

ПОСТРОЕНИЕ И НЕЙРОСЕТЕВАЯ СТАБИЛИЗАЦИЯ ОПТИМАЛЬНЫХ ПО БЫСТРОДЕЙСТВИЮ ПРОГРАММНЫХ ДВИЖЕНИЙ МАНИПУЛЯТОРА С УЧЕТОМ КОНСТРУКТИВНЫХ ОГРАНИЧЕНИЙ

А. В. Тимофеев, А. В. Завьялов

Санкт-Петербургский институт информатики и автоматизации РАН
199178, Санкт-Петербург, 14-я линия В.О., д.39
adil@iias.spb.su

УДК 681.51

А. В. Тимофеев, А. В. Завьялов. Построение и нейросетевая стабилизация оптимальных по быстродействию программных движений манипулятора с учетом конструктивных ограничений // Труды СПИИРАН. Вып. 1, т. 2. — СПб: СПИИРАН, 2002.

Аннотация. Проблема построения и стабилизации оптимальных по быстродействию программных движений (ПД) манипулятора рассмотрена в контексте применения нейросетей для аппроксимации обратной модели динамики. Представление ПД в виде траекторий на фазовой плоскости использовано для получения явных зависимостей обобщенных координат от времени, краевых условий и конструктивных ограничений. Кратко рассмотрен вопрос синхронизации движений различных звеньев манипулятора. Нелинейный закон управления, использующий в качестве обратной модели динамики ее нейросетевое представление, проиллюстрирован при помощи моделирования на примере пятизвенного манипулятора (без учета динамики приводов, т.е. управление в виде обобщенных сил). Метод адресует оптимальное по быстродействию управление многосвязной нелинейной системой в режиме реального времени. — Библи. 5 назв.

UDC 681.51

A. V. Timofeev, A. V. Zavyalov. Programming and neural-net stabilizing of manipulator time-optimal trajectories with construction constraints. // SPIIRAS Proceedings. Issue 1, v. 2. — SPb: SPIIRAS, 2002.

Abstract. The problem of programming and neural-net stabilizing of manipulator time-optimal trajectories is taken into consideration within context of neural-nets threatening for inverse dynamic model approximation. Trajectories are put into phase plane and generalized coordinates direct dependences on time and constraints are obtained. The problem of links movements' synchronization is addressed in brief. A nonlinear control law with neural-net representation of inverse dynamic model is illustrated with simulations of 5 DOF manipulator control (without consideration of servo dynamics, i.e. control signals are generalized forces). Method addresses time-optimal real-time control of multilinked nonlinear system. — Bibl. 5 items.

1. Постановка задачи

Рассмотрим аналитическое описание динамики исполнительных механизмов m -звенного робота-манипулятора в виде нелинейных дифференциальных уравнений Лагранжа II-го рода:

$$A(\xi, q)\ddot{q} + B(\xi, q, \dot{q}) = U + \pi, \quad q(t_0) = q_0, \quad \dot{q}(t_0) = \dot{q}_0, \quad (1)$$

где $m \times m$ матрица-функция $A(\xi, q)$ характеризует массо-инерционные свойства манипулятора, m — вектор-функция $B(\xi, q, \dot{q})$ характеризует действие сил тяжести, кориолисовых и центробежных сил, $q \in R^m$ — вектор обобщенных координат, $\xi \in R^p$ — вектор параметров, $U \in R^m$ — вектор управляющих моментов, развиваемых в приводах манипулятора, $\pi \in R^m$ — вектор внешних возмущений.

Требуется синтезировать оптимальное по быстродействию программное движение (ПД) и закон стабилизации (регулятор) переходных процессов (ПП), обеспечивающий перевод исполнительного механизма из заданного начального состояния в целевое состояние:

$$q(t_0) = q_0, \quad \dot{q}(t_0) = \dot{q}_0; \quad q(t_T) = q_T, \quad \dot{q}(t_T) = \dot{q}_T \quad (2)$$

за минимальное время:

$$T = t_T - t_0 \rightarrow \min, \quad (3)$$

где t_0 — время начала движения, t_T — время окончания движения. Необходимо также учесть конструктивные ограничения:

$$|\ddot{q}_i| \leq \alpha_i, \quad i = 1, 2, \dots, m, \quad (\text{ограничения на модуль обобщенного ускорения } i\text{-го звена}),$$

$$|\dot{q}_i| \leq \beta_i, \quad (\text{ограничения на модуль обобщенной скорости}),$$

$$q_i \in [q_i^{\min}, q_i^{\max}], \quad (\text{ограничения на диапазон изменения обобщенной координаты}). \quad (4)$$

Постановка задачи расширена по сравнению с [1], так как дополнительно учитываются ограничения на обобщенные скорости и координаты. Требуется также обеспечить синхронизацию ПД всех звеньев манипулятора для того, чтобы попадание в целевое состояние произошло одновременно. При этом важно исследовать нетривиальный случай ненулевых обобщенных скоростей в целевом состоянии. В статье предлагаемые алгоритмы рассматриваются в нейросетевом контексте.

2. Представление ПД в виде траекторий на фазовой плоскости

Предположим, что значения обобщенных координат и обобщенных скоростей для всех звеньев манипулятора могут быть измерены в любой момент времени. Пусть необходимо максимально быстро перевести манипулятор из текущего состояния в целевое состояние, для которого задано положение и ориентация схвата в рабочем пространстве манипулятора, а также скорости их изменения. Предположим, что обратная задача кинематики решена каким-либо способом для целевого состояния, и соответствующие ему значения обобщенных координат и обобщенных скоростей звеньев представлены в виде (2). В этом случае построение оптимальных по быстродействию ПД может быть произведено в пространстве обобщенных координат отдельно для каждого звена [1].

При наличии ограничений на модуль обобщенной скорости (4) оптимальное по быстродействию ПД формируется следующим образом: сначала производится движение с максимальным по модулю обобщенным ускорением, затем, в случае достижения максимального по модулю значения обобщенной скорости, производится движение с нулевым ускорением, и, наконец, с противоположным по знаку обобщенным ускорением [2]. Моменты ступенчатого изменения ускорений будем называть переключениями управления. Оптимальное ПД может содержать два, одно или ни одного переключения управления в зависимости от взаимного расположения начальной и целевой точек на фазовой плоскости (рис. 1):

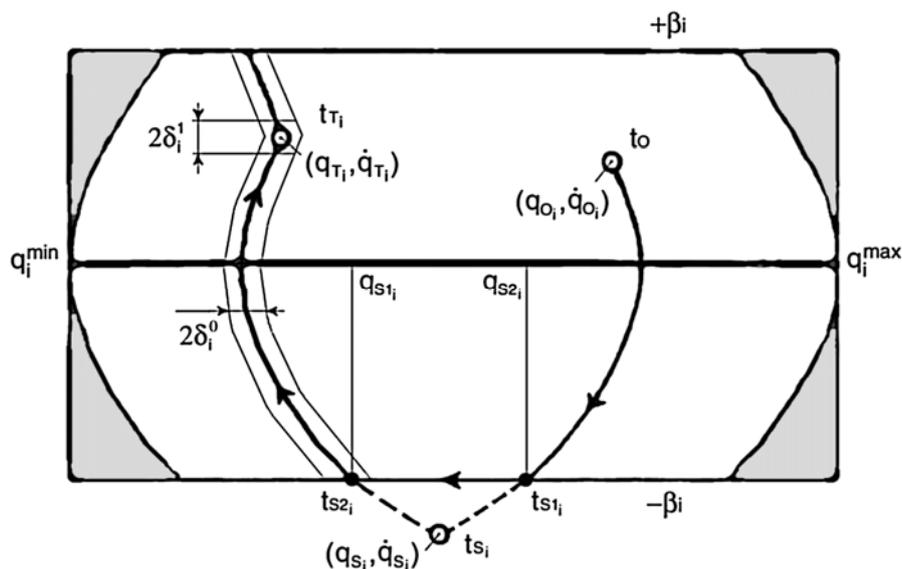


Рис. 1. Фазовая плоскость

Такое ПД представляется на фазовой плоскости состоящим из кусков парабол с вершинами на оси абсцисс и ветвями, направленными влево или вправо в зависимости от знака ускорения. В случае если достигается ограничение по скорости, появляется горизонтальный участок движения с постоянной, максимальной по модулю скоростью.

Затемненные области на рис. 1 соответствуют недостижимым из-за ограничений на ускорения положениям на фазовой плоскости. При достижении конечного ограничителя по некоторой координате скорость должна быть равна нулю, наличие же ограничения на ускорение не позволяет иметь слишком большую скорость вблизи конечного ограничителя. Таким образом, выбор целевых положений на фазовой плоскости обусловлен следующими соотношениями:

$$(q_{Ti}, \dot{q}_{Ti}): \quad |\dot{q}_{Ti}| \leq \beta_i; \quad q_{Ti} \geq q_i^{\min} + \frac{\dot{q}_{Ti}^2}{2\alpha_i}; \quad q_{Ti} \leq q_i^{\max} - \frac{\dot{q}_{Ti}^2}{2\alpha_i}. \quad (5)$$

При построении ПД зададимся допустимыми отклонениями от целевых значений обобщенных скоростей и координат, настолько малыми, чтобы не вносить существенную долю в итоговую погрешность попадания в целевое состояние. Обозначим их, соответственно, δ_i^1 и δ_i^0 (рис. 1).

Два отрезка парабол, непосредственно приводящие в целевое состояние образуют кривую переключения ускорения. Аналитически положение точки фазовой плоскости, соответствующей моменту начала ПД, относительно кривой переключения определяется функцией:

$$\Omega_i \equiv -\text{sign} \left[\Delta_{\delta_i^0} \left(q_{Oi} - q_{Ti} + \frac{\dot{q}_{Oi}^2 - \dot{q}_{Ti}^2}{2\alpha_i} \text{sign}(\dot{q}_{Oi} - \dot{q}_{Ti}) \right) \right], \quad (6)$$

где Δ — функция, реализующая зону нечувствительности заданной ширины δ_i^0 .

Таким образом, точка находится левее кривой переключения, если $\Omega_i = +1$, правее, если $\Omega_i = -1$, и в δ_i^0 -окрестности кривой переключения, если $\Omega_i = 0$.

Вертикальное положение точки начала движения относительно целевого положения определяется значением функции:

$$\Theta_i \equiv \text{sign} \left[\Delta_{\delta_i^1} (\dot{q}_{Ti} - \dot{q}_{Oi}) \right], \quad (7)$$

где δ_i^1 - допустимая погрешность \dot{q}_i .

Таким образом, точка находится выше целевой, если $\Theta_i = -1$, ниже, если $\Theta_i = +1$ и в δ_i^1 -окрестности по вертикали, если $\Theta_i = 0$. Если $\Omega_i = 0$ и $\Theta_i = 0$ одновременно, точка фазовой плоскости находится в окрестности целевого положения по обеим осям, и $T = t_T - t_O = 0$.

Введем вспомогательную функцию:

$$\eta_i^0 \equiv \begin{cases} \alpha_i (q_{Ti} - q_{Oi}) \Omega_i + \frac{1}{2} (\dot{q}_{Ti}^2 + \dot{q}_{Oi}^2), & \Omega_i \neq 0 \\ 0, & \Omega_i = 0 \end{cases} \quad (\eta_i^0 \geq 0). \quad (8)$$

Случай $\sqrt{\eta_i^0} \leq \beta_i$ соответствует такому ПД, при совершении которого ограничение по модулю обобщенной скорости не достигается, а при $\dot{q}_{Si} > \beta_i$ — достигается.

3. Вычисление параметров оптимальных ПД

3.1. Случай $\Omega_i = 0$

Оптимальное по быстродействию движение производится с постоянным, максимальным по модулю обобщенным ускорением. Переключение управления не производится.

$$t_{Ti} = t_O + \frac{1}{\alpha_i} (\dot{q}_{Ti} - \dot{q}_{Oi}) \Theta_i, \quad (9)$$

где t_{Ti} - момент завершения оптимального движения i -го звена.

Явный вид зависимости обобщенной координаты от времени при совершении оптимального по быстродействию ПД:

$$q_i^{opt}(t) = q_{Oi} + \dot{q}_{Oi}(t - t_O) + \frac{1}{2} \alpha_i (t - t_O)^2 \Theta_i, \quad t \in [t_O, t_{Ti}]. \quad (10)$$

3.2. Случай $\Omega_i \neq 0$ и $\sqrt{\eta_i^0} \leq \beta_i$

Оптимальное по быстродействию движение производится с однократным переключением знака обобщенного ускорения. Модуль ускорения поддерживается постоянным и максимальным.

Переключение управления необходимо производить при обобщенной скорости:

$$\dot{q}_{Si} \equiv \Omega_i \sqrt{\eta_i^0}, \quad (11)$$

что следует из рассмотрения пересечения парабол на фазовой плоскости, проходящих через начальную и конечную точки оптимальной траектории с разными по знаку обобщенными ускорениями (рис. 1). Соответствующую обобщенную координату получим из уравнения параболы на фазовой плоскости [1], проходящей через начальную точку движения, подставляя \dot{q}_{Si} :

$$q_{Si} = q_{oi} + \frac{\dot{q}_{Si}^2 - \dot{q}_{oi}^2}{2\alpha_i} \Omega_i. \quad (12)$$

Момент времени, в который должно производиться переключение управления, а также момент окончания оптимального движения выражаются соответственно:

$$t_{Si} = t_o + \frac{1}{\alpha_i} (\dot{q}_{Si} - \dot{q}_{oi}) \Omega_i, \quad (13)$$

$$t_{Ti} = t_{Si} + \frac{1}{\alpha_i} (\dot{q}_{Si} - \dot{q}_{Ti}) \Omega_i = t_o + \frac{1}{\alpha_i} \left[2\sqrt{\eta_i^0} - (\dot{q}_{oi} + \dot{q}_{Ti}) \Omega_i \right], \quad (14)$$

Зависимость обобщенной координаты от времени для оптимального ПД имеет вид:

$$q_i^{opt}(t) = \begin{cases} q_{oi} + \dot{q}_{oi}(t-t_o) + \frac{1}{2} \alpha_i (t-t_o)^2 \Omega_i, & t \in [t_o, t_{Si}] \\ q_{Si} + \dot{q}_{Si}(t-t_{Si}) - \frac{1}{2} \alpha_i (t-t_{Si})^2 \Omega_i, & t \in (t_{Si}, t_{Ti}] \end{cases} \quad (15)$$

3.3. Случай $\Omega_i \neq 0$ и $\sqrt{\eta_i^0} > \beta_i$

Оптимальное по быстродействию движение производится с последовательным переключением обобщенного ускорения сначала из максимального по модулю в нулевое значение, затем из нулевого в противоположное по знаку начальному.

Оба переключения производятся при максимальной по модулю обобщенной скорости (рис. 1):

$$\dot{q}_{S1i} = \dot{q}_{S2i} = \beta_i \Omega_i. \quad (16)$$

Соответствующие координаты получим из уравнений парабол на фазовой плоскости, проходящих через начальную и конечную точки движения, подставляя \dot{q}_{S1i} и \dot{q}_{S2i} соответственно:

$$q_{S1i} = q_{oi} + \frac{\beta_i^2 - \dot{q}_{oi}^2}{2\alpha_i} \Omega_i, \quad q_{S2i} = q_{Ti} + \frac{\dot{q}_{Ti}^2 - \beta_i^2}{2\alpha_i} \Omega_i. \quad (17)$$

Моменты времени перехода в режим постоянной максимальной по модулю обобщенной скорости, выхода из режима постоянной скорости и завершения движения выражаются соответственно формулами:

$$t_{S1i} = t_o + \frac{1}{\alpha_i} (\beta_i - \dot{q}_{oi} \Omega_i), \quad (18)$$

$$t_{S2i} = t_{S1i} + \frac{1}{\alpha_i \beta_i} (\eta_i^0 - \beta_i^2) = t_o + \frac{1}{\alpha_i} \left(\frac{\eta_i^0}{\beta_i} - \dot{q}_{oi} \Omega_i \right), \quad (19)$$

$$t_{Ti} = t_{S2i} + \frac{1}{\alpha_i} (\beta_i - \dot{q}_{Ti} \Omega_i) = t_o + \frac{1}{\alpha_i} \left(\frac{\eta_i^0}{\beta_i} + \beta_i - (\dot{q}_{oi} + \dot{q}_{Ti}) \Omega_i \right). \quad (20)$$

Зависимость обобщенной координаты от времени для оптимального ПД имеет вид:

$$q_i^{opt}(t) = \begin{cases} q_{oi} + \dot{q}_{oi}(t-t_o) + \frac{1}{2} \alpha_i (t-t_o)^2 \Omega_i, & t \in [t_o, t_{S1i}] \\ q_{S1i} + \beta_i (t-t_{S1i}) \Omega_i, & t \in (t_{S1i}, t_{S2i}] \\ q_{S2i} + \beta_i (t-t_{S2i}) \Omega_i - \frac{1}{2} \alpha_i (t-t_{S2i})^2 \Omega_i, & t \in (t_{S2i}, t_{Ti}] \end{cases}. \quad (21)$$

3.4. Зависимость времени совершения оптимального ПД от положения на фазовой плоскости точки начала движения при постоянном положении целевой точки

Зависимость представлена на рис. 2. Полупрозрачная поверхность соответствует случаю, когда не учитываются ограничения на модуль обобщенной скорости. Очевидно, что учет ограничений приводит к увеличению общего времени совершения ПД:

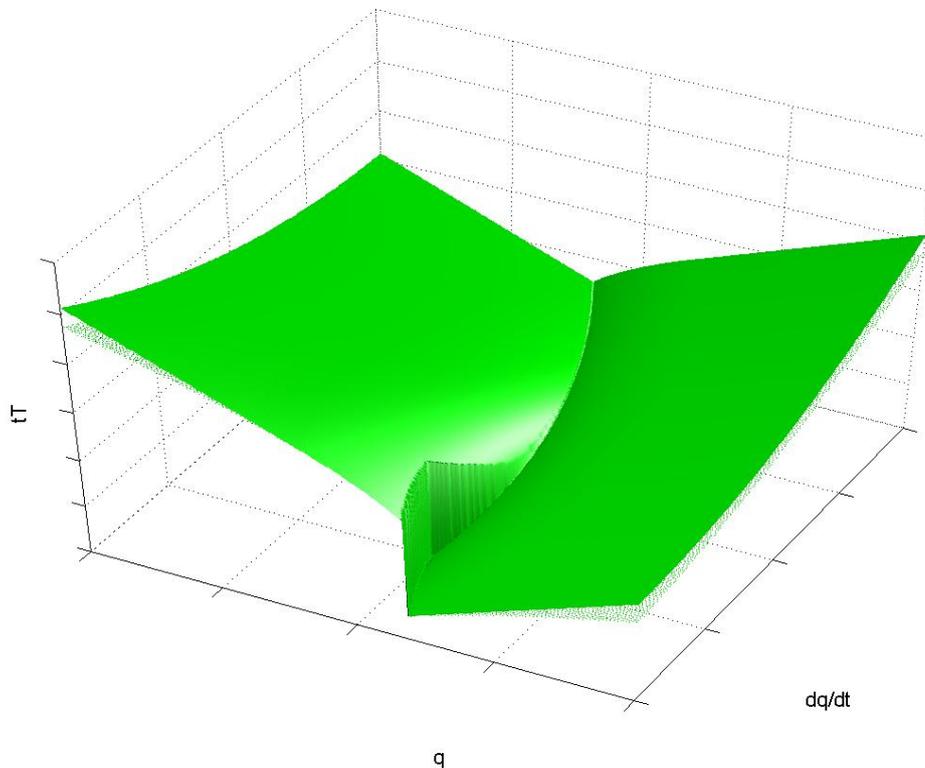


Рис. 2. Время совершения оптимального ПД как функция расположения на фазовой плоскости точки начала движения относительно целевой точки

4. Синхронизация ПД звеньев манипулятора

Время совершения оптимальных по быстродействию ПД звеньями манипулятора различно. В том случае, если целевая скорость не равна нулю, более быстрые звенья успеют покинуть целевые состояния еще до того, как медленные звенья достигнут своих целевых состояний. Таким образом, существует необходимость скорректировать ПД всех звеньев, за исключением того, у которого t_{Ti} максимально. В результате трансформации все ПД должны завершиться в один момент времени t_T .

В зависимости от взаимного расположения начальной и целевой точек на фазовой плоскости выбирается один из вариантов трансформации оптимальной траектории по i -ой степени свободы в направлении увеличения конечного времени прихода в целевое состояние. Трансформация траектории производится таким образом, чтобы завершающий этап движения протекал в режиме постоянной скорости, а не постоянного ускорения, что должно снизить динамическую погрешность попадания в целевое состояние. На фазовой плоскости трансформация выражается, в частности, в образовании горизонтальной площадки перед точкой (q_{Ti}, \dot{q}_{Ti}) .

Этот вариант синхронизации легко поддается параметризации, что немаловажно, так как необходимо иметь возможность по известному значению t_T получить искомые значения параметров трансформации оптимальных ПД путем не слишком громоздких вычислений.

Среди возможных путей увеличения t_{Ti} можно рассматривать также ужесточение ограничений на модули обобщенных ускорений или минимизацию мощности управления или уменьшение модулей производных от обобщенных сил и другие варианты. Их общей особенностью является сложность параметризации, хотя практическая ценность, несомненно, выше.

Итак, по требуемому времени завершения движения t_T аналитически вычисляется значение специального параметра трансформации (например, ширины горизонтальной площадки на фазовой плоскости), затем по значению параметра трансформации в явном виде вычисляются зависимости обобщенных ускорений, скоростей и координат от времени. Соответствующие соотношения здесь не приводятся.

5. Обратная модель динамики манипулятора

Используя представление массогабаритных характеристик манипулятора в виде параметров Денавита-Хартенберга [3], обратная модель динамики может быть представлена в следующем виде:

$$\sum_{i=j}^m \sum_{k=1}^i \text{tr}(U_{ij} H_i U_{ik}^T) \ddot{q}_k + \sum_{i=j}^m \sum_{k=1}^i \sum_{l=1}^i \text{tr}(U_{ij} H_i U_{ikl}^T) \dot{q}_k \dot{q}_l - \sum_{i=j}^m m_i G^T U_{ij} R_i^* = u_j. \quad (22)$$

Приведем (22) к виду (1):

$$A(\xi, q)_{j,k} = \begin{cases} \sum_{i=j}^m \text{tr}(U_{ij} H_i U_{ik}^T), & k \leq j, \\ \sum_{i=k}^m \text{tr}(U_{ij} H_i U_{ik}^T), & k > j, \end{cases} \quad (23)$$

$$B(\xi, q, \dot{q})_j = \sum_{i=j}^m \sum_{k=1}^i \sum_{l=1}^i \text{tr}(U_{ij} H_i U_{ikl}^T) \dot{q}_k \dot{q}_l - \sum_{i=j}^m m_i G^T U_{ij} R_i^*. \quad (24)$$

Согласно [4], стабилизирующий закон управления манипулятором, использующий ОМД, выражается в следующем виде:

$$u(t) = A(\xi, q) [\ddot{q}^{opt} + \Gamma_1 (\dot{q} - \dot{q}^{opt}) + \Gamma_0 (q - q^{opt})] + B(\xi, q, \dot{q}), \quad (25)$$

где Γ_1, Γ_0 — устойчивые матрицы.

6. Нейросетевая идентификация обратной модели динамики манипулятора

Нейросетевой аппроксиматор отображения (1) для i -той координаты может быть реализован в виде персептрона [5] с одним скрытым слоем и линейным нейроном на выходе. Каждая такая нейросеть может аппроксимировать отображение $w = \ddot{q}^{opt} + \Gamma_1 (\dot{q} - \dot{q}^{opt}) + \Gamma_0 (q - q^{opt})$, q и \dot{q} в u_i . Таким образом, необходимо количество нейросетей, равное числу степеней свободы манипулятора. Нейросетевая идентификация ОМД может производиться одним из известных алгоритмов.

Для подтверждения теоретических положений был проведен ряд опытов по моделированию оптимального по быстродействию управления пятизвенным манипулятором при помощи регулятора (25) с применением нейросетевой аппроксимации ОМД. Нейросети были обучены при помощи метода обратного распространения ошибки [5] по обучающим выборкам, сформированным априори на заданных траекториях по известной ОМД. Выгода от использования нейросетей в этом случае связана с преимуществами параллельных вычислений. В случае большой избыточности манипулятора матричные вычисления по известной ОМД могут оказаться столь громоздкими, что выполнять их в реальном времени станет невыгодно или вовсе невозможно.

Для аппроксимации ОМД пятизвенного манипулятора было использовано 5 MLP [5] с 15 нейронами во входном слое (с сигмоидальными функциями активации), 31 нейроном в скрытом слое [1] (с сигмоидальными функциями активации) и одним выходным нейроном (с линейной функцией активации). Дополнительно для улучшения условий обучения нейросетей применена фильтрация ступенчатых сигналов оптимальных обобщенных ускорений, входящих в обучающие выборки, учитывая, что более гладкая функция аппроксимируется нейросетью с лучшим качеством и с лучшими обобщающими свойствами. Кроме того, входные и выходные сигналы, вошедшие в обучающие выборки, нормировались таким образом, чтобы они принимали значения в интервале от -1 до 1 . В таких условиях обучение протекает более успешно, чем без нормализации. Для улучшения обобщающих свойств нейросетей в обучающие выборки вводились шумы с возрастающей по мере обучения дисперсией.

Моделирование проведено средствами MATLAB R12 (MathWorks, inc.). Результаты представлены на рис. 4 (числовые значения приведены в СИ) для вращательного сочленения между первым и вторым звеньями манипулятора

(кинематическая схема на рис. 3) и на рис. 5 для поступательного сочленения между вторым и третьим звеньями.

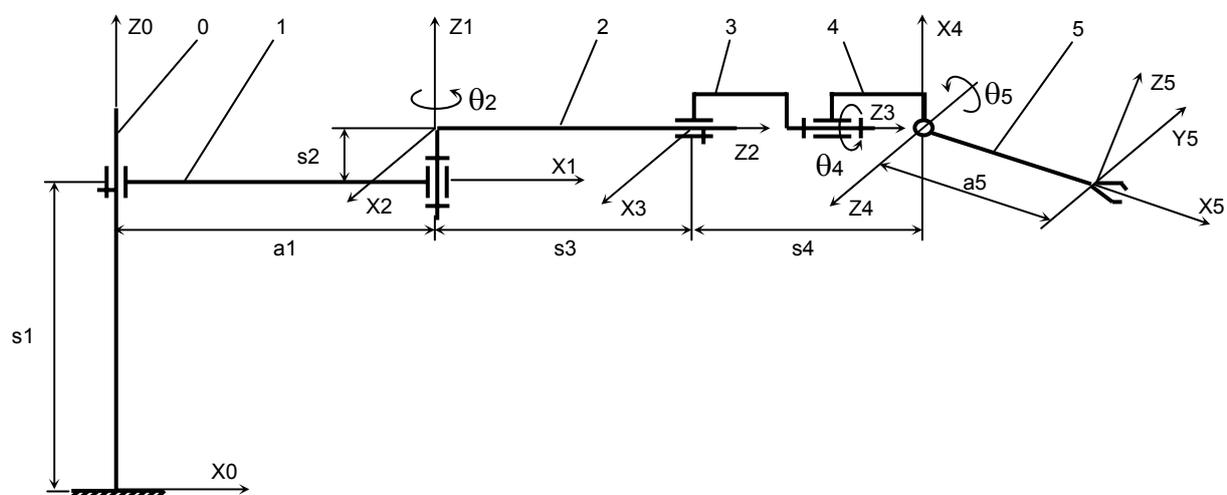


Рис. 3. Кинематическая схема пятизвенного робота-манипулятора, использованная при моделировании (с нумерацией звеньев, обозначениями параметров Денавита-Хартенберга и осей координат, связанных со звеньями)

Литература

- [1] Тимофеев А. В., Богданов А. А. Синтез нейросетевых регуляторов для оптимального управления роботами и мехатронными системами // Робототехника и мехатроника: периодич. науч.-тех. сб. тр., Изд-во РАН, РНК по ТММ РАН, БГТУ, вып. 1. — М.—СПб, 1996. — с. 115-125.
- [2] Болтянский В. Г. Математические методы оптимального управления. — М.: Наука, 1969. — 408 с.
- [3] Фу К., Гонсалес Р., Ли К. Робототехника. — М.: Мир, 1989. — 620 с.
- [4] Тимофеев А. В. Управление роботами. — Л.: Изд-во ЛГУ, 1986. — 239 с.
- [5] Haykin Simon S. Neural Networks: A Comprehensive Foundation. — Prentice Hall, 1998. — 842 p.

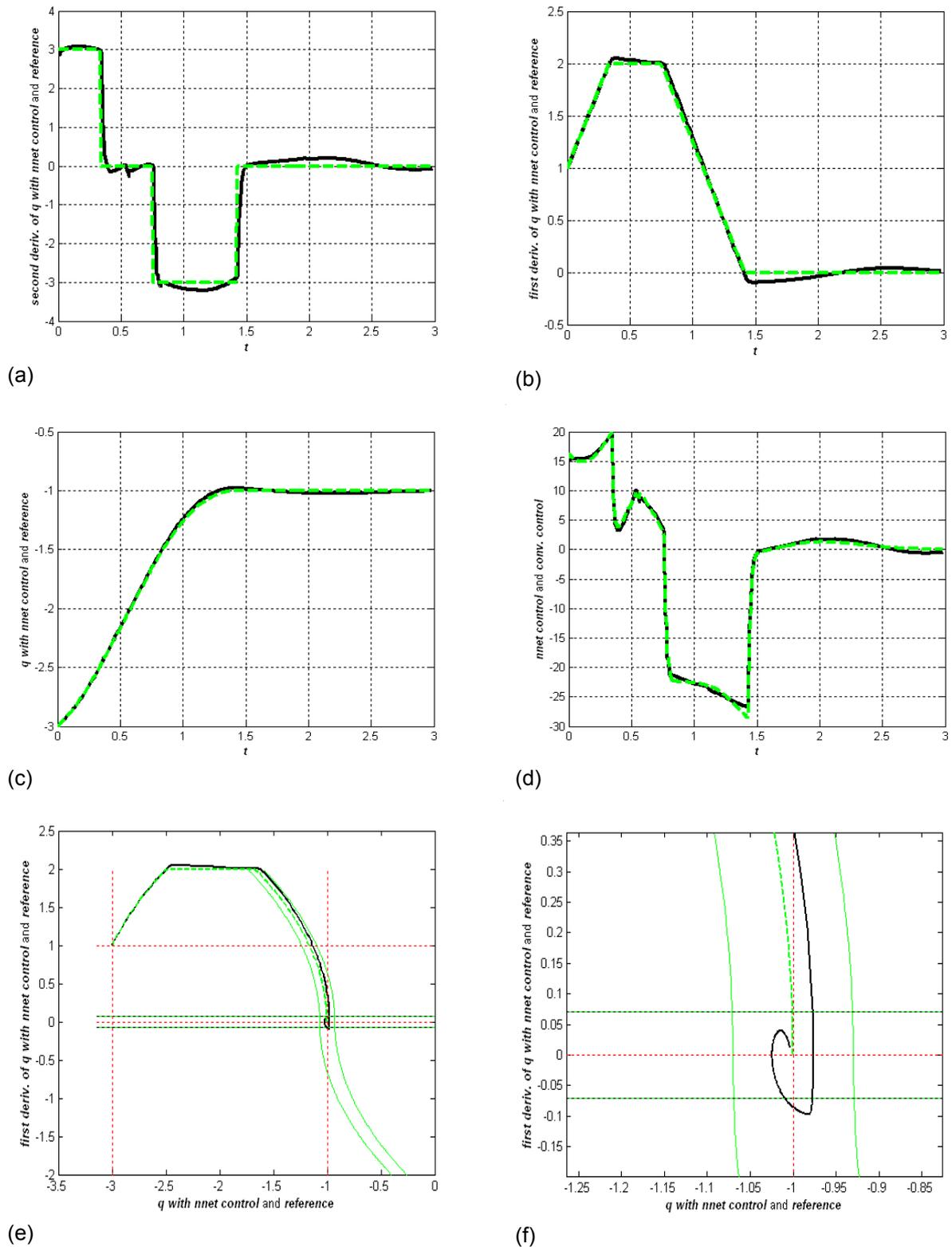


Рисунок 4. Результаты моделирования. (a), (b) и (c) — обобщенное ускорение, скорость и координата от времени (штрихом — желаемые значения), (d) — управление от времени (штрихом — с точным вычислением ОМД), (e) и (f) — траектория на фазовой плоскости (f — увеличено, штрихом — оптимальная траектория).

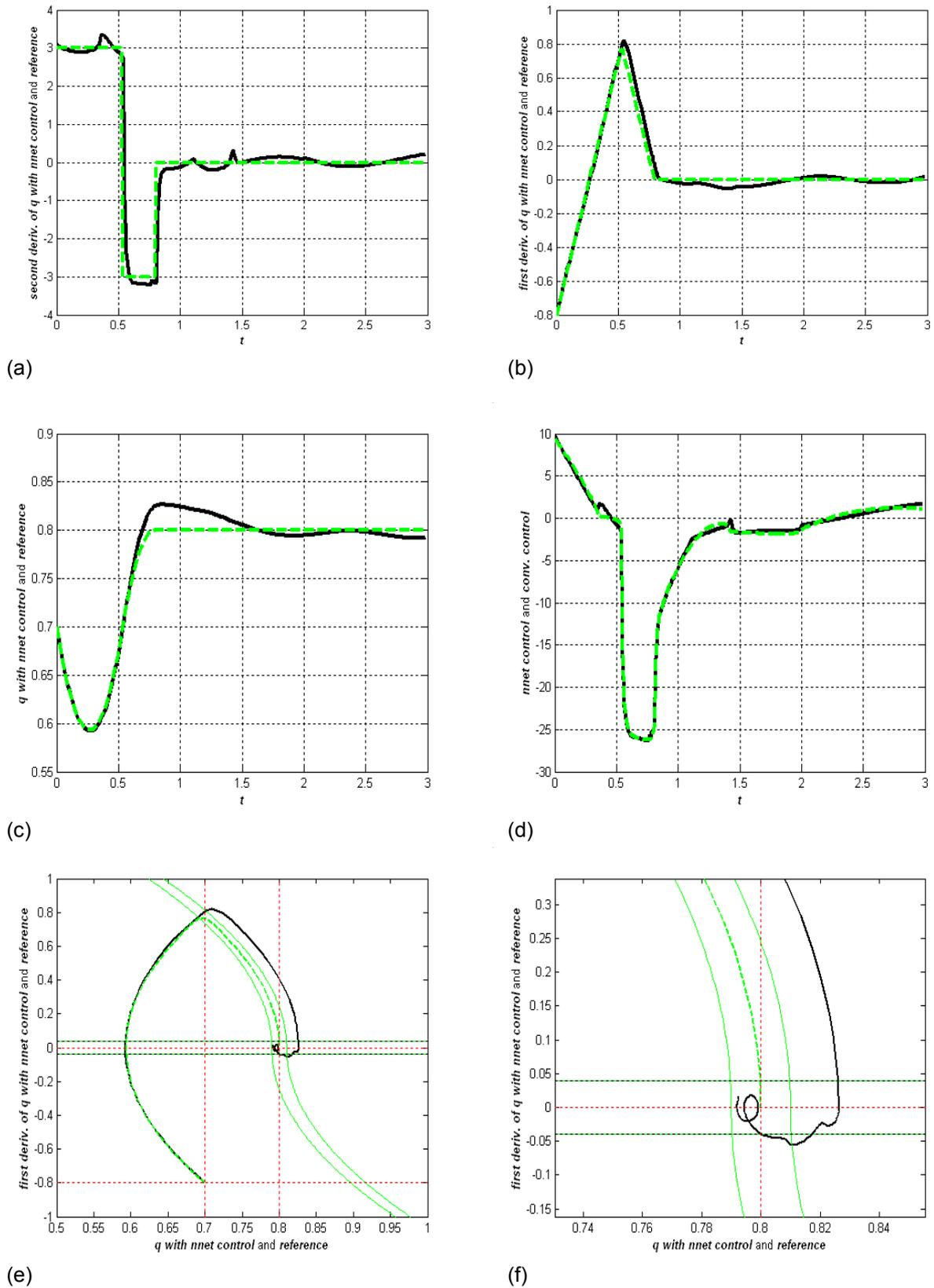


Рисунок 5. Результаты моделирования. (a), (b) и (c) — обобщенное ускорение, скорость и координата от времени (штрихом — желаемые значения), (d) — управление от времени (штрихом — с точным вычислением ОМД), (e) и (f) — траектория на фазовой плоскости (f — увеличено, штрихом — оптимальная траектория).