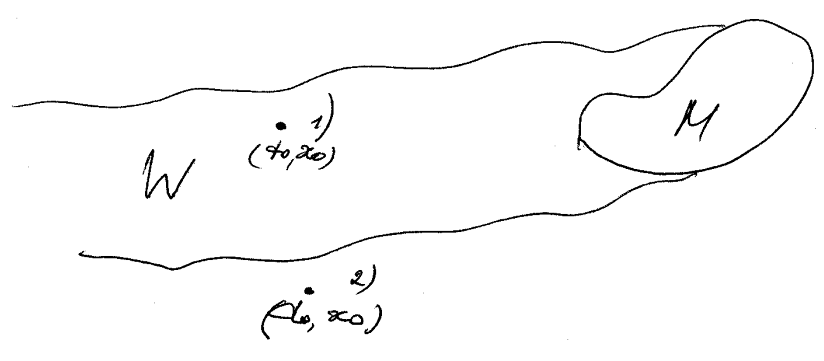
М, если мы зафиксировали некоторую начальную позицию (t_0, x_0). А именно, теорема говорит о том, что если зафиксирована начальная позиция (t_0, x_0), то могут быть только два варианта:

1) первый игрок гарантирует перевод на М (чтобы ни делал в пределах своих возможностей второй игрок),

2) существует окрестность множества М, от которой второй игрок гарантирует уклонение.

Таким образом: либо, либо. Подчеркнём, что 2) не является логическим отрицанием 1). Отрицание звучало бы так: первый игрок не может гарантировать перевод на М.

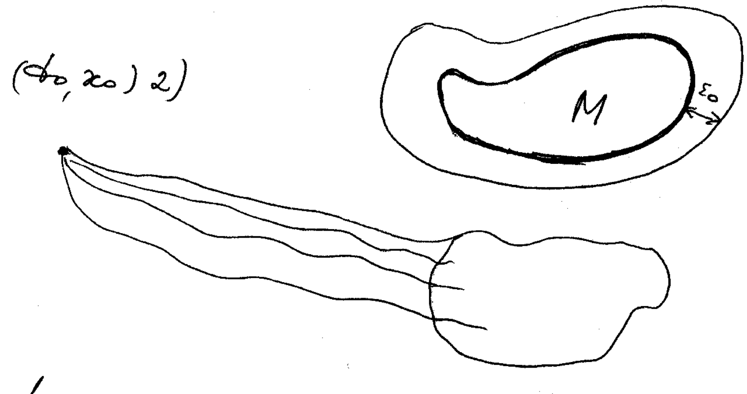
Опираясь на наши предыдущие знания, мы можем догадаться, откуда происходит эта теорема. А именно, если задано множество М, то подсоединяем к нему максимальный стабильный мост W.



Логически возможны только два варианта:

1) $(t_0, x_0) \in W$, 2) $(t_0, x_0) \notin W$. В первом случае, используя экстремальную к W стратегию V первый игрок приводит на M. А во втором?

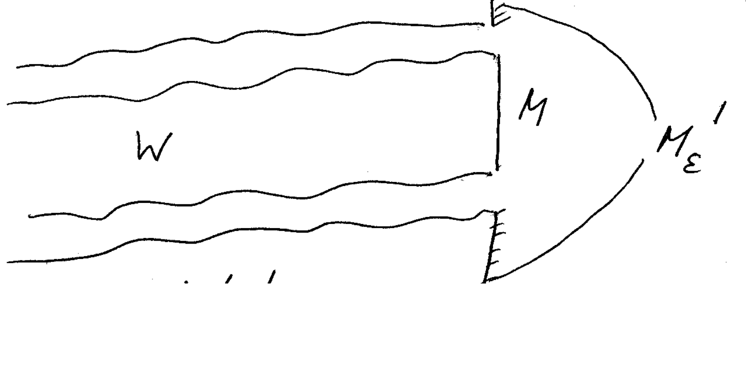
Теорема об альтернативе как раз и говорит о том, что во втором случае существует некоторая ε_0-окрестность M, от которой уклоняется второй игрок. Т.е. у второго игрока есть способ управления V, зависящий от (t_0, x_0), такой что, используя его при малом шаге Δ дискретной схемы второй игрок уводит весь пучок движений от попадания на ε_0-окрестность множества M.



Можем догадаться, как возникает вариант 2). Множество M — замкнутое. Навесив ε-окрестность, получаем открытое множество $M + B(\varepsilon)$, где $B(\varepsilon)$ — открытый шар радиуса ε. Но тогда дополнение нашего пространства до $M + B(\varepsilon)$ — замкнутое множество. Обозначим его M'_ε. Будем считать, что второй игрок стремится привести систему на M'_ε, а первый препятствует этому. Другими словами, просто переобозначим игроков.

Вообразим максимальный стабильный (лучше сказать, v-стабильный) мост W', корсоединенный в такой задаче к M'_ε. Если он захватил позицию (t_0, x_0), то второй игрок при помощи своей экстремальной к W'_ε стратегии приведёт движение на M'_ε. На это можно надеяться, если ε мало (по сравнению с расстоянием от (t_0, x_0) до W). При малом ε мост W'_ε захватит позицию (t_0, x_0).

Именно такая идея лежит в доказательстве теоремы об альтернативе. При этом динамика $\dot{x} = f(t, x, u, v)$ может быть весьма произвольной (но удовлетворяющей условию наличия седловой точки в маленькой игре), множество M может быть произвольным замкнутым множеством в пространстве $\{t, x\}$. Но если момент окончания T зафиксирован и множество M лежит в гиперплоскости $\{t = T, x\}$ (т.е. оно в пространстве x, но попадание на него фиксируется только в момент T), то всё намного нагляднее:



Если условие седловой точки в маленькой игре не выполнено, то теория становится более сложной, хотя и во многом аналогичной.

Теорему об альтернативе (а она важна и сама по себе) применяют для исследования дифференциальных игр с заданной функцией платы. Продемонстрируем это на задаче с фиксированным моментом окончания T и терминальной платой $\varphi(x(T))$, которую минимизирует первый игрок и максимизирует второй.

Функция цены в задаче с фиксированным моментом окончания и терминальной платой

Пусть

$$\dot{x} = f(t, x, u, v), \quad x \in R^n, \quad u \in P, \ v \in Q.$$

Предположим, что выполнены стандартные предположения о функции f, например, глобальное условие Липшица по x и непрерывность по t, u, v. Дополнительно считаем выполненным условие седловой точки в маленькой игре (условие Айзекса).

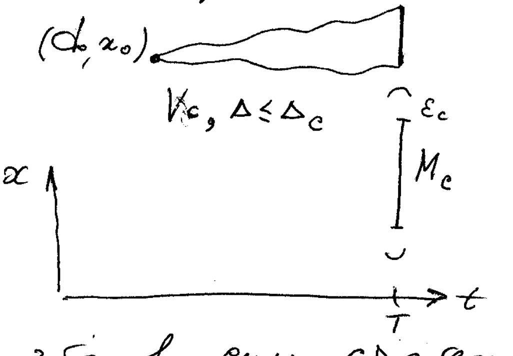
Рассмотрим задачу с фиксированным моментом окончания T и непрерывной функцией платы $\varphi(x)$, подсчитываемой в момент окончания. Первый игрок минимизирует значение $\varphi(x(T))$, второй максимизирует.

Покажем, что наилучшие гарантированные результаты игроков совпадают для любой начальной позиции (t_0, x_0): $\Gamma^{(1)}(t_0, x_0) = \Gamma^{(2)}(t_0, x_0)$. Установим также, как можно конструировать оптимальные стратегии U^0 и V^0 при заданной начальной точке.

Символом M_c обозначим множество Лебега $\{x : \varphi(x) \le c\}$ функции платы.

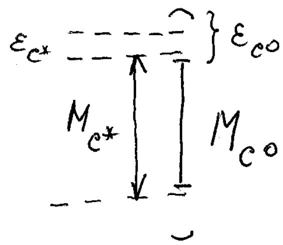
① Мысленно зафиксируем позицию (t_0, x_0).

Через c^0 обозначим верхнюю грань тех значений c, для которых второй игрок уклоняет из (t_0, x_0) движения от попадания на M_c в момент T, т.е. верхнюю грань тех значений c, для которых выполнено второе свойство альтернативы.



Возьмём $c < c^0$. Тогда из (t_0, x_0) есть стратегия V_c второго игрока, ограничение на шаг Δ_c дискретной схемы, размер ε_c окрестности множества M_c такие, что в силу стратегии V_c с шагом $\Delta \le \Delta_c$ дискретной схемы управления пучок $X^{(2)}(T; t_0, x_0, V_c, \Delta)$, полученный при переборе всех $u(\cdot)$, доведённых до момента T, не пересекается с множеством $M_c + B(\varepsilon_c)$.

Для предельного значения c^0 такое уклонение уже невозможно. От противного, пусть оно есть. Но тогда можно взять $c^* > c^0$ (именно, $c^* > c^0$!), близкое к c^0, что при V_{c^0}, $\Delta \le \Delta_{c^0}$ уклонение будет от некоторой окрестности $\varepsilon_{c^*} \le \varepsilon_{c^0}$ (ε_{c^*} близко к ε_{c^0}) множества $M_{c^*} + B(\varepsilon_{c^*})$:



Приходим к противоречию с определением c^0.

Обращаясь к теореме об альтернативе, заключаем, что для начальной позиции (t_0, x_0) при стабильном множестве M_{c^0} второе свойство альтернативы не выполнено. Следовательно, имеет место первое: $\exists$ стратегия U_{c^0} первого игрока такая, что при $\forall \varepsilon$ $\exists \Delta(\varepsilon)$: весь пучок $X^{(1)}(T; t_0, x_0, U_{c^0}, \Delta)$, доведённый до момента T, содержится в ε-окрестности множества M_{c^0}.

Более того, в качестве стратегии U_{c^0} можно взять стратегию, экстремальную к максимальному стабильному мосту $W^{(1)}_{c^0}$, обрывающемуся в момент T на множестве M_{c^0}.

Итак, $\Gamma^{(1)}(t_0, x_0) \le c^0$.

Поскольку при $c < c^0$ есть уклонение второго игрока от M_c, то получаем, что

$$\Gamma^{(1)}(t_0, x_0) = c^0 \qquad (1)$$

② Теперь через c_0 обозначим нижнюю грань тех значений c, для которых первый игрок уклоняет систему от попадания на $cl\, M_c'$ в момент T. Здесь $M_c' = R^n \setminus M_c$, а cl - символ замыкания.

Фактически, начинаем повторять рассуждения из ①, только теперь первый игрок уклоняет от дополняющих множеств [не забыли их замкнуть!]. Получаем, что $\exists$ стратегия V_{c_0} второго игрока, которая приводит на $cl\, M_{c_0}'$. В качестве такой стратегии можно взять экстремальную стратегию к максимальному стабильному мосту $W_{c_0}^{(2)}$ за второго игрока, нерасширенному к $cl\, M_{c_0}'$.

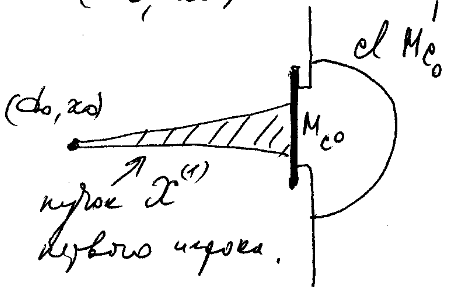
Имеем $\Gamma^{(2)}(t_0, x_0) \geq c_0$.

Поскольку при $c > c_0$ есть уклонение первым игроком от $cl\, M_c'$, то получаем

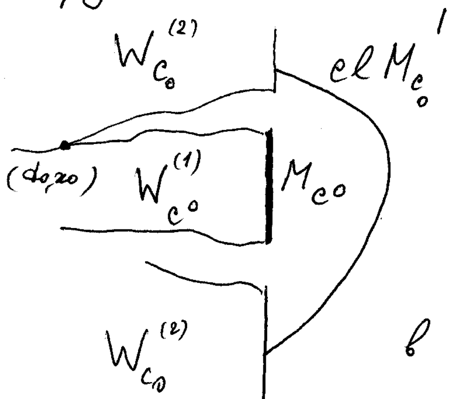
$$\Gamma^{(2)}(t_0, x_0) = c_0. \qquad (2)$$

③ Как соотносятся c^0 и c_0?

Предположим, что $c_0 < c^0$. Это невозможно, поскольку в силу определения c_0 первый уклоняется от $cl\, M_c'$ при $c > c_0$, т.е. приводит весь пучок на M_c. Получаем противоречие с тем, что c^0 есть минимальное c, при котором первый игрок приводит из (t_0, x_0) на M_c (что следует из определения c^0).



Предположим, что $c_0 > c^0$. Это невозможно, ибо здесь по отношению к M_{c^0} выполнено в альтернативе и 1), и 2).



Таким образом $c^0 = c_0$ и значит, в силу формул (1), (2) получаем $\Gamma^{(1)}(t_0, x_0) = \Gamma^{(2)}(t_0, x_0)$. При этом оптимальные результаты реализуются на экстремальных стратегиях U^0 к множеству $W^{(1)}_{c^0}$ и V^0 к множеству $W^{(2)}_{c^0}$. Говорят, что на U^0, V^0 реализуется седловая точка.

Подчеркнём, что стратегии U^0, V^0 зависят от начальной позиции (t_0, x_0). Т.е. нужно помнить, что в обозначения $U^0(t, x)$, $V^0(t, x)$ следовало бы вписать какое-либо напоминание об этом.

Универсальные стратегии

Стратегия V^0 называется универсальной оптимальной по отношению к некоторому множеству F начальных позиций, если она годится как оптимальная для из каждой начальной позиции $(t_0, x_0) \in F$.

Аналогично за второго игрока. Наличие универсальных оптимальных стратегий — большая редкость. Но они есть. Рассмотрим такой случай.

Универсальная оптимальная стратегия первого игрока в игре с фиксированным моментом окончания, линейной динамикой, выпуклой функцией платы. Управление первого игрока скалярное.

(1) Пусть

$$\dot{x} = A(t)x + B(t)u + C(t)v \qquad (3)$$

$$u \in P,\ v \in Q,\ x \in R^n,\ T.$$

Предположим, что P — отрезок, а функция платы φ выпуклая. Пусть она зависит от m координат в момент окончания и уходит в ∞ при $|x| \to \infty$.

Переобозначением в $B(t)u$, $u \in P$ можно добиться, что $u \in R^1$. Кроме того, условимся (чтобы не терять общность), что P — это отрезок с концами

$-\mu$, $+\mu$, т.е. $|u| \le \mu$.

Переходим к эквивалентной системе без фазовых переменных в правой части:

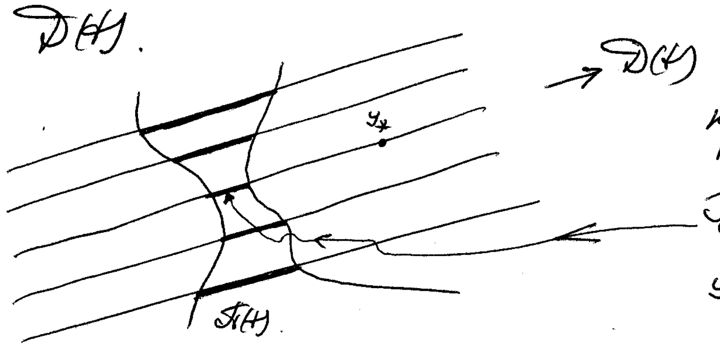
$$\dot y = D(t)\,u + E(t)\,v \qquad (4)$$

$$u \in P,\ v \in Q,\ T,\ \varphi$$

Имеем $y(T) = x_m(T)$.

Мы знаем, что максимальный стабильный мост W_c (мост за первого игрока), построенный от выпуклого целевого множества $M_c = \{y : \varphi(y) \le c\}$ [Мы сохраним для функции цены в (4) такое же обозначение, что и для цены в (3)] имеет выпуклые t-сечения $W_c(t)$. Это означает, что функция цены $\Gamma(t, y)$ при фиксированном t является выпуклой. [Формально, это влечёт только квазивыпуклость. Надо обратиться к неравенствам].

Имея направление $D(t)$, рассмотрим всё пространство R^m точек y на прямые, параллельные $D(t)$.



Через каждую точку y_* проходит прямая $E(t, y_*)$. Возьмём совокупность точек y, на которых реализуется $\min\limits_{y \in E(t,y_*)} \Gamma(t, y)$. В силу выпуклости функции $y \to \Gamma(t,y)$ такая совокупность — отрезок [точки min $\exists$, ибо $\Gamma(t,y) \to \infty$ при $|y| \to \infty$]. Полученное множество в R^m обозначим $\mathcal{M}(t)$.

Возьмём поверхность $\Pi(t) \subset \mathcal{M}(t)$, которая делит R^m на три части: $\Pi(t)$, $\Pi^-(t)$, $\Pi^+(t)$, где $\Pi^-(t)$ — часть по направлению $D(t)$, а $\Pi^+(t)$ — против направления $D(t)$. Тогда стратегия первого игрока

$$U(t,y) = \begin{cases} -\mu, & y \in \Pi^-(t), \\ +\mu, & y \in \Pi^+(t), \\ \forall u \in [-\mu, \mu], & y \in \Pi(t) \end{cases}$$

является оптимальной универсальной для первого игрока. Поверхность $\Pi(t)$ назовём поверхностью переключения для момента t. Варьируя по t, получим поверхность в пространстве t, y.

В исходных координатах x стратегия будет выглядеть следующим образом:

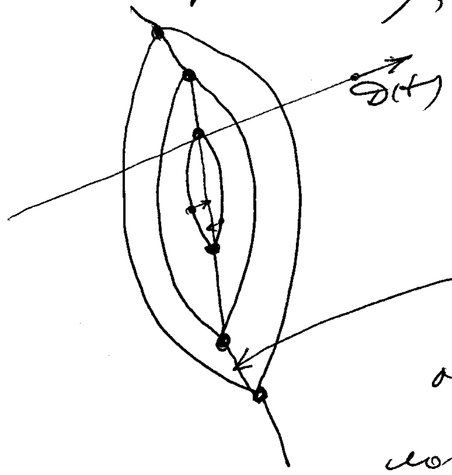
$$U(t,x) = \begin{cases} -\mu, & X_m(T,t)x \in \Pi^-(t), \\ \mu, & X_m(T,t)x \in \Pi^+(t), \\ \forall u \in [-\mu, \mu], & X_m(T,t)\,x \in \Pi(t). \end{cases}$$

Доказывается также, что указанные стратегии устойчивы к малым погрешностям измерения положения $x(t)$ и к малым погрешностям.

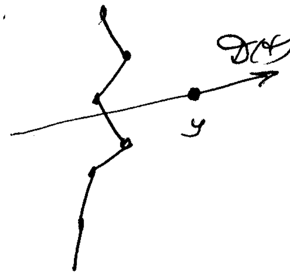
построения $\Pi(t)$.

Очень важно, что у нас нет проблем с определением, где находится конкретное положение $x(t)$: в $\Pi^-(t)$ или в $\Pi^+(t)$. Для определения знака используем вектор $D(t)$.

② Рассмотрим случай $n=2$. Сечение $W_c(t)$ упорядочены по c. Можем взять малую сетку по c и разбить при каждом c границу $\partial W_c(t)$ на части: на одной вектор $D(t)$ и „втыкается" в $W_c(t)$ при $u=-\mu$, на другой – при $u=+\mu$. Исходя

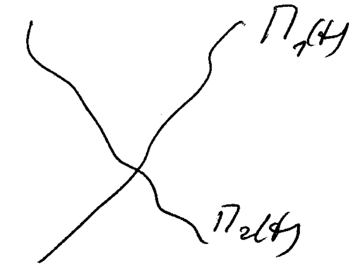


из этого, строим линию переключения $\Pi(t)$. При этом запоминаем только дискретный набор точек (упорядоченных по c в одну сторону от точки, соответствующей $c_{min}(t)$ и аналогично в другую). Каждому c соответствуют две точки на линии переключения. Соединяем узловые точки отрезками. Получаем приближение к „истинной", „идеальной" линии переключения.



Используя $D(t)$ легко определить, где находится конкретная точка y (в $\Pi^-(t)$ или в $\Pi^+(t)$).

(3) Если $P = \{u : u = \sum_{i=1}^{k} u_i,\ |u_i| \le \mu_i\}$, то можем строить поверхности переключения в количестве k штук. Если они попарно не зацепляются (помимо зацепления в $W_{c\,min}(t)$), то всё сохраняется. Здесь



после перехода к системе вида (4) получим $D(t)u = \sum_{i=1}^{k} D_i(t) u_i$

$$\underbrace{\begin{pmatrix} | & | & | \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} u \end{pmatrix}}_{u \in R^3} = D_1(t)u_1 + D_2(t)u_2 + D_3(t)u_3$$

(4) Фактически мы получаем разбиение пространства y при каждом t на клетки, в каждой из которых действует своя вершина параллелограмма (при $k > 2$ — многомерного параллелепипеда). Клеток столько-сколько вершин.

(5) В случае выпуклой цели и ограничения Q в виде отрезка или параллелепипеда так можно делать и при построении оптимальных стратегий второго игрока. Но [!!] они уже не будут обладать свойством устойчивости по отно—

щенно к малым информационным погрешностям, ибо там, в скалярном случае, на самой поверхности пресечения следует положить $V(t,y) = \pm\nu$, где $\pm\nu$ — крайние значения отрезка Q. Произвольное v, отличное от $\pm\nu$ уже нельзя выбрать. Аналогично на каждой компоненте в случае Q в виде параллелепипеда.

(6) Все сохраняется, если φ — квазивыпуклая функция (т.е. её множества Лебега выпуклые).

Если φ — вогнутая функция, то просто преобразуем игроков: бывший первый становится вторым и наоборот. То же, если φ — квазивогнутая функция.

(7) Идею разбиения на клетки можно использовать и в невыпуклом случае. Но [!!] надо следить за тем, являются ли соседними вершины P для соседних клеток. Кроме того, следует делать анализ устойчивости полученного способа управления.

[Показываются мультфильмы для управляемой материальной точки, осцилляторы, а также для задачи преследования двумя догоняющими одного убегающего]

Адаптивное управление на основе методов теории дифференциальных игр

В нашем курсе мы рассматривали задачи теории дифференциальных игр, в постановке которых чётко оговорены ограничения на управления первого и второго игроков.

Сейчас посмотрим, как это можно применить к задачам, где по постановке оговорено только ограничение P на управление первого игрока, а ограничение Q на управление второго заранее не задано.

Типичный случай — управление самолётом в условиях ветрового возмущения. „Стандартный" самолёт имеет четыре органа управления: силу тяги, руль высоты, руль направления, элероны. Оговорены пределы возможных отклонений управляющих органов от номинальных значений. Тем самым обоснованно можно считать заданным ограничение P на полезное управляющее воздействие.

Однако нельзя чётко обосновать ограничение Q на ветровое возмущение. Можно задать средние значения скорости ветра по координатам оси: номинала продольного ветра, вертикального и бокового. Но как объяснить, что отклонение от номинала, например, по боку должно быть не больше 10 м/с, а не 12 м/с.

Хотелось бы иметь постановки и методы решения, в которых ограничение на второго игрока заранее не оговаривается.

К недостаткам (с инженерной точки зрения) антагонистических дифференциальных игр следует отнести также и то, что первый игрок, минимизируя заданный показатель, работает на крайних управлениях. В случае с самолётом это приводит к резкой переброске рулей с одного крайнего значения на другое. В практических задачах это нежелательно. Приходится придумывать зоны линейного перехода или ещё что-то весьма кустарное.

Адаптивное управление, о котором сейчас будем говорить, не предполагает, что ограничение Q жёстко задано; кроме того, на „слабую" помеху оно реагирует „слабым" управлением.

1 Общая идея. Установим некоторое соответствие между виртуальным уровнем помехи и уровнем отрабатывающего её управления. Пусть

Q_k (где $k \geq 0$ — числовой параметр) — мысленно заданное ограничение на помеху. Условимся, что Q_k монотонно увеличивается с ростом k. Каждому k поставим в соответствие ограничение P_k на полезное управление. Оно увеличивается с ростом $k \in [0, 1]$, при $k = 1$ совпадает с ограничением P, заданным по постановке. При $k \geq 1$ полагаем $P_k = P$.

Паре Q_k, P_k сопоставим в пространстве t, x стабильный мост W_k, в котором управление, выбираемое из P_k, может удерживать движение, какой бы ни была помеха из Q_k.

Предположим, что множества W_k увеличиваются с ростом k.

Измеряя в момент t текущее фазовое состояние $x(t)$, идентифицируем индекс $\bar{k}$ трубки $W_{\bar{k}}$, на границе которой находится в рассматриваемый момент система, и на малом промежутке времени (шаг дискретной схемы управления) применяем подходящее управление (формируя его, например, по правилу экстремального приближения на мост $W_{\bar{k}}$) из множества $P_{\bar{k}}$.

Если движение выходит из $W_{\bar{k}}$, это означает, что на данном малом промежутке времени действовала помеха большего уровня, чем $Q_{\bar{k}}$. [Конечно, мы работаем с некоторым порогом ε выхода из $W_{\bar{k}}$. Иначе будет происходить (точнее, возможен) выход просто из-за дискретности Δ шага выбора полезного управления]. Тогда на следующем дискрете времени будем выбирать управление из множества $P_{\tilde{k}}$, связанного с трубкой $W_{\tilde{k}}$, $\tilde{k} > \bar{k}$, на границе которой находится система в начальный момент следующего дискрета.

Если же движение опускается во внутренность трубки $W_{\bar{k}}$, то на следующем дискрете времени применяем управление, уровень которого соответствует трубке $W_{\tilde{k}}$, где $\tilde{k} < \bar{k}$ и т.д.

Таким образом, уровень используемого управления „подстраивается" под уровень трубки W_k, на границе которой находится текущая точка. Следовательно, уровень вырабатываемого управления будет зависеть от того, где находится начальное состояние (в начальный момент во значение k_0 большое или нет) и от того, какого уровня идёт возмущение (управление второго игрока) при $t \geq t_0$.

② Реализация идеи для задач с линейной динамикой. Описанная идея — весьма общая. В настоящее время она эффективно реализована для линейных задач с динамикой вида

$$\dot{x} = A(t)x + B(t)u + C(t)v \qquad (5)$$

$$x \in R^n,\ u \in P,\ t \in [t_0, T]$$

Здесь P — выпуклое компактное ограничение на управление первого игрока, $[t_0, T]$ — промежуток процесса управления. Отметим, что P содержит нуль пространства R^p, в котором оно лежит.

Первый игрок заинтересован в приводе m выделенных координат системы (5) в момент T на терминальное множество M^*, которое предполагается выпуклым компактом в пространстве выделенных m компонент фазового вектора x. Условимся, что M^* содержит непустую окрестность начала координат этого пространства. Начало координат примем за „центр" множества M^*. Первую m выделентных компонент вектора x как можно ближе к центру множества M соответствует интересам первого игрока.

Стандартно переходим к системе без фазовых переменных в правой части:

$$\dot{y} = D(t)u + E(t)v \qquad (6)$$
$$y \in R^m, \; u \in P$$

2.1 Стабильные мосты. Рассматриваем семейство антагонистических дифференциальных игр

$$\dot{y} = D(t)u + E(t)v \qquad (7)$$
$$y \in R^m, \; \mathcal{M} \subset R^m, \; u \in \mathcal{P}, \; v \in Q$$

с фиксированным моментом окончания T. Множества $\mathcal{M}, \mathcal{P}, Q$ предполагаются выпуклыми компактами и рассматриваются как параметры игры. Каждой тройке $\mathcal{M}, \mathcal{P}, Q$ можно поставить в соответствие максимальный стабильный мост.

2.2 Выберем множество Q_{max}, трактуемое как максимальное „разумное" ограничение на управление второго игрока. Через W_{main}

обозначим максимальный стабильный мост для системы (7), соответствующий параметрам $\mathcal{P} = P$, $Q = Q_{max}$, $\mathcal{M} = M^{*}$. Назовём его главным мостом. Отметим, что Q_{max} содержит нуль и каждое сечение $W_{main}(t)$ также содержит нуль.

2.3 Дополнительно введём трубку W_{add}, каждое сечение $W_{add}(t)$ которой есть множество достижимости системы (7) при $u \equiv 0$, $v \in Q_{max}$ в момент t с начальным множеством $O(\varepsilon)$ (окрестность нуля радиусом ε), где $\varepsilon > 0$ мало, взятым в момент t_0. Подчеркнём, что $W_{add}(t)$ – множество достижимости второго игрока. Его можно интерпретировать как сечение максимального стабильного моста для системы (7) при

$$\mathcal{P} = \{0\}, \quad Q = Q_{max}, \quad \mathcal{M} = W_{add}(T),$$

2.4 Рассмотрим семейство трубок W_k, сечения которых определяются соотношением

$$W_k(t) = \begin{cases} k \cdot W_{main}(t), & 0 \le k \le 1 \\ W_{main}(t) + (k-1)\, W_{add}(t), & k > 1. \end{cases}$$

Множества $W_k(t)$ компактны, выпуклы и строго вложены друг в друга по возрастанию k.

Доказано, что каждая трубка W_k при $k \in [0, 1]$ есть максимальный стабильный мост для системы (7), соответствующий ограничению $k P$ на управление первого игрока, ограничению $k Q_{max}$ на управление второго и терминальному множеству $k M^*$. При $k > 1$ множество W_k есть стабильный мост для параметров

$$\mathcal{P} = P, \quad \mathcal{Q} = kQ_{max}, \quad \mathcal{M} = M^* + (k-1)\, W_{add}(T).$$

Опираясь на эту систему стабильных трубок, строим адаптивное управление.

Подробно метод адаптивного управления описан в трёх статьях: 1) Труды ИММ, т. 15, N 3, 2009, 2) ПММ, 2006, т. 70, вып. 5, 3) ПММ, 2009, т. 73, вып. 4.

[В конце лекции показывалась работа адаптивного управления для задачи посадки самолёта]