

ГЛАВА 7 • МЕТОД ПОСЛЕДОВАТЕЛЬНЫХ ПРИБЛИЖЕНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

§ 1. Простейший вариант метода

Пусть управляемая система описывается уравнениями с начальными условиями и ограничениями

$$dx/dt = f(t, x, u), \quad x(t_0) = x^0, \quad u(t) \in U, \quad t \geq t_0. \quad (7.1.1)$$

Здесь $x = (x_1, \dots, x_n)$ — n -мерный вектор фазовых координат, $u = (u_1, \dots, u_m)$ — m -мерный вектор управляющих функций, t — время, $f = (f_1, \dots, f_n)$ — заданная вектор-функция, x^0 — постоянный вектор, t_0 — начальный момент времени, U — замкнутое множество m -мерного пространства. Допустимыми управлениями будем называть кусочно-непрерывные функции $u(t)$, удовлетворяющие ограничению (7.1.1).

Поставим задачу об определении допустимого управления $u(t)$, минимизирующего функционал

$$J = (c, x(T)), \quad T > t_0. \quad (7.1.2)$$

Здесь T — заданный момент времени, $c = (c_1, \dots, c_n)$ — ненулевой постоянный вектор. Скобками обозначено скалярное произведение векторов. Будем предполагать, что поставленная задача имеет решение в классе допустимых управлений $u(t)$, это решение будем называть оптимальным управлением. Введем дополнительных фазовых координат широкий класс функционалов приводится к виду (7.1.2).

Введем n -мерный вектор $p = (p_1, \dots, p_n)$ сопряженных переменных (импульсов) и функцию Гамильтона H ; запишем сопряженную систему и условия трансверсальности:

$$H(t, x, p, u) = (p, f(t, x, u)), \quad (7.1.3)$$

$$dp_i/dt = -\partial H/\partial x_i = -\sum_{j=1}^n p_j \partial f_j(t, x, u)/\partial x_i, \quad p(T) = -c. \quad (7.1.4)$$

Согласно принципу максимума [75], искомое оптимальное управление доставляет функции H из (7.1.3) максимум по $u \in U$ при любом $t \in [t_0, T]$, если x и p удовлетворяют уравнениям и граничным условиям (7.1.1), (7.1.4).

Для решения поставленной задачи оптимального управления в работе [10] был предложен метод последовательных приближений — простейший вариант которого состоит в следующем. В каче-

стве начального приближения $u^{(1)}(t)$ зададим некоторое допустимое управление; при его выборе обычно используются различные физические соображения. Метод состоит из последовательных итераций; k -я итерация заключается в следующем ($k = 1, 2, \dots$):

- 1) решая задачу Коши (7.1.1) с управлением $u = u^{(k)}(t)$, определяем траекторию $x = x^{(k)}(t)$ на интервале $[t_0, T]$;
- 2) решая задачу Коши (7.1.4) «справа налево» от момента $t = T$ до $t = t_0$ при $u = u^{(k)}(t)$, $x = x^{(k)}(t)$, определяем сопряженные переменные $p^{(k)}(t)$ на интервале $[t_0, T]$;
- 3) определяем управление $u^{(k+1)}(t)$ на интервале $[t_0, T]$ из условия

$$H(t, x^{(k)}(t), p^{(k)}(t), u^{(k+1)}(t)) = \max_{u \in U} H(t, x^{(k)}(t), p^{(k)}(t), u). \quad (7.1.5)$$

Если это условие определяет $u^{(k+1)}(t)$ неединственным образом, то выбираем любое из возможных значений. После этого переходим к следующей итерации и т. д.

Если процесс последовательных приближений сходится, то продолжим его до тех пор, пока последующие приближения не будут отличаться друг от друга в пределах заданной точности. Полученное после сходимости решение будет удовлетворять принципу максимума.

Описанный простейший вариант итерационного процесса, конечно, далеко не всегда сходится. Если функции f_i линейны по x_j и имеют вид

$$\dot{x}_i = \sum_{j=1}^n a_{ij}(t) x_j + b_i(t, u), \quad i = 1, \dots, n$$

(где a_{ij}, b_i — заданные функции), то метод уже на второй итерации дает оптимальную траекторию при любом начальном приближении.

В самом деле, здесь x и u не войдут в сопряженную систему (7.1.4). Поэтому решение задачи Коши (7.1.4) для $p(t)$ не будет зависеть от выбора начального приближения, и первое приближение $p^{(1)}(t)$ будет совпадать с искомым вектором сопряженных переменных $p(t)$ вдоль оптимальной траектории. Далее, в рассматриваемом случае функция H из (7.1.3) представится в виде $H = H_1 + H_2$, где H_1 не зависит от u , а H_2 не зависит от x и зависит лишь от t, p, u . Следовательно, при максимизации H по u , согласно (7.1.5), полученное управление $u^{(2)}(t)$ будет искомым оптимальным управлением (это управление может быть и неединственным). Таким образом, доказано, что в рассматриваемом случае метод сходится уже на второй итерации.

Сходимость изложенного метода для одного класса линейных задач с квадратичным функционалом была показана в работе [11]

Некоторые итерационные методы, отличные от описанного выше, но близкие по идеям, рассмотрены в книге [31].

Прежде чем переходить к приемам улучшения сходимости, отметим некоторые достоинства описанного алгоритма, которые желательно по возможности сохранить при его модификациях.

1. Метод сравнительно прост для программирования и не требует большой памяти машины. В процессе счета достаточно хранить лишь таблицу значений управляющей функции $u(t)$ и траекторию $x(t)$ для текущего приближения. Сопряженные переменные запоминать не нужно, так как они используются лишь при вычислении нового управления, согласно (7.1.5). Это вычисление можно проводить на каждом шаге интегрирования сопряженной системы. Вообще говоря, можно не хранить в памяти и траекторию $x(t)$, запоминая лишь конечное значение $x(T)$ и интегрируя затем систему (7.1.1) одновременно с (7.1.4) от $t = T$ до $t = t_0$. При этом, однако, увеличивается объем вычислений. Таким образом, с точки зрения потребной памяти и времени счета, одна итерация метода последовательных приближений соответствует одной итерации градиентных методов (см. [96, 97], обзор этих методов см. в [28]). Отличие состоит в операции (7.1.5), но она обычно сводится к вычислению по замкнутым формулам, так как зависимость функции f от u и область U имеют достаточно простой вид во многих реальных задачах.

2. При применении описанного метода на каждой итерации главная часть приращения функционала отрицательна и максимальна по абсолютной величине среди главных частей приращений, получаемых при всевозможных допустимых вариациях управления [46]. Это следует из формулы [78] для приращения δJ функционала (7.1.2), которое он получает, если допустимое управление $u(t)$ заменить на другое допустимое управление $u(t) + \delta u(t)$; формула для δJ имеет вид

$$\delta J = \delta_1 J + \varphi, \quad \delta_1 J = - \int_{t_0}^T [H(t, x(t), p(t), u + \delta u) - H(t, x(t), p(t), u)] dt. \quad (7.1.6)$$

Здесь $\delta_1 J$ — главная часть приращения. Для остаточного члена справедлива оценка [78]

$$|\varphi| \leq A \|\delta u\|^2, \quad \|\delta u\|^2 = \int_{t_0}^T \sum_{k=1}^m |\delta u_k(t)|^2 dt, \quad (7.1.7)$$

где постоянная A зависит от констант Липшица для функций f_i и их производных. Функции H , $x(t)$, $p(t)$ в формуле (7.1.6) определены соотношениями (7.1.3), (7.1.1), (7.1.4) для $u = u(t)$.

Из формул (7.1.6), (7.1.7) сразу следует сформулированное выше свойство алгоритма: главная часть приращения $\delta_1 J$ отрицательна и максимальна по абсолютной величине. При этом, если остаточный член φ мал по сравнению с $\delta_1 J$, то $\delta_1 J < 0$, а функционал в ходе итераций будет убывать. Отметим, что в отличие от градиентных методов [28, 96, 97], основанных на формулах классического вариационного исчисления, описанный метод на принципе максимума и формулах (7.1.6), (7.1.7) включает игольчатые вариации (малость φ не требует малости абсолютной величины $\delta u(t)$).

3. Пусть функции f_i и множество U таковы, что максимум H по u достигается всегда на границе множества U . Это будет так, например, если f линейно по u , а U — замкнутая ограниченная область. Тогда, если исключить особые управления, оптимальное управление будет граничным (в частности, сюда относятся разрывные релейные управления). Очевидно, что при применении описанного метода на каждой итерации будет получаться граничное управление, т. е. приближение к оптимальному управлению будет идти в классе разрывных граничных или релейных управлений, что представляется удобным. При использовании градиентных методов [28, 96, 97] картина, вообще говоря, будет иной: релейное управление будет аппроксимироваться непрерывными управлениями.

§ 2. Способы улучшения сходимости

Рассмотрим некоторые приемы [48], позволяющие добиться сходимости итераций или ускорить их сходимость в тех случаях, когда простейший вариант метода из § 1 расходится или сходится медленно. Для простоты будем по-прежнему иметь в виду задачу со свободным правым концом (7.1.1) — (7.1.2).

Обозначим через Φ оператор, который каждому допустимому управлению $u(t)$ ставит в соответствие новое приближение для управления по схеме § 1. Вычисление этого оператора сводится к решению двух задач Коши для систем (7.1.1), (7.1.4) и к определению управления из условия (7.1.5). Тогда простейший вариант итерационного процесса из § 1 можно записать в виде

$$u^{(k+1)}(t) = \Phi(u^{(k)}(t)), \quad k = 1, 2, \dots \quad (7.2.1)$$

1. Первый способ улучшения сходимости состоит в том, что в систему (7.1.1) вводится некоторый параметр ε так, чтобы при $\varepsilon = 1$ получалась исходная система, а при $\varepsilon \rightarrow 0$ итерационный процесс (7.2.1) быстро сходиллся. Например, можно ввести параметр одним из следующих способов:

$$dx/dt = \varepsilon f(t, x, u), \quad dx/dt = f(t, x, \varepsilon u), \quad 0 \leq \varepsilon \leq 1. \quad (7.2.2)$$

При $\varepsilon \rightarrow 0$ системы (7.2.2) становятся слабо управляемыми, и поэтому итерационный процесс (7.2.1) будет быстро сходиться [при тех же начальных условиях, ограничениях и функционале (7.1.4), (7.1.2)]. Это следует из того, что в обозначениях (7.1.6), (7.1.7) будем иметь по порядку величин $|\delta_1 J| \sim \varepsilon$, $|\varphi| \sim \varepsilon^2$, т. е. $|\varphi| \ll |\delta_1 J|$ и $\delta J \approx \delta_1 J$.

Получив сходимостъ при некотором ε , будем затем постепенно увеличивать ε , беря каждый раз в качестве начального приближения для итерационного процесса оптимальное управление для предыдущего значения ε . Дойдя таким образом шагами по ε до значения $\varepsilon = 1$, получим решение исходной задачи. Отметим, что при достаточно малом ε в уравнениях (7.2.2) приближенное решение задачи оптимального управления можно получить и не прибегая к процессу итераций. Задачи оптимального управления для систем с малым параметром, подобных (7.2.2), были рассмотрены в работе [95], где дан способ их приближенного аналитического решения. Возможны, конечно, и другие способы введения параметра, помимо (7.2.2). Например, в задачах управляемого полета с аэродинамическим управлением параметр ε вводился обычно в виде коэффициента при аэродинамических силах.

2. Вторая группа способов состоит в том, что на основе итерационного процесса (7.2.1) строятся другие процессы, в которых новое приближение для управления в каждый момент времени выражается через две или более предыдущие итерации. Простейшим вариантом является следующая схема итераций:

$$u^{(k+1)}(t) = (1 - \alpha)u^{(k)}(t) + \alpha\Phi(u^{(k)}(t)), \quad t_0 \leq t \leq T, \quad 0 \leq \alpha \leq 1, \\ k = 1, 2, \dots, \quad (7.2.3)$$

применяемая в случае выпуклого множества U . Здесь постоянную α можно выбирать из условия минимума функционала по α на данной итерации, как в методе скорейшего спуска, а можно ограничиться требованием монотонного убывания функционала на каждой итерации. При подборе α требуется пересчитывать только систему (7.1.1) для вычисления функционала. Неплохие результаты давал описанный здесь способ, в котором α на каждой итерации последовательно принимало значения $1, 1/2, 1/4, \dots$, и т. д., пока не наступало уменьшение функционала. Применялся в расчетах также метод Эйткена (см. [18], т. II, стр. 146—148), в котором каждая четвертая итерация строится на основе трех предыдущих. Однако при этом нужно хранить в памяти машины больше информации — таблицы значений управляющих функций для нескольких приближений.

† Отметим, что при применении этой группы способов приближение к релейному оптимальному управлению не будет идти в

классе релейных управлений, так что свойство 3 из § 1 здесь не имеет места.

3. В соотношении (7.2.3) величину α можно считать функцией времени, ограниченной неравенством $0 \leq \alpha(t) \leq 1$ на интервале $[t_0, T]$. Наиболее простой способ, обладающий к тому же свойством 3 из § 1 и не требующий выпуклости множества U , получится, если считать $\alpha(t)$ кусочно-постоянной функцией, принимающей значения 0 и 1. Так приходим к третьему способу улучшения сходимости, который описывается итерационным процессом

$$u^{(k+1)}(t) = \Phi(u^{(k)}(t)), \quad t \in [t', t''], \quad (7.2.4)$$

$$u^{(k+1)}(t) = u^{(k)}(t), \quad t \in [t', t''], \quad k = 1, 2, \dots$$

Здесь интервал $[t', t'']$, входящий в интервал $[t_0, T]$, выбирается так, чтобы обеспечить условие

$$J^{(k+1)} < J^{(k)}, \quad (7.2.5)$$

где $J^{(k)}, J^{(k+1)}$ — значения функционала, отвечающие управлениям $u^{(k)}, u^{(k+1)}$. Подбор чисел t', t'' на каждой итерации происходит методом проб с проверкой условия (7.2.5). Вначале полагаем $t' = t_0, t'' = T$. Если на какой-то итерации условие (7.2.5) нарушено, сокращаем интервал, пока оно не будет выполнено. Если при некоторых значениях t', t'' изменение функционала несущественно, т. е. $|J^{(k)} - J^{(k+1)}| < \varepsilon_0$, где ε_0 — малое число, характеризующее точность решения, то можно испробовать увеличение интервала $[t', t'']$. Процесс итераций заканчивается либо когда на полном интервале ($t' = t_0, t'' = T$) получено несущественное (менее чем на ε_0) изменение функционала, либо когда даже при малом интервале $[t', t'']$, длина которого равна шагу таблицы управляющих функций, получается увеличение функционала.

Первый случай означает сходимостъ, а второй случай означает, что шаг таблицы слишком велик, и счет надо производить с более мелким шагом. Изменение величин $[t', t'']$ в процессе их подбора можно проводить по различным алгоритмам. Проще и экономичнее всего положить $t'' = T$ и управлять одним числом t' , изменяя его в соответствии с описанной выше логикой. При этом пересчет траектории, т. е. интегрирование системы (7.1.1), достаточно проводить лишь на интервале $[t', T]$, на котором происходит изменение управления согласно (7.2.4).

4. Все три способа улучшения сходимости содержат некоторые параметры ($\varepsilon, \alpha, t', t''$), регулировкой которых можно добиваться сходимости. Если устремить $\varepsilon \rightarrow 0, \alpha \rightarrow 0, t'' - t' \rightarrow 0$, то все три способа будут обеспечивать малость остаточного члена φ из формулы (7.1.7). В другом предельном случае ($\varepsilon = 1, \alpha = 1, t' = t_0, t'' = T$) все эти способы переходят в простейший вариант

метода из § 1, который будет наиболее эффективным вблизи минимума. Процесс итераций во всех случаях следует оканчивать тогда, когда две последующие итерации простейшего варианта (7.2.1) будут практически совпадать.

Расчеты многих конкретных задач показали, что с помощью каждого из описанных способов часто удается добиться сходимости в тех случаях, когда простейший вариант метода расходится. По-видимому, наиболее простым и общим является третий способ, который сочетается с методом последовательных приближений, не нарушая свойства 3 из § 1.

Описанные способы улучшения сходимости можно сочетать друг с другом. В приводимой в главе 10 программе, взятой из работы [48], применяются второй и третий способы, т. е. полагается

$$u^{(k+1)}(t) = \begin{cases} u^{(k)}(t), & 0 \leq t < t', \\ (1 - \alpha)u^{(k)}(t) + \alpha\Phi(u^{(k)}(t)), & t' \leq t \leq T. \end{cases} \quad (7.2.6)$$

Числа α и t' подбираются на каждой итерации так, чтобы обеспечить условие (7.2.5). Если оно не выполняется при каких-то α , t' , то сначала α уменьшается делением пополам до заданного предельного значения. Если это не приводит к выполнению условия (7.2.5), то интервал $[t', T]$ сокращается вдвое, и снова α начинает последовательно принимать значения $1, 1/2$ и т. д. В случае выполнения условия (7.2.5) на следующей итерации принимается сначала $\alpha = 1$, а интервал $[t', T]$ либо сохраняется прежним, либо увеличивается по определенному правилу (если уменьшение функционала не превосходило заданной точности ϵ_0).

Программа главы 10 реализует, таким образом, довольно гибкий алгоритм, в котором сочетаются вложенные друг в друга второй и третий способы улучшения сходимости. При этом второй способ может быть вообще отключен по желанию программиста (см. главу 10).

Первый способ также нетрудно применять, используя программу главы 10 для решения задачи для системы типа (7.2.2) при каждом фиксированном ϵ .

§ 3. Некоторые обобщения

1. Метод последовательных приближений вместе с его модификациями может быть применен для более широкого класса задач, чем те, что рассмотрены выше. Пусть управляемая система по-прежнему описывается соотношениями (7.1.1), а условия в конце процесса и минимизируемый функционал имеют вид

$$\begin{aligned} h(T, x(T)) = 0, \quad g_i(T, x(T)) = 0, \quad i = 1, \dots, r; \\ J = F(T, x(T)), \quad 0 \leq r < n. \end{aligned} \quad (7.3.1)$$

Здесь h, g_i, F — заданные функции. Первое равенство (7.3.1) служит условием, определяющим момент T окончания процесса, причем в качестве T берется первый момент времени $T > t_0$, при котором удовлетворяется это условие. Равенства $g_i = 0$ из (7.3.1) представляют собой дополнительные краевые условия, которые могут и отсутствовать (если $r = 0$). Случай интегрального функционала легко сводится к задаче (7.3.1) путем введения дополнительной фазовой координаты (см., например, [75]).

Для задачи оптимального управления (7.1.1), (7.3.1) сопряженные переменные по-прежнему удовлетворяют системе (7.1.4), функция Гамильтона имеет прежний вид (7.1.3), а искомое оптимальное управление доставляет максимум по $u \in U$ функции Гамильтона H при всех $t \in [t_0, T]$. Вместо условия (7.1.4) для $p(t)$ будем иметь следующие условия трансверсальности при $t = T$

$$p(T) = \left(dF/dt - \sum_{i=1}^r \lambda_i dg_i/dt \right) (dh/dt)^{-1} dh/\partial x + \sum_{i=1}^r \lambda_i \partial g_i/\partial x - \partial F/\partial x, \quad (7.3.2)$$

$d/dt = \partial/\partial t + (f, \partial/\partial x)$, которые легко выводятся из общих условий (см. [25, 75, 78]). Здесь $\partial/\partial x$ — оператор градиента по фазовым координатам x , d/dt — полная производная вдоль траекторий системы (7.1.1). Момент T и параметры λ_i в (7.3.2) неизвестны и должны определяться при помощи краевых условий (7.3.1). Предполагается, что полная производная dh/dt в нуль не обращается.

2. Если $r = 0$, то суммы, содержащие множители λ_i , в соотношении (7.3.2) нужно опустить. В этом случае метод последовательных приближений применяется почти без изменений. Интегрирование системы (7.1.1) идет до выполнения условия $h = 0$, определяющего момент T в данном приближении. После этого вычисляется $p(T)$ по формуле (7.3.2), причем производные d/dt в точке T проще всего определять через конечно-разностные отношения по последнему шагу интегрирования системы (7.1.1). Затем интегрируется система (7.1.4) от момента T до t_0 и определяется новое управление из условия (7.1.5). При счете следует лишь иметь в виду, что длительность процесса от итерации к итерации может возрастать, поэтому нужно на этот случай доопределять управление при $t > T$, полагая, например, $u(t) = u(T)$ при $t \geq T$. Именно такое доопределение предусмотрено в программе, приводимой в главе 10.

3. Если $r > 0$, то простейший способ учета краевых условий заключается в применении метода штрафных функций, который широко используется в градиентных методах (см. [28, 96]).

Вводим новый функционал J' , полагая

$$J' = F(T, x(T)) + \sum_{i=1}^r a_i [g_i(T, x(T))]^2, \quad a_i > 0, \quad i = 1, \dots, r \quad (7.3.3)$$

и опуская краевые условия $g_i = 0$. Задача о минимуме функционала (7.3.3) при единственном условии $h = 0$ в конце процесса будет задачей с $r = 0$, которая решается по описанной выше схеме. После сходимости итераций каждый раз увеличиваем постепенно коэффициенты штрафов a_i , пока краевые условия не будут удовлетворены с требуемой точностью.

4. Другой способ учета краевых условий также родственен тому, что применяется в градиентных методах (см. [28, 97]), но основан на формуле (7.1.6) для вариации функционала. Пусть $u(t)$ — некоторое допустимое управление, $x(t)$ — соответствующая ему, согласно системе (7.1.1), траектория, момент T определен из первого уравнения (7.3.1). Обозначим через $q(t)$, $q_i(t)$ решения сопряженной системы (7.1.4) при выбранных $u(t)$ и $x(t)$, удовлетворяющие следующим начальным условиям в момент T :

$$q(T) = \frac{dF}{dt} \left(\frac{dh}{dt} \right)^{-1} \frac{\partial h}{\partial x} - \frac{\partial F}{\partial x}, \quad t = T; \quad (7.3.4)$$

$$q_i(T) = - \frac{dq_i}{dt} \left(\frac{dh}{dt} \right)^{-1} \frac{\partial h}{\partial x} + \frac{\partial g_i}{\partial x}, \quad i = 1, \dots, r.$$

Запишем главную часть приращения функционалов $J = F(T, x(T))$ и $g_i(T, x(T))$, которую они получают при замене управления $u(t)$ на другое допустимое управление $u(t) + \delta u(t)$. Эти вариации, как нетрудно проверить, будут задаваться формулами, аналогичными (7.1.6):

$$\delta_1 J = - \int_{t_0}^T (q, f(t, x, u + \delta u) - f(t, x, u)) dt, \quad (7.3.5)$$

$$\delta_1 g_i = - \int_{t_0}^T (q_i, f(t, x, u + \delta u) - f(t, x, u)) dt, \quad i = 1, \dots, r,$$

в которых функции $q(t)$ и $q_i(t)$ удовлетворяют сопряженной системе (7.1.4) и условиям (7.3.4). Из формул (7.3.2), (7.3.4) вытекает, в силу линейности системы (7.1.4),

$$p(t) = q(t) + \sum_{i=1}^r \lambda_i q_i(t), \quad t_0 \leq t \leq T. \quad (7.3.6)$$

Сопоставляя равенства (7.3.5), (7.3.6) и (7.1.3), получим

$$\delta_1 J + \sum_{i=1}^r \delta_1 g_i = - \int_{t_0}^T [H(t, x, p, u + \delta u) - H(t, x, p, u)] dt. \quad (7.3.7)$$

Теперь можно предложить следующий алгоритм. Пусть $u(t)$ — некоторое допустимое управление. Опишем одну итерацию метода.

1. Решаем задачу Коши (7.1.1) с управлением $u(t)$, определяем траекторию $x(t)$, момент T из первого условия (7.3.1) и величины $g_i(T, x(T))$, $F(T, x(T))$.

2. Интегрируем $r + 1$ раз «справа налево» сопряженную систему (7.1.4) с начальными условиями (7.3.4). Получим $r + 1$ вектор-функций $q(t)$, $q_i(t)$ на интервале $[t_0, T]$.

3. Задаемся желаемыми значениями вариаций $\Delta_i = \delta_1 g_i$, $i = 1, \dots, r$. Проще всего положить $\Delta_i = -g_i$, стремясь уже в следующей итерации компенсировать ошибку удовлетворения краевых условий, или задать $\Delta_i = -\alpha_i g_i$, где числа α_i лежат в пределах $0 < \alpha_i \leq 1$, $i = 1, \dots, r$.

4. Подбираем постоянные λ_i так, чтобы $\delta_1 g_i = \Delta_i$ при $i = 1, \dots, r$, одновременно определяя новое управление $u + \delta u$. Для этого, задавшись набором λ_i , по формуле (7.3.6) определяем $p(t)$, затем находим $u + \delta u$ из условия максимума H из (7.1.3) по $u + \delta u \in U$ в каждый момент t , после чего вычисляем квадратуры (7.3.5) для $\delta_1 g_i$. Таким образом, $\delta_1 g_i$ есть в конечном счете нелинейные функции от λ_i , и определение λ_i из условий $\delta_1 g_i = \Delta_i$ сводится к решению системы трансцендентных уравнений. Подбор λ_i можно производить, например, по методу Ньютона или другим численным способом. На каждом шаге этого способа определяется управление $u + \delta u$ и пересчитываются квадратуры (7.3.5). Определив λ_i , найдем тем самым также и новое управление $u + \delta u$, после чего переходим к п. 1 и так продолжаем до сходимости итераций.

Если на каком-то шаге все краевые условия удовлетворены, то постоянные λ_i будут определяться из условий $\delta_1 g_i = \Delta_i = 0$ при $i = 1, \dots, r$. Подынтегральное выражение в (7.3.7) неотрицательно, так как $u + \delta u$ максимизирует H в каждый момент t . Тогда из (7.3.7) следует $\delta_1 J \leq 0$. Описанный метод в этом случае обладает всеми свойствами 1–3 простейшего варианта из § 1 и является его непосредственным обобщением. Если норма (7.1.7) приращения δu достаточно мала, то метод будет обеспечивать уменьшение функционала при сохранении краевых условий, т. е. $\delta J \approx \delta_1 J \leq 0$, $\delta g_i \approx \delta_1 g_i = 0$ с точностью до малых высшего порядка относительно $\|\delta u\|$.

Если метод не сходится (функционал возрастает или краевые условия не выполняются), то можно применить, например, третий способ улучшения сходимости из § 2. Введя интервал $[t', t''] \subset [t_0, T]$, будем производить варьирование управления только на этом интервале, полагая $\delta u = 0$ вне интервала $[t', t'']$. Если же на какой-то итерации не сходится процесс подбора чисел λ_i (см. п. 4 описанного алгоритма), можно уменьшить коэффициенты α_i в п. 3.

Так как описанный способ включает на каждой итерации дополнительный итерационный процесс подбора λ_i , то он может быть рациональным лишь при небольших r . В программе главы 10 для удовлетворения краевых условий использован метод штрафных функций, который обладает большей простотой.

5. Если задача оптимального управления содержит фазовые ограничения, то их можно учесть обычным методом штрафных функций (см., например, [28, 96]). Так, пусть оптимальная траектория $x(t)$ должна удовлетворять неравенству

$$G(x(t), t) \leq 0, \quad t_0 \leq t \leq T \quad (7.3.8)$$

на всем интервале движения. Здесь $G(x, t)$ — заданная скалярная функция. Вводим новую фазовую координату соотношениями:

$$dx_{n+1}/dt = \chi(G(x, t)), \quad x_{n+1}(t_0) = 0. \quad (7.3.9)$$

Здесь $\chi(G)$ — функция, обладающая следующими свойствами: она достаточно гладкая при всех G ; $\chi(G)$ равно или близко к нулю при $G \leq 0$; $\chi(G)$ принимает большие значения при $G > 0$ и быстро растет с ростом G . Например, можно принять

$$\chi(G) = Ae^{bG}, \quad (7.3.10)$$

где A, b — достаточно большие положительные постоянные. Нетрудно видеть, что величина фазовой координаты x_{n+1} в конечный момент T будет, согласно (7.3.9), некоторой интегральной мерой нарушения условия (7.3.8): чем сильнее нарушено это условие и чем больше интервал нарушения, тем больше величина $x_{n+1}(T)$. Вместо минимизации исходного функционала J будем минимизировать функционал вида $J'' = J + x_{n+1}(T)$, причем после сходимости итераций будем постепенно изменять параметры функции $\chi(G)$ (например, увеличивать числа A, b в соотношении (7.3.10)), пока условие (7.3.8) не будет выполнено с требуемой точностью на всем интервале.

6. Метод последовательных приближений применим и к некоторым задачам оптимального управления, в которых минимизируемый функционал представляет собой экстремум (по времени) некоторой заданной функции $\psi(t, x)$ от фазовых координат и вре-

мени. Рассмотрим решение подобных задач, следуя работе [49]. Пусть управляемая система по-прежнему описывается соотношениями (7.1.1), а минимизируемый функционал задан в виде

$$J = \min_{t_1 \leq t \leq t_2} \psi(t, x(t)), \quad (7.3.11)$$

где t_1, t_2 — заданные числа ($t_0 \leq t_1 \leq t_2$), ψ — заданная скалярная непрерывно-дифференцируемая функция. Основная трудность задач с функционалами типа (7.3.11) состоит в том, что функция $\psi(t, x(t))$ может иметь при различных допустимых управлениях, вообще говоря, различное и притом сколь угодно большое число экстремумов на заданном интервале $[t_1, t_2]$.

Отметим частные случаи этой задачи, когда экстремум функции ψ на интервале $[t_1, t_2]$ заведомо единственен. Если при любых допустимых управлениях функция $\psi(t, x(t))$ монотонна на интервале $[t_1, t_2]$, то минимум ψ по t достигается при $t = t_1$ или при $t = t_2$, и задача становится обычной задачей оптимального управления со свободным концом и фиксированным временем (см. п. 2 данного параграфа).

Рассмотрим несколько более интересный случай единственного минимума функции ψ внутри интервала $[t_1, t_2]$. Введем функцию

$$f_0(t, x, u) = d\psi(t, x)/dt = \partial\psi/\partial t + (\partial\psi/\partial x, f(t, x, u)). \quad (7.3.12)$$

Здесь использовано обозначение (7.3.2) для полной производной. Сделаем два допущения:

1) функция f_0 из (7.3.12) явно не зависит от управления u , т. е.

$$\partial f_0 / \partial u_i = 0, \quad i = 1, \dots, m \quad (7.3.13)$$

или

$$f_0 = f_0(t, x) \quad (7.3.14)$$

при всех x и $u \in U$;

2) при всех допустимых $u(t)$ функция $f_0(t, x(t))$ на интервале $[t_1, t_2]$ строго возрастает и один раз проходит через нуль.

При сделанных допущениях функция $\psi(t, x(t))$ при любом допустимом управлении имеет на интервале $[t_1, t_2]$ один экстремум и притом минимум. Обозначая точку минимума через T , получим, согласно (7.3.11), (7.3.12), (7.3.14),

$$f_0(T, x(T)) = 0, \quad J = \psi(T, x(T)). \quad (7.3.15)$$

Таким образом, задача свелась к задаче, описанной в п. 2 данного параграфа. В обозначениях (7.3.1), (7.3.2) получим, учи-

тывая соотношения (7.3.15),

$$\begin{aligned} r = 0, \quad h(t, x) = f_0(t, x) = d\psi/dt, \quad F(t, x) = \psi(t, x), \quad J = \\ = F(T, x(T)) = \psi(T, x(T)), \quad p(T) = \frac{d\psi}{dt} \left(\frac{df_0}{dt} \right)^{-1} \frac{\partial f_0}{\partial x} - \frac{\partial \psi}{\partial x} \quad \text{при } t = T. \end{aligned} \quad (7.3.16)$$

Дальнейший ход решения задачи при условиях (7.3.16) будет таким, какой описан в п. 2 данного параграфа. Аналогично может быть рассмотрена задача о максимизации функционала (7.3.11), а также задачи о минимизации или максимизации функционала типа (7.3.11), в котором \min заменен на \max . Кроме того, в момент $t = T$ могут быть заданы дополнительные краевые условия типа (7.3.1); тогда будем иметь $r > 0$, и условия (7.3.16) также изменятся в соответствии с (7.3.2).

Заметим, что второе из сделанных выше допущений можно ослабить, потребовав, чтобы функция $f_0(t, x(t))$ строго возрастала и проходила через нуль на интервале $[t_1, t_2]$ не при всех допустимых управлениях, а лишь при оптимальном управлении и при близких к нему (в смысле метрики L_2) управлениях. При численном решении задач, когда это условие трудно проверить, можно в ряде случаев поступить следующим образом. Предположив, что минимум ψ достигается при $T = t_1$, при $T = t_2$ или при $t_1 < T < t_2$, и решив три задачи оптимального управления, соответствующие этим предположениям, нужно затем отобрать решение, дающее меньшее значение функционала.

Допущения 1, 2, конечно, существенно упрощают рассматриваемые задачи. Отметим, однако, что допущение 1 довольно часто выполняется в задачах механики управляемых движений. Если функция ψ зависит только от обобщенных координат движущегося объекта и не зависит от его скоростей, а управление осуществляется при помощи действующих на объект сил, то функция f_0 из (7.3.12) не будет зависеть от u , и условие (7.3.13) будет выполнено. Примером подобной задачи является задача о минимизации минимального расстояния между управляемой точкой и неуправляемой целью (задача о минимизации промаха). Простейшая задача о минимизации промаха была решена в работе [49]. Другой пример будет дан в § 5.

7. Описанный метод формально может быть применен и для численного определения траекторий дифференциальных игр. В этом случае вместо операции максимума (7.1.5) нужно вычислять минимум по управлениям противников.

8. Наконец, метод последовательных приближений можно использовать и для оптимизации некоторых процессов с распределенными параметрами, описываемых уравнениями в частных

производных. Для возможности применения метода существенно лишь, чтобы для решаемой задачи можно было сформулировать принцип максимума или записать первую вариацию функционала в форме, аналогичной (7.1.6). При этом можно идти двумя путями: либо записать алгоритм метода для задачи с частными производными и затем решать полученные краевые задачи теми или иными методами, либо сразу применить дискретизацию и свести задачу к системе обыкновенных уравнений (например, при помощи метода прямых), после чего она будет решаться в точности так же, как и рассмотренные выше задачи. Таким образом можно численно решать задачи оптимизации управления линейными и нелинейными колебательными, тепловыми и другими процессами. О постановке задач оптимального управления для систем с распределенными параметрами см., например, книгу [20].

§ 4. Задача о максимальной дальности планирования в сопротивляющейся среде

В качестве первой задачи, решенной при помощи метода последовательных приближений, изложим решенную в работе [46] модельную задачу о достижении максимальной дальности планирования аппарата в сопротивляющейся среде.

Рассмотрим плоское движение материальной точки под действием силы тяжести и аэродинамических сил: подъемной силы и силы сопротивления. Процесс управления осуществляется при помощи двух функций, одна из которых непрерывна, а другая — релейного типа; эти функции будут введены ниже. Уравнения движения имеют вид:

$$\begin{aligned} dx/dt = v \cos \theta, \quad dy/dt = v \sin \theta, \quad m \cdot dv/dt = -R - mg \sin \theta, \\ mv \cdot d\theta/dt = Y - mg \cos \theta. \end{aligned} \quad (7.4.1)$$

Здесь t — время, x, y, v, θ — соответственно дальность, высота, скорость точки и угол наклона траектории точки к горизонтальной оси x , m — масса тела, g — ускорение силы тяжести, R и Y — сила сопротивления и подъемная сила (фиг. 7.1). Как обычно, принимаем $R = 0.5 \rho v^2 S c_x$, $Y = 0.5 \rho v^2 S c_y$, где ρ, S, c_x и c_y — плотность среды (постоянная), характерная площадь тела и аэродинамические коэффициенты (c_x — коэффициент сопротивления, c_y — коэффициент подъемной силы). Будем считать, что тело симметрично относительно продольной и поперечной осей. Тогда, очевидно, c_x является четной, а c_y — нечетной периодической функцией угла атаки α с периодом π (угол атаки — угол между вектором скорости и продольной осью тела). Для простоты положим $c_x = A - B \cos 2\alpha$, $A > B > 0$, $c_y = C \sin 2\alpha$, $C > 0$, что означает аппроксимацию полярных (зависимости между c_x и c_y)

эллипсом. При малых значениях угла атаки α эти формулы переходят в общепринятые (c_1, c_2, c_3 — постоянные) $c_x = c_1 + c_2\alpha^2$, $c_y = c_3\alpha$.

Вместо трех постоянных A, B, C введем максимальное качество $K = \max_{\alpha} [c_y(\alpha)/c_x(\alpha)]$ и угол атаки α_0 , при котором оно достигается. Включая третью постоянную множителем в S , окончательно получим после элементарных преобразований:

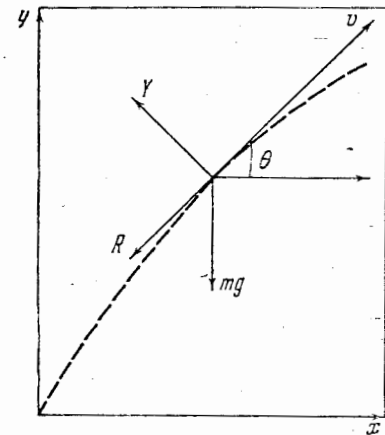
$$c_x = 1 - \cos 2\alpha_0 \cos 2\alpha, \quad c_y = K \sin 2\alpha_0 \sin 2\alpha. \quad (7.4.2)$$

Пусть в процессе движения можно изменять характерную площадь S , придавая ей одно из двух значений (в зависимости от величины η): $S = S_0(1 + \eta b)$, $S_0 > 0, b > 0, \eta = 0; 1$.

Изменение S означает, что тело может изменять свою геометрию в процессе полета (например, увеличивать характерную площадь крыльев).

Угол атаки α и параметр η считаем управляющими функциями, причем первая из них может принимать произвольные значения, а вторая — одно из двух допустимых значений.

Начальные условия берем в виде: $t = 0, x = y = 0, v = v_0, \theta = \theta_0 > 0$. Вариационная задача заключается в определении такого закона управления $\alpha(t)$ и $\eta(t)$, при котором величина дальности



Фиг. 7.1

полета $x(T)$ была бы максимальной в момент T , когда тело снова пересечет ось абсцисс (т. е. при $y(T) = 0$). Таким образом, это — вариационная задача со свободным правым концом и нефиксированным временем.

Перейдем к безразмерным переменным, выбрав в качестве единицы скорости начальную скорость v_0 , единицы времени — величину v_0/g и единицы массы — m . Тогда система уравнений (7.4.1) примет следующий вид (безразмерные переменные обозначены теми же буквами, что и соответствующие им размерные):

$$\begin{aligned} dx/dt &= v \cos \theta, \quad dy/dt = v \sin \theta, \quad dv/dt = -\sigma v^2 c_x (1 + \eta b) - \sin \theta, \\ d\theta/dt &= \sigma v c_y (1 + \eta b) - (\cos \theta)/v. \end{aligned} \quad (7.4.3)$$

Начальные условия при этом преобразуются в следующие:

$$t = 0, \quad x = 0, \quad y = 0, \quad v = 1, \quad \theta = \theta_0. \quad (7.4.4)$$

Здесь σ — безразмерный параметр, характеризующий отношение аэродинамической силы к весу, $\sigma = \rho S v_0^2 / 2mg$.

Запишем функцию Гамильтона (7.1.3) для поставленной вариационной задачи

$$\begin{aligned} H &= p_x v \cos \theta + p_y v \sin \theta - p_v \sigma v^2 c_x (1 + \eta b) - p_\theta \sin \theta + \\ &+ p_0 \sigma v c_y (1 + \eta b) - p_0 \frac{\cos \theta}{v}. \end{aligned} \quad (7.4.5)$$

Максимизируется функционал $J = x(T)$ при условии $y(T) = 0$. В обозначениях (7.3.1) имеем $r = 0, h = y, F = -x$. Условия трансверсальности (7.3.2) здесь примут вид:

$$\begin{aligned} p_x(T) &= 1, \quad p_y(T) = 0, \quad p_\theta(T) = 0, \quad p_v(T) = - \left(\frac{dx/dt}{dy/dt} \right)_T = \\ &= - \operatorname{ctg} \theta(T). \end{aligned}$$

Так как H из (7.4.5) не зависит явно от x и y , то $\dot{p}_x = \dot{p}_y = 0$, поэтому $p_x \equiv 1, p_y \equiv - \operatorname{ctg} \theta(T)$. Для импульсов p_v и p_θ остаются два уравнения:

$$\begin{aligned} dp_v/dt &= - \cos \theta + \operatorname{ctg} \theta(T) \sin \theta + 2p_v \sigma c_x (1 + \eta b) - \\ &- p_0 \sigma c_y (1 + \eta b) - p_0 \frac{\cos \theta}{v^2}, \end{aligned} \quad (7.4.6)$$

$$dp_\theta/dt = v \sin \theta + v \cos \theta \operatorname{ctg} \theta(T) + p_v \cos \theta - p_0 (\sin \theta)/v$$

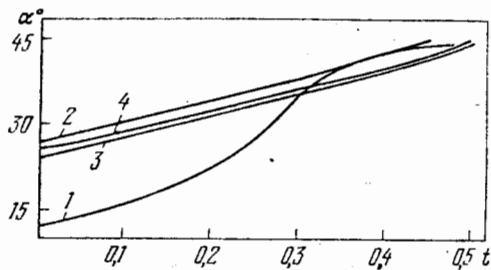
с начальными условиями

$$t = T, \quad p_v = 0, \quad p_\theta = 0. \quad (7.4.7)$$

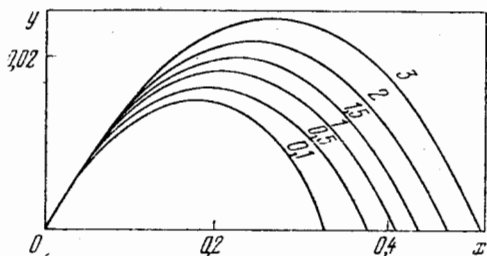
Система уравнений (7.4.3), (7.4.6) при условиях (7.4.4), (7.4.7) решалась описанным в § 1 методом с использованием первого и второго способов улучшения сходимости, описанных в § 2 данной главы. На каждом шаге процесса управляющие функции α, η находились из условия максимальности выражения

$$(1 + \eta b)(p_0 c_y - v p_v c_x), \quad (7.4.8)$$

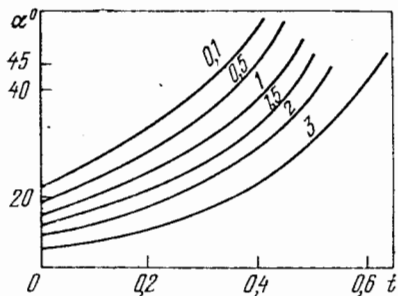
полученного из H (7.4.5) отбрасыванием слагаемых и множителей, не зависящих от управления. Однако при $t = T$, в силу граничных условий, имеем $p_v = p_\theta = 0$, и непосредственное определение α и η из условия (7.4.8) невозможно. Заметим, что из уравнений (7.4.6), с учетом естественного условия $0 > \theta(T) > (-\pi)/2$, следует для момента $t = T: dp_v/dt = 0, dp_\theta/dt = v \sin \theta < 0$. Следовательно, при малых отрицательных значениях разности $t - T$ величина p_θ отрицательна, и по порядку величины много больше, чем p_v . Поэтому при малых $t - T$ условие



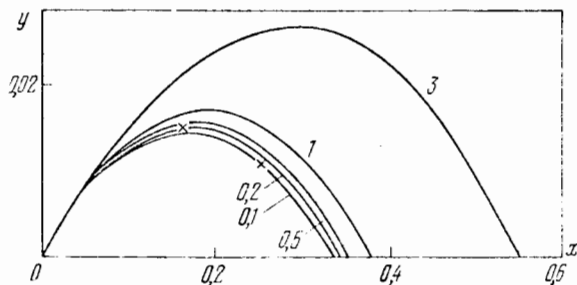
Фиг. 7.2



Фиг. 7.3



Фиг. 7.4



Фиг. 7.5

максимальности (7.4.8) приводится с точностью до малых высшего порядка к условиям $\eta = 1$ и $c_y(\alpha) = c_{y \max}$, что, в силу (7.4.2), соответствует значению $\alpha(T) = \pi/4$. Эти условия вполне понятны физически: для достижения максимальной дальности в последний момент выгоднее увеличить максимально подъемную силу, не заботясь о сопротивлении.

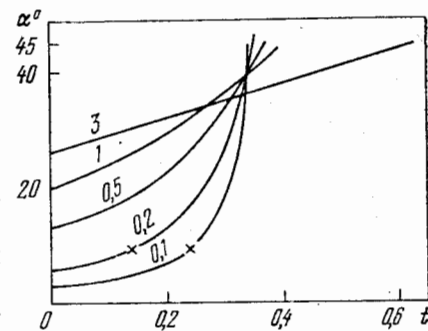
Ниже приводятся результаты численного решения поставленной задачи на ЭВМ для значений параметров: $b = 0.2$, $\alpha_0 = 10^\circ$, $\theta_0 = 10^\circ$.

Параметры K и σ варьировались в широких пределах, причем в качестве первого приближения для какого-либо варианта задавалось оптимальное управление, найденное для предыдущего варианта с близкими значениями K и σ .

На фиг. 7.2 изображен типичный пример сходимости процесса последовательных приближений. На ней представлены зависимости управления $\alpha(t)$ для одного из вариантов, причем номер кривой указывает номер приближения. Пятое и последующие приближения практически не отличались от четвертого. Значения максимизируемого функционала (дальности) для последовательных итераций были следующими: $x_1 = 0.426$, $x_2 = 0.441$, $x_3 = 0.450$, $x_4 = 0.451$.

На фиг. 7.3 представлены оптимальные траектории при различных значениях σ , указанных на графике, и фиксированном значении максимального качества $K = 1$. Оптимальные управления $\alpha(t)$ для этой серии вариантов даны на фиг. 7.4. С увеличением σ (отношения аэродинамических сил к весу) оказывается выгодным уменьшать углы атаки, проигрывая в подъемной силе, но зато добиваясь уменьшения лобового сопротивления. Дальность полета возрастает с увеличением σ . Для всех приведенных здесь вариантов оказалось $\eta(t) \equiv 1$, что соответствует наибольшему из двух возможных значений характерной площади.

Серия оптимальных траекторий при фиксированном значении $\sigma = 0.5$ и различных значениях максимального качества K , указанных цифрами на графике, изображена на фиг. 7.5. Дальность полета, естественно, возрастает с увеличением качества. На фиг. 7.6 представлены соответствующие оптимальные управления. При малых значениях качества K движение происходит на малых углах атаки. Кроме того, здесь на начальной части траектории



Фиг. 7.6

(до точек, указанных звездочкой для $K = 0.1$ и $K = 0.2$) управляющая функция $\eta = 0$ (характерная площадь минимальна). Оба этих фактора уменьшают силу сопротивления (они одновременно уменьшают и подъемную силу, но она незначительна при малых K). Для больших K , наоборот, сила сопротивления становится менее существенной по сравнению с подъемной силой. Поэтому здесь движение по оптимальной траектории происходит при больших углах атаки и с $\eta = 1$.

Отметим, что приближенные решения этой задачи при малых σ были построены методом малого параметра [95], причем получено хорошее совпадение численного и аналитического решений при малом σ .

§ 5. Одна задача об оптимальном маневре в центральном поле

В качестве второго численного примера применения метода последовательных приближений приведем задачу, решенную в работе [49].

Рассмотрим плоское движение спутника (материальной точки) под действием силы притяжения Земли и управляющей реактивной силы. Уравнения движения спутника имеют вид

$$\begin{aligned} \dot{r} &= v \sin \theta, \quad \dot{\varphi} = \frac{v}{r} \cos \theta, \quad \dot{v} = -\frac{1}{r^2} \sin \theta + a \cos \gamma, \quad \dot{\theta} = \frac{v}{r} \cos \theta - \\ & - \frac{1}{vr^2} \cos \theta + \frac{a}{v} \sin \gamma. \end{aligned} \quad (7.5.1)$$

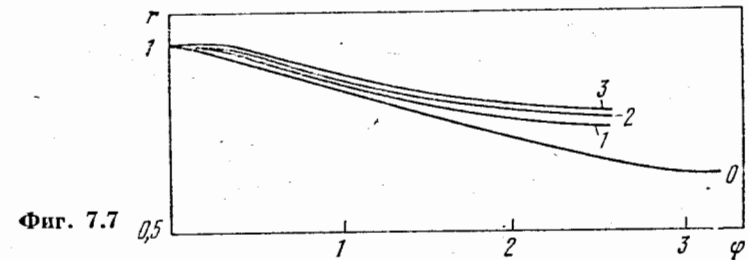
Здесь r, φ — полярные координаты спутника в плоскости орбиты (расстояние от центра Земли и полярный угол), v — модуль его скорости, θ — угол между вектором скорости и перпендикуляром к радиусу-вектору, a — реактивное ускорение, γ — угол между вектором тяги и вектором скорости. Величину a считаем постоянной, а γ — управляющая функция, которая может принимать любые значения. Все величины безразмерны. Соответствующие размерные величины, отмеченные звездочкой, можно получить по формулам:

$$r^* = rr_0^*, \quad v^* = v \sqrt{\mu/r_0^*}, \quad t^* = tr_0^{3/2}/\mu^{1/2}, \quad a^* = a\mu/(r_0^*)^2,$$

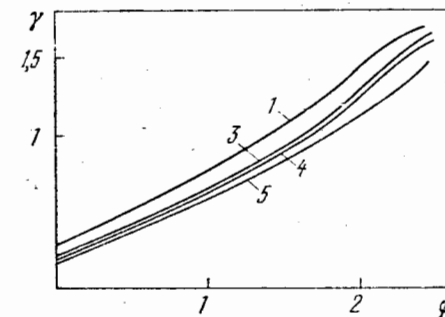
где r_0^* — расстояние от центра Земли до спутника в начальный момент, а μ — произведение массы Земли на постоянную всемирного тяготения.

Известно, что время существования неуправляемого спутника зависит главным образом от высоты его перигея: чем выше перигей орбиты, тем меньше влияние торможения в верхних слоях атмосферы. Поэтому естественно поставить вариационную задачу

о максимальном увеличении высоты перигея за время маневра. Считаем, что время маневра (управляемого движения) не превышает одного оборота, и пренебрегаем изменением массы спутника и влиянием атмосферы за это время. Поскольку перигей — это точка орбиты, где $r(t)$ минимально, то поставленная задача сводится к определению оптимального управления $\gamma(t)$, доставляю-



Фиг. 7.7



Фиг. 7.8

щего максимум функционалу

$$J = \min_t r(t).$$

Задача свелась к постановке, описанной в п. 6 § 3, причем в обозначениях § 3 имеем $\varphi = r$, и допущение 1 [условие (7.3.13) или (7.3.14)] здесь выполнено. Допущение 2 из п. 6 § 3 также выполняется, если реактивное ускорение a достаточно мало. Начальная орбита считается эллиптической, и максимизация подтежит высота первого перигея, т. е. первого минимума функции $r(t)$. Начальные условия задавались в виде:

$$r(0) = 1, \quad \varphi(0) = 0, \quad v(0) = v_0, \quad \theta(0) = \theta_0. \quad (7.5.2)$$