Синтез оптимального управления в линейных задачах быстродействия

На прошлом занятии мы говорили о возможности построения оптимального (или близкого к оптимальному) управления обратной связи в задачах управления на основе умения решения программных задач. Такую проблему стали называть задачей синтеза.

Изучая множества достижимости, мы отдельно выделили случай линейных по динамике проблем. Именно с таких задач и началось изучение проблемы синтеза.

Следует отметить, что построение управления обратной связи в задачах оптимального быстродействия началось лет за 5-7 раньше возникновения принципа максимума Понтрягина. В США этим занимался Д. Бушоу (диссертация защищена в 1952 г), в нашей стране А. А. Фельдбаум (статьи в русских журналах в начале 50-х годов). Таким образом, образцы построения оптимального управления обратной связи были. Но с возникновением принципа максимума Понтрягина (ПМП) на всё это взглянули по-новому.

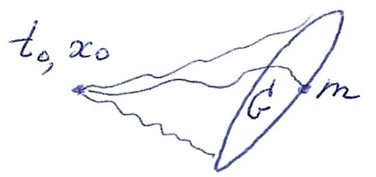
① Рассматриваются задачи быстродействия на плоскости, причём целевое множество — это точка, совпадающая с началом координат. Предполагается стационарность линейной управляемой системы:

$$\dot{x} = Ax + Bu \qquad (1)$$
$$u \in P$$

Здесь матрицы A и B не зависят от t.

В этом случае разумно задачу ставить не в классе функций $u(t, x)$, а в классе функций $u(x)$. Опускаем вопрос о допустимых управлениях: в известных уже примерах оптимальная функция $u^0(x)$ получалась разрывной; поэтому разрывные функции нельзя выкидывать из допустимых; но что понимать под решением (1), если вместо u подставлять $u(x)$ разрывное? Будем "синтезировать" $u(x)$ в точке x в виде начального значения управления $t \to u^*(t)$, оптимального по быстродействию в начало координат, если в качестве начальной взята точка x. Посмотрим, что получается, когда перебираем точки x на плоскости. Потом придумаем, как объяснять решение (движение (1)) при использовании функции $u^0(x)$. Примерно так рассуждали в конце 50-х годов.

! (2) Очень важно, что была доказана замкнутость множества достижимости $G(t; t_0, x_0)$ для стационарных линейных систем в классе кусочно-непрерывных управлений. При этом множество P считалось выпуклым многогранником. В самом деле, если бы замкнутости множества достижимости не было, то мы могли бы говорить лишь о последовательности кусочно-непрерывных программных управлений, реализующих оптимальный результат.



Пусть замыкание G впервые налетело на целевую точку m. Но если $G(t)$ не является замкнутым, то какое управление (кусочно-непрерывное) ведёт в m. При отсутствии замкнутости $G(t)$ его может не быть?

Кроме того, полезным является свойство единственности оптимального программного управления. Как выделять начальное значение оптимального кусочно-непрерывного управления, если оно не является единственным? Это связано с наличием "площадок" на границе множества достижимости? Сформулируем теорему о единственности кусочно-непрерывного программного управления (см. книги "Понтрягин + соавторы", "Болтянский")

Наконец, важным является и утверждение о том, что если мы нашли кусочно-непрерывное управление, осуществляющее перевод и удовлетворяющее ПМП, то оно является оптимальным. Т.е. надо запретить такой вариант развития множества достижимости

Множество достижимости содержит точку m в разные моменты на своей границе.

Если точка m для задачи быстродействия — начало координат $x=0$, то после первого попадания в множество достижимости точка m остаётся внутри его. Это такое "грубое" достаточное условие оптимальности в виде ПМП. Соответствующее утверждение также есть в указанных выше двух книгах.

③ Мы начнём с управляемой материальной точки на прямой. Начальные сведения о такой задаче были в лекции 2. Воспользуемся ими.

Управляем силой:

Закон Ньютона

$$\ddot{y} = u \quad , \quad |u| \le \mu$$

Переходим к стандартной записи вида

$$\dot{x} = Ax + Bu \quad , |u| \le \mu .$$

Для этого полагаем $x_1 = y$, $x_2 = \dot{y}$. Тогда

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u \quad , \quad |u| \le \mu \end{cases}$$

Т.е. в нашем случае $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Сопряжённая система

$$\dot{\psi} = -A^T \psi \quad , \qquad A^T = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

или

$$\begin{cases} \dot{\psi}_1 = 0 \\ \dot{\psi}_2 = -\psi_1 \end{cases}$$

Решение сопряжённой системы

$$\psi_1(t) = \text{const}$$

$$\psi_2(t) = -\psi_1(t - t_0) + \psi_2(t_0) .$$

Принцип максимума

$$\max_{|p| \le \mu} \psi^T(t)(Ax(t) + Bp) = \psi^T(t)(Ax(t) + Bu(t)) \qquad (2)$$

или

$$\max_{|p| \le \mu} \psi^T(t) B p = \psi^T(t) B u(t) . \qquad (3)$$

! Видим, что в (3) нет A. Это всегда так в линейных системах.

Будем считать $t_0 = 0$.

В нашем случае соотношение (3) принимает вид

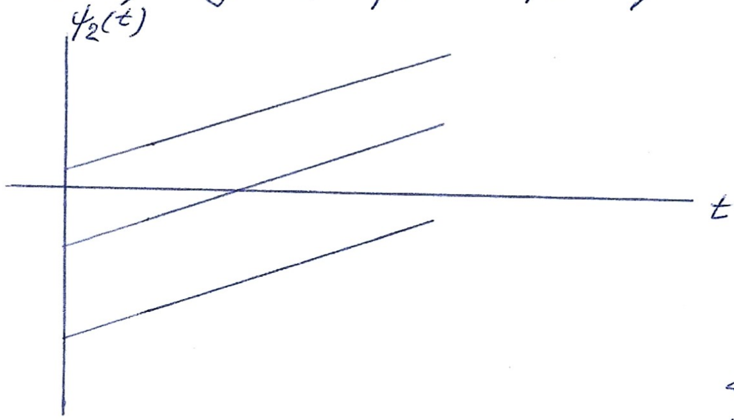
$$\max_{|p|\le\mu} (\overset{\text{const}}{\psi_1}\cdot 0 + \psi_2(t)\cdot 1\cdot p) = (\psi_1\cdot 0 + \psi_2(t)\cdot 1\cdot u(t))$$

т.е.

$$\max_{|p|\le\mu} (-\psi_1 t + \psi_2(0))p = (-\psi_1\cdot t + \psi_2(0))u(t)$$

Стало быть,

$$u(t) = \mu\, \mathrm{sign}\,(-\psi_1 t + \psi_2(0))$$



$\psi_2(t)$ - линейная функция. Поэтому функция $\mathrm{sign}\,\psi_2(t)$ меняет знак не более одного раза ⟹ управление, удовлетворяющее принципу максимума имеет не более одного переключения.

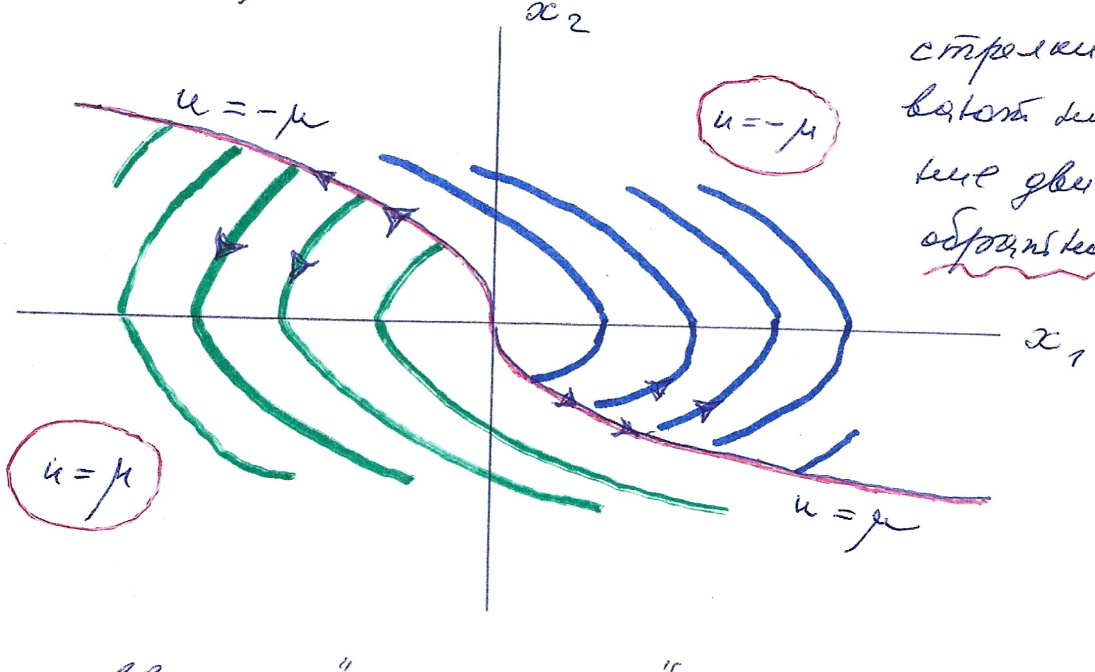
√ [Проверить совпадение с формулами на стр. 13 лекции 2 (знак)]

④ Теперь важно, что, в зависимости от начального состояния x_0, множество достижимости $G(t, t_0, x_0)$ налегает на точку m при всевозможных значениях ψ_1 и ψ_2 (компоненты вектора нормали l к границе множества достижимости в точке налегания)

Стало быть, в последний момент управление оптимальное может быть как μ, так и $-\mu$.

При этом, если мы "пятимся" от последнего момента, то либо есть переключение на противоположное по знаку крайнее управление, либо его нет.

Поэтому из точки m (начало координат) выпустим в обратном времени два движения: в силу $u=\mu$ и $u=-\mu$. Поскольку работаем со 2 законом Ньютона, то получаемые траектории имеют вид парабол, ориентированных вдоль горизонтальной оси x_1: $x_1 = \dfrac{x_2^2}{\pm 2\mu}$. Имеем две ветви



стрелки показывают направление движения в обратном времени.

Теперь вводим "скользящий" момент переключения и продолжая движения в обратном времени, заполняем траекториями нижнюю (на основе верхней дуги) и верхнюю (на основе нижней дуги) часть плоскости.

Таким образом, вся плоскость получилась разбитой "линией переключения" на две части. Ниже оптимальное управление равно $u=\mu$, выше $u=-\mu$. Синтез построен.

5 Теперь рассмотрим управляемый осциллятор:

$$\ddot{y} = -y + u \,, \qquad |u| \le \mu.$$

Переходим к стандартной записи

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -x_1 + u \,, \quad |u| \le \mu \end{cases} \tag{4}$$

Имеем

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

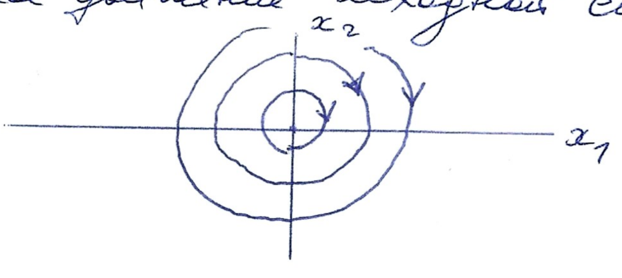
Матрица A^T для сопряжённой системы

$$A^T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Rightarrow \dot{\psi} = -A^T \psi \,, \quad \text{или}$$

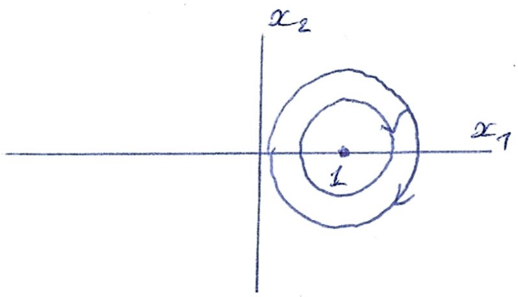
$$\begin{cases} \dot{\psi}_1 = \psi_2 \\ \dot{\psi}_2 = -\psi_1 \end{cases}$$

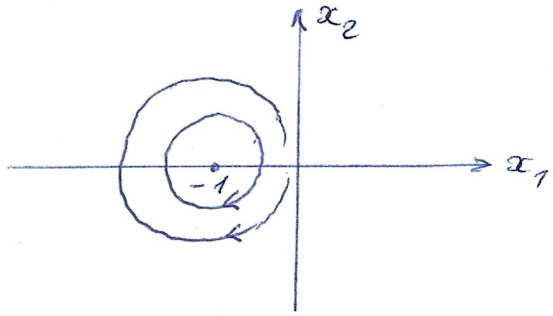
Это тоже колебательная система.

Свободное движение исходной системы: окружность с центром в нуле.



Если в исходной системе действует $u=1$, то в силу (4) траекторией становится окружность в точке $x_2=0$, $x_1=1$. Если $u=-1$, то $x_2=0$, $x_1=-1$.





Движение сопряжённой системы — тоже окружность.

Условие максимума:

$$\max_{|p|\le\mu} \psi^T(t) B p = \psi^T(t) B u(t).$$

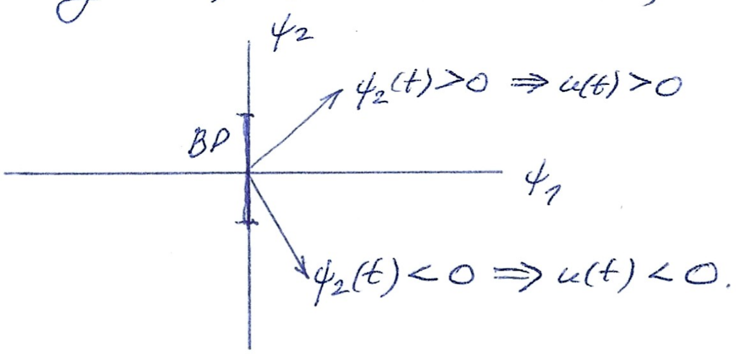
В нашей конкретике:

$$\max_{|p|\le\mu} \psi_2(t) p = \psi_2(t) u(t).$$

Таким образом, опять

$$u(t) = \mu \operatorname{sign} \psi_2(t)$$

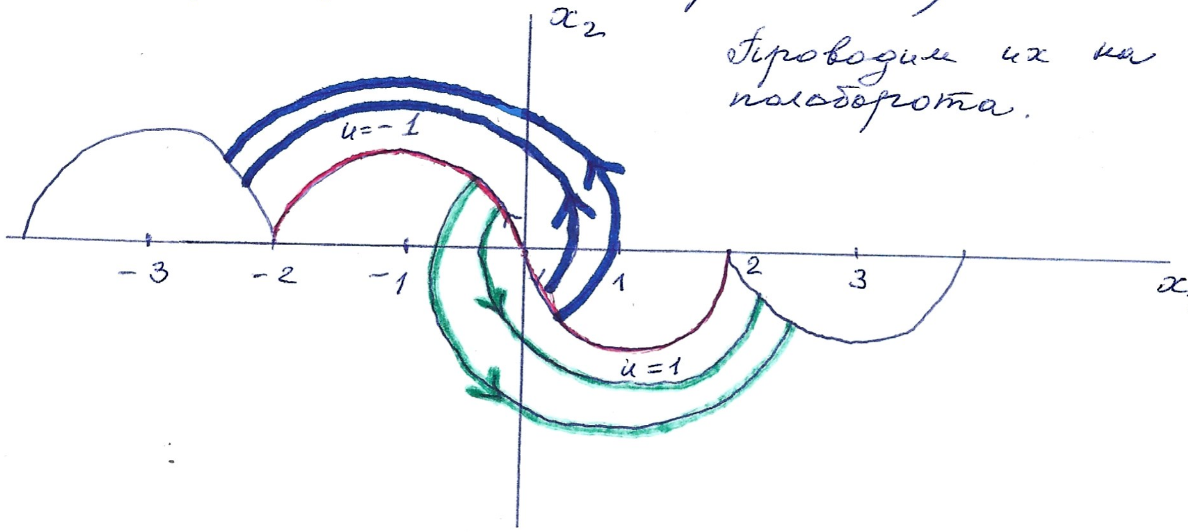
Поэтому $u(t)$ изменяет знак, когда $\psi_2(t)=0$:



При этом смена знака происходит, когда $\psi_2(t)$ проходит нулевое положение, т. е. очередная смена через пол оборота вектора $\psi(t)$.

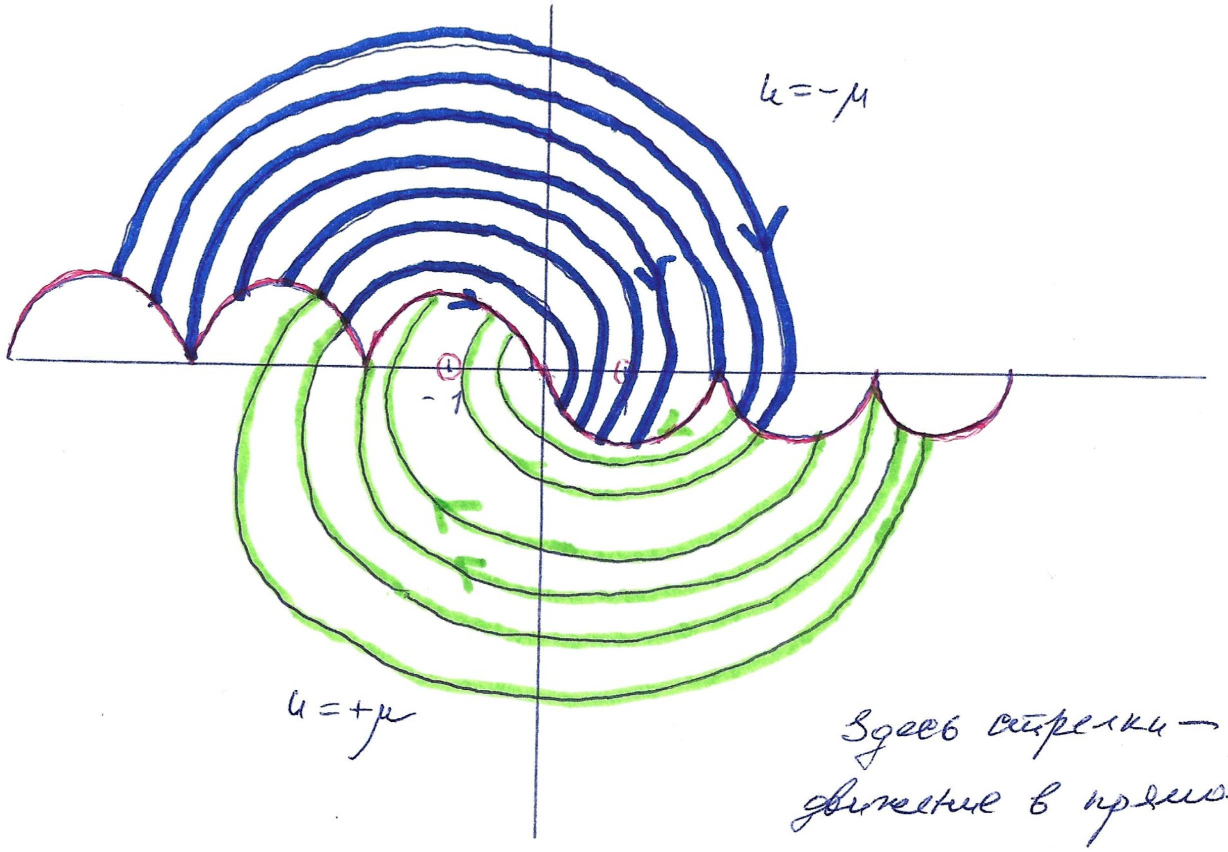
6) Опять очень важно, что в момент окончания могут быть любые значения (по знаку) компоненты $\psi_2(t)$. Поэтому из конечной точки m (начало координат) выпускаем в обратном времени два движения при $u = \mu$ и $u = -\mu$

Проводим их на пол оборота.



Теперь на полученных дугах делаем скользящий момент переключения и переходим на противоположное по знаку крайнее управление и держим его в течение половины оборота.

Продолжая такой процесс, заполняем траекториями всю плоскость с одной кусочно-гладкой линией переключения. Выше этой линии оптимальное управление обратной связи $u^0(x) = -\mu$, ниже $u^0(x) = \mu$.



Здесь стрелки — движение в прямом времени.

Движение по линии переключения — только на двух дугах, примыкающих к нулю. Остальные дуги линии переключения пробегаются.

Далее в лекции идёт разговор по книге В.Г. Болтянского.