В. М. АЛЕКСЕЕВ
В. М. ТИХОМИРОВ
С. В. ФОМИН

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

*Допущено Министерством
высшего и среднего специального образования СССР
в качестве учебного пособия
для студентов математических специальностей
высших учебных заведений*

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1979

22.18
А 48
УДК 519,6

Оптимальное управление. Алексеев В. М., Тихомиров В. М., Фомин С. В. — М.: Наука. Главная редакция физико-математической литературы, 1979.

Книга написана на основе преподавания курса «Оптимальное управление» на механико-математическом факультете МГУ. Она состоит из трех концентров: 1) элементарный вывод основных условий экстремума и решение конкретных задач; 2) применение теорем дифференциального исчисления в банаховых пространствах к доказательству необходимых условий экстремума; 3) дополнительные вопросы теории экстремальных задач. Особенностью книги является единый подход к различным задачам иа экстремум.

Книга согласована с учебником А. Н. Колмогорова и С. В. Фомина «Элементы теории функций и функционального анализа».

А $\frac{20204-170}{053(02)\text{-}79}$ 2-79. 1702070000

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Одной из характерных особенностей современной эпохи является все возрастающее внимание к проблемам управления. Как никогда прежде, ощущается потребность в плодотворном и эффективном использовании природных богатств, огромных людских ресурсов, материальных и технических средств. Говоря о наиболее приметных явлениях научно-технического прогресса в XX веке, обычно называют расщепление атома, освоение космоса, создание электронной вычислительной техники. На этом фоне теория управления выглядит пока менее эффектно, хотя в развитии современной цивилизации она уже играет выдающуюся роль, и есть основание думать, что в будущем эта роль станет еще значительней.

Всюду, где имеется возможность активного участия человека, возникает проблема отыскания наилучшего, или, как говорят, *оптимального* из возможных управлений. Вызванные к жизни потребностями экономики и техники, оптимизационные проблемы потребовали в свою очередь создания новых разделов математики.

В 40-х годах исследование задач экономики породило новое направление анализа, получившее название линейного и выпуклого программирования. В те же годы приобрели актуальность задачи управления летательными аппаратами и технологическими процессами сложной структуры. Соответствующая математическая теория была создана в середине пятидесятых годов и получила название теории оптимального управления. Выдающуюся роль сыграл в этом «принцип максимума» Л. С. Понтрягина. В теории оптимального управления произошел синтез идей и методов исследования, с одной стороны восходящих к классикам вариационного исчисления, а с другой — вполне современных. Ее развитие самым существенным образом связано с именами советских математиков.

Эта книга задумана как учебное пособие по различным курсам оптимизации, читаемым в университетах и вузах с повышенной математической подготовкой. Расскажем вкратце об общем замысле и плане книги.

История исследования задач на экстремум, или, как сейчас обычно говорят, «экстремальных задач» началась не в наши дни — в той или иной мере такие задачи всегда привлекали внимание математиков. Мы хотели раскрыть перед читателями с самой разной предварительной подготовкой эту связь времен и неразрывность научной дороги. Это сделано в первой главе, рассчитанной на широкую аудиторию. Хотя используемый здесь математический аппарат минимален, мы старались выдержать стиль точного математического текста, ограничиваясь наиболее выразительными, но элементарными фрагментами истории изучения экстремальных задач. Наша цель — связать воедино первоначальные замыслы И. Кеплера и П. Ферма, задачи, поставленные Х. Гюйгенсом, И. Ньютоном и И. Бернулли, идеи и методы Ж. Лагранжа, Л. Эйлера, и К. Вейерштрасса с современным этапом теории, непосредственно продолжающей исследования великих предшественников. Кроме того, в первой главе описаны методы решения конкретных задач и приведены примеры решения на базе единой идеологии задач, поставленных в разные времена учеными разных направлений.

Остальная часть книги адресована по преимуществу математикам. Возникновение новой теории стимулировало развитие и старых и новых разделов математического анализа. Не все эти разделы должным образом представлены в современном математическом образовании.

Нам представляется, что фрагмент классического анализа, объединенный вокруг темы «неявная функция», играет ныне исключительную роль во всех аспектах конечномерного и бесконечномерного анализа. То же самое можно сказать и об основах выпуклого анализа. Наконец, для теории оптимального управления существенно, что основные факты теории дифференциальных уравнений остаются верными и для уравнений с разрывными правыми частями.

Названные выше три раздела анализа и геометрии излагаются во второй главе.

Изложение теории собственно экстремальных задач представлено в третьей и четвертой главах. Их главную часть составляет содержание обязательного полугодового курса, читавшегося авторами на механико-математическом факультете МГУ. Текст препарирован так, что доказательства основных теорем занимают не больше одной лекции. Всюду выдержан принцип полноты изложения. Мы старались везде избежать пропусков, ссылок на очевидность и т. д.

Некоторые стандартные обозначения (из теории множеств, функционального анализа и т. п.) употребляются в тексте без пояснений. Поэтому для облегчения ориентировки читателей в конце книги приведен список основных обозначений с краткой их расшифровкой.

Параграфы, номера которых отмечены в тексте и в оглавлении звездочкой, призваны подвести читателя к современным методам теории экстремальных задач. Здесь изложение более соответствует монографическому стилю, допускаются отдельные, правда минимальные, ссылки на теоремы, хотя и ставшие классическими, но находящиеся пока за пределами традиционных программ. Эти параграфы написаны на основе специальных курсов «Выпуклый анализ», «Дополнительные главы теории экстремальных задач» и других, также читавшихся на механико-математическом факультете МГУ в течение ряда лет.

Более подробно о содержании читатель может узнать из оглавления, где выделены и названы все важнейшие результаты и факты, нашедшие отражение в книге.

Из сказанного следует, что мы рассчитываем на разную читательскую аудиторию: прежде всего — на студентов университетов и вузов с повышенным математическим курсом, но также и на инженеров, экономистов и математиков, сталкивающихся с необходимостью решать различные экстремальные задачи. Для этих читателей написано введение, а для тех, кто интересуется теорией экстремальных задач более серьезно, в конце помещен путеводитель по литературе — монографической и обзорной.

В постановке и лекционной разработке на механико-математическом факультете МГУ курса «Оптимальное управление» очень большая заслуга принадлежит Сергею Васильевичу Фомину; по его же инициативе создана и настоящая книга. В ней использован первоначальный конспект лекций, написанный С. В. Фоминым совместно с В. М. Тихомировым. Безвременная кончина Сергея Васильевича в расцвете сил и таланта не дала ему возможности увидеть завершение своих замыслов.

Мы выражаем свою искреннюю признательность коллективу кафедры общих проблем управления механико-математического факультета МГУ, принимавшему активное участие в обсуждении как методики преподавания курса «Оптимальное управление», так и характера изложения отдельных затрагиваемых в книге вопросов.

Мы считаем своим обязательным долгом отметить, что на формирование математических концепций, положенных в основу этой книги, значительное влияние оказала творческая деятельность А. А. Милютина.

Мы благодарны А. И Маркушевичу за ценные консультации, а также А. П. Буслаеву и Г. Г. Магарил-Ильяеву, внимательно прочитавшим рукопись и сделавшим ряд весьма полезных замечаний.

В. М. Алексеев,
В. М. Тихомиров

ГЛАВА I

ВВЕДЕНИЕ

§ 1.1. Как возникают экстремальные задачи?

Людям свойственно стремление к лучшему, и если им приходится выбирать из нескольких возможностей, то желание найти среди них *оптимальную* представляется вполне естественным.

Слово «оптимальный» происходит от латинского optimus, что значит — наилучший, совершенный. Для того чтобы найти оптимальную из возможностей, приходится решать задачи на отыскание *максимума* или *минимума*, т. е. наибольших и наименьших значений каких-то величин. Оба эти понятия — максимум и минимум — объединяются единым термином *«экстремум»* (от латинского extremum, означающего «крайнее»). Поэтому задачи на отыскание максимума или минимума называют *экстремальными задачами*.

Методы решения и исследования различного рода экстремальных задач составляют специальные разделы математического анализа. Почти тот же смысл вкладывается в название *«задачи оптимизации»*; в последнем более отчетливо прослеживается связь с практическими применениями математики.

Цель этой книги познакомить читателя с теорией и приемами решения экстремальных задач. Но прежде, чем переходить к формальному, логически последовательному изложению этой ветви математики, мы хотим обратиться к прошлому, чтобы лучше понять причины, побуждающие ставить и решать экстремальные задачи в их прикладном аспекте, т. е. как задачи оптимизации.

Mercatique solum, facti
de nomine Byrsam
Taurino quantum possent
circumdare tergo

P. Vergilius Maro «Aeneas» [1])

1.1.1. Классическая изопериметрическая задача. Задача Дидоны. Задачи отыскания наибольших и наименьших величин были поставлены впервые античной наукой. Древнейшей из известных экстремальных задач является, пожалуй, классическая изопериметрическая задача. Трудно сказать, когда впервые была высказана мысль о наибольшей «вместимости» окружности и сферы среди всех замкнутых кривых одной и той же длины, или поверхностей одной и той же площади. Один из последних учеников афинской школы платоников Симплиций (VI в. н. э.), составивший незадолго до окончательного краха античной цивилизации обширный комментарий к трудам Аристотеля (IV в. до н. э.), пишет: «Доказано до Аристотеля, ибо он пользуется этим, как известным, а затем более полно — Архимедом и Зенодором, что среди изопериметрических фигур наиболее вместимым является круг, а среди изопифанных — шар». В этих словах обозначена постановка следующих экстремальных задач: *среди плоских замкнутых кривых, имеющих заданную длину, найти кривую, охватывающую наибольшую площадь, и среди пространственных замкнутых поверхностей, имеющих заданную площадь, найти поверхность, охватывающую наибольший объем.* Для философа-платоника такая постановка задачи естественна и связана с поисками идеальных форм. Недаром круг и шар были в древности символами геометрического совершенства.

Более прозаическую мотивировку той же изопериметрической и ряда близких к ней задач мы находим, пусть даже в несколько наивной, но достаточно отчетливой форме в легенде о Дидоне. Напомним ее, следуя «Энеиде» римского поэта Вергилия, две строки которого приведены выше в качестве эпиграфа.

Финикийская царевна Дидона и с ней небольшой отряд жителей города Тира, спасаясь от преследований тирана — брата Дидоны, покинули родной город и в поисках счастья отправились на кораблях на запад вдоль

[1]) Столько купили земли и дали ей имя Бирса, сколько смогли окружить бычьей шкурой. (*П. Вергилий Марон* «Энеида»).

берегов Средиземного моря. Выбрав на африканском побережье удобное место (нынешний Тунисский залив), Дидона и ее спутники решили основать здесь поселение. По-видимому, эта идея не вызвала энтузиазма у местных жителей, но все же Дидоне удалось уговорить их предводителя Ярба, и он неосторожно согласился уступить Дидоне клочок земли, «который можно окружить бычьей шкурой». Не сразу понял простодушный Ярб хитрость и коварство финикиянки. Разрезав шкуру на тонкие полоски, Дидона связала их в один длинный ремень и, окружив им значительную территорию, заложила на ней город Карфаген[1]). В память об этой истории карфагенская цитадель получила название Бирса[2]). Все эти события легенда относит к 825 (или к 814) г. до н. э.

Анализируя ситуацию, мы обнаруживаем несколько возможностей поставить здесь задачу оптимизации.

А) Требуется указать оптимальную форму участка земли, который при заданной длине периметра L имеет наибольшую площадь S.

Ясно, что это та же самая классическая изопериметрическая задача[3]). Ее решением является круг.

Упражнение. Считая бычью шкуру прямоугольником 1×2 м и приняв ширину ремешка 2 мм, найдите L и максимальное S.

(Авторам не удалось найти точные размеры Бирсы. Расположенная на высоком (63 м) холме, она вряд ли была особенно большой. Для сравнения — длина стен Московского Кремля 2235 м.)

Решение изопериметрической задачи заключено в следующем утверждении:

Если спрямляемая кривая длины L ограничивает плоскую фигуру площади S, то

$$L^2 \geqslant 4\pi S, \tag{1}$$

причем равенство имеет место тогда и только тогда, когда кривая — окружность.

Неравенство (1) называется *изопериметрическим*; его доказательство можно найти в [21].

Б) Другие постановки задачи получаются, если, как

---

[1]) Финикийское Картадашт — новый город.

[2]) На языке пунийцев (так римляне называли жителей Карфагена) — шкура. Это название употребляется до сих пор.

[3]) Еще одна реальная ситуация, приводящая к той же задаче, описана Л. Н. Толстым в рассказе «Много ли человеку земли нужно». Разбор этого рассказа с точки зрения геометрии см. Я. И. Перельман «Занимательная геометрия». — М.—Л.: Гостехиздат, 1950, гл. 12.

это естественно считать, Дйдона хотела сохранить выход к морю. В отличие от классической изопериметрической, мы будем называть эти задачи задачами Дидоны. Для простоты рассмотрим сначала случай прямолинейного берега (рис. 1).

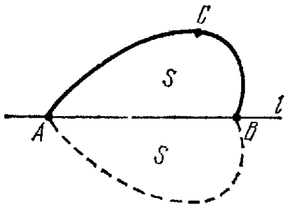


Рис. 1.

Первая задача Дидоны. *Среди всех дуг длины L, содержащихся в полуплоскости, ограниченной прямой l и с концами A, $B \in l$, найти такую, которая вместе с отрезком $[AB]$ ограничивает фигуру наибольшей площади S.*

Решение. Пусть ACB — произвольная допустимая дуга с концами A, $B \in l$, ограничивающая фигуру площади S (рис. 1). Отразив ее симметрично относительно l, мы получим замкнутую кривую длины $2L$, ограничивающую фигуру площади $2S$. Согласно изопериметрическому неравенству

$$(2L)^2 \geqslant 4\pi 2S, \tag{2}$$

откуда

$$S \leqslant L^2/(2\pi). \tag{3}$$

Следовательно, максимальным значением S может быть только $L^2/(2\pi)$, и это значение действительно достигается, если ACB — полуокружность, опирающаяся на диаметр $[AB]$. Задача имеет единственное решение с точностью до сдвига отрезка вдоль прямой (почему?).

В) В предыдущей задаче положение концов A и B искомой дуги можно было выбирать на прямой l произвольно. Что произойдет, если эти концы будут заданы?

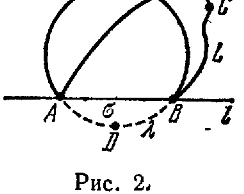


Рис. 2.

Вторая задача Дидоны. *Среди всех дуг длины L, лежащих в полуплоскости, ограниченной прямой l, и с заданными концами A, $B \in l$ найти такую, которая вместе с отрезком $[AB]$ ограничивает фигуру наибольшей площади.*

Решение. Ясно, что задача имеет смысл только при $L > |AB|$ (в противном случае либо нет ни одной дуги, удовлетворяющей условиям задачи, либо (при $L = |AB|$) такая дуга только одна — сам отрезок $[AB]$). Естественно ожидать по аналогии с предыдущим, что решением будет дуга окружности, для которой $[AB]$ является хордой. Такая дуга $A\hat{C}B$ определяется единственным образом. Дополним ее до полной окружности дугой ADB (рис. 2). Длину дуги ADB обозначим λ, а площадь сегмента, ограниченного этой дугой и отрезком $[AB]$,— σ.

Пусть теперь ACB—произвольная дуга, удовлетворяющая условиям задачи и ограничивающая вместе с $[AB]$ площадь S. Замкнутая кривая $ACBD$ имеет длину $L+\lambda$ и ограничивает площадь $S+\sigma$. Согласно (1) $4\pi(S+\sigma)\leqslant \leqslant(L+\lambda)^2$, откуда

$$S\leqslant\frac{1}{4\pi}(L+\lambda)^2-\sigma.$$

Как и в (1), равенство, а значит, и максимальная S достигаются тогда и только тогда, когда кривая $ACBD$ является окружностью, т. е. когда дуги равны: $ACB=A\hat{C}B$.

Обратим внимание на следующее различие в двух разобранных задачах. В первой задаче Дидоны запас конкурирующих кривых больше, поскольку положение точек A и B не задано. Впрочем, без ограничения общности одну из них, скажем A, можно считать фиксированной. Положение точки B тогда определяется дополнительным условием: $A\hat{C}B$ не просто дуга окружности, как во второй задаче Дидоны, но это полуокружность. В эквивалентной форме: в своих концах искомая дуга подходит к прямой l под углом 90°. Мы увидим потом, что здесь проявляется общий принцип: предоставляя концам искомой кривой некоторую свободу, мы должны потребовать, чтобы в них выполнялись некоторые условия, называемые *условиями трансверсальности*. Форма же искомой кривой в обеих задачах одинакова, она определяется некоторым уравнением (*уравнением Эйлера*), которое должно выполняться вдоль кривой. В нашем случае во всех точках искомая кривая должна иметь одну и ту же кривизну.

Г) Рассмотрим теперь, что произойдет, если берег не является прямой линией, ограничась случаем фиксированных концов, Легко сообразить, что если между точками A и B берег мало отличается от прямой, то решением остается та же дуга окружности $A\hat{C}B$, что и раньше.

Предыдущее доказательство применимо полностью, только буквой σ следует обозначать площадь, заштрихованную на рис. 3. Отсюда же видно, что будет в случае, когда между A и B находится глубокий залив. Пусть, например, между A и B берег прямой, но из точки D перпендикулярно AB прорыт канал DC. Считая, что граница города должна проходить по суше, мы видим, что решением задачи будет та же дуга $A\hat{C}B$, пока точка D остается внутри нее (рис. 4, *а*). Если же канал DC пересек дугу $A\hat{C}B$, но еще $|AC|+|CB|<L$, решением является кривая, составленная из двух дуг окружностей AC и CB

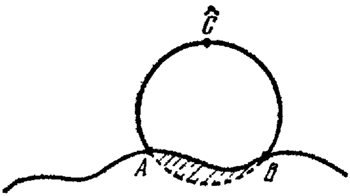


Рис. 3.

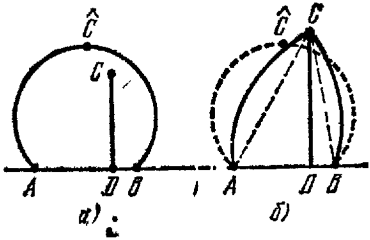


Рис. 4.

(рис. 4, *б*). В предельном случае $|AC|+|CB|=L$ решением является ломаная ACB, а при $|AC|+|CB|>L$ решение не существует.

Рассмотренный вариант можно было бы назвать *задачей Дидоны с фазовыми ограничениями*.

Д) Наконец, остановимся еще на одном варианте задачи Дидоны. Пусть по каким-то соображениям (например, ввиду запрета жрецов бога Эшмуна, храм которого впоследствии находился в Бирсе) стены города нельзя вести более чем под углом 45° к линии берега, который мы снова предполагаем прямым. Такая задача была бы уже *задачей оптимального управления*. Ее решение можно найти при помощи *принципа максимума Понтрягина*, с которым мы познакомимся ниже. В типичном случае решение выглядит, как на рис. 5. Отрезки AC и BE образуют с линией берега угол 45°, а CDE — дуга окружности.

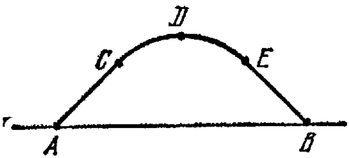


Рис. 5.

1.1.2. Другие старинные экстремальные задачи в геометрии. Уделив изопериметрической задаче достаточно много места, мы остановимся на некоторых других экстремальных задачах с геометрическим содержанием, которые рассматривались математиками разных веков. Такие задачи встречаются, в частности, в трудах величайших математиков античности—Евклида, Архимеда и Аполлония.

В «Началах» Евклида (IV в. до н. э.) мы находим лишь одну задачу на максимум (книга 6, предложение 27). В современной редакции она выглядит так.

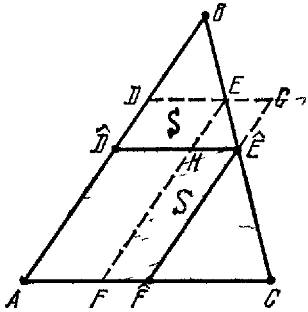


Рис. 6.

Задача Евклида. *В данный треугольник ABC вписать параллелограмм $ADEF$* (рис. 6) *наибольшей площади.*

Решение. У искомого параллелограмма точки $\hat{D}$, $\hat{E}$ и $\hat{F}$ являются серединами соответствующих сторон данного треугольника. Доказать это можно разными способами. Например, легко показать, что площади параллелограммов $\hat{D}DG\hat{E}$ и $FH\hat{E}\hat{F}$ одинаковы. Отсюда следует, что площадь параллелограмма $ADEF$ меньше площади параллелограмма $A\hat{D}\hat{E}\hat{F}$, ибо последняя равна площади фигуры $ADG\hat{E}HF$, содержащей параллелограмм $ADEF$. ■

В дошедших до нас сочинениях Архимеда (III в. до н. э.) изопериметрическая задача не упоминается. Доныне неизвестно, что было сделано Архимедом в этой области, и потому приведенные в п. 1.1.1 слова Симплиция остаются пока загадочными. Решение же одной *изопифанной* (т. е. относящейся к фигурам равной площади) задачи имеется в сочинении Архимеда «О шаре и цилиндре». Там ставится и решается *задача о максимальном объеме, который могут иметь сегменты шаров одинаковой по площади боковой поверхности.*

Ответом в этой задаче является полушар (подобно тому, как полукруг является решением второй задачи Дидоны).

«Коника» или «Конические сечения»—так называется величайшее творение Аполлония (III—II в. до н. э.). К нашей теме относится пятая книга «Конических сечений». Вот что пишет Ван дер Варден: Аполлоний ставит «задачу о том, как провести из одной точки O к коническому сечению самый длинный и самый короткий прямолинейные отрезки. Однако он дает больше, чем обещает: он определяет все проходящие через O прямые, которые пересекают коническое сечение под прямым углом (в настоящее время их называют нормалями), разбирает, при каком положении O задача имеет два, три или четыре решения». Перемещая точку O, он «определяет ординаты граничных точек G_1 и G_2, где число проходящих через O нормалей разом переходит с 2 на 4 и обратно»[1]) (рис. 7).

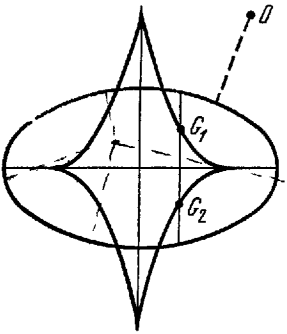


Рис. 7.

Вместе с гибелью античной цивилизации научная деятельность в Европе замирает примерно до XV века. В XVI в. закладываются основы алгебры и появляются первые экстремальные задачи алгебраического содержания.

Вот, например, задача, предлагавшаяся Н. Тартальей (XVI в.): *разделить число восемь на две части так, чтобы произведение произведения этих частей на их разность было максимальным.*

До XVII в. не было выработано никаких общих приемов решения экстремальных задач, и каждая из них решалась специально для нее разработанным приемом. В 1615 г. вышла книга И. Кеплера «Новая стереометрия винных бочек»[2]). Кеплер начинает книгу так. «В тот год, как я женился, урожай винограда был хороший и вино дешево, а потому мне, как хорошему хозяину, следовало запастись вином. Я купил его несколько бочонков. Через некоторое время пришел продавец — измерить вместимость бочонков, чтоб назначить цену за вино. Для этого он опускал в каждый бочонок железный прут, и, не прибегая ни к какому вычислению, немедленно объявлял, сколько в бочке вина». (Этот «метод» мы иллюстрируем на рис. 8.)

Кеплер очень удивился. Ему показалось странным, как с помощью одного измерения можно вычислить вме-

---

[1]) Ван дер Варден Б. Л. Пробуждающаяся наука.— М.: Физматгиз, 1959.

[2]) Кеплер И. Новая стереометрия винных бочек.— М.— Л.: ОНТИ — ГТТИ, 1935.

стимость бочек разной формы. «Я счел для себя подходящим,— пишет Кеплер,— взять новый предмет математических занятий и исследовать геометрические законы такого удобного измерения, и выяснить его основания». Для разрешения поставленной задачи Кеплер закладывает основания интегрального и дифференциального исчисления и заодно дает первые общие правила решения экстремальных задач. Он пишет (цитируем по книге Е. А. Предтеченского)[1]) «Кеплер, его жизнь и научная деятельность»: «Под влиянием благодатного гения, бывшего, без сомнения, хорошим геометром, бочары стали придавать бочкам ту форму, которая при данной длине линии, измеренной мерщиком, дает возможность судить о наибольшей вместимости бочки, а так как **вблизи всякого максимума изменения бывают нечувствительными**, то небольшие случайные отклонения не оказывают заметного влияния на емкость».

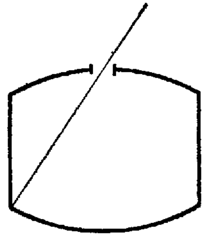


Рис. 8.

В выделенных нами словах и заложен тот основной алгоритм нахождения экстремумов, который впоследствии был оформлен в точную теорему сначала (для многочленов) Ферма (1629 г.), а затем — Ньютоном и Лейбницем и получил название *теоремы Ферма*.

Отметим еще, что Кеплер решил несколько конкретных экстремальных задач, в частности задачу *о цилиндре наибольшего объема, вписанном в шар*.

В заключение этого пункта приведем еще одну геометрическую задачу, которой в XVII в. интересовались многие математики (Кавальери, Вивиани, Торичелли, Ферма и др.). В XIX в. ею занимался немецкий геометр Штейнер, и потому эту задачу часто называют *задачей Штейнера*[2]).

Задача Штейнера. *В плоскости треугольника найти точку, сумма расстояний от которой до вершин треугольника минимальна.*

Решение. Приведем изящное геометрическое решение этой задачи для остроугольного треугольника. Пусть

---

[1]) Предтеченский Е. А. Кеплер, его жизнь и научная деятельность.— Петербург: изд. З. И. Гржебина, 1921.

[2]) Отметим заодно, что впоследствии подобные задачи стали возникать при строительстве дорог, нефтепроводов и городских коммуникаций.

в треугольнике ABC (рис. 9) величина угла $C \geqslant 60°$. Совершим поворот треугольника ABC вокруг точки C на угол $60°$. Получим треугольник $A'B'C$. Возьмем теперь любую точку D в треугольнике ABC, а через D' обозначим образ D при нашем повороте. Тогда сумма длин $|AD|+|BD|+|CD|$ равна длине ломаной $|BD|+|DD'|+$ $+|D'A'|$, ибо треугольник CDD' равносторонний и $|D'A'|=$ $=|DA|$.

Пусть теперь $\hat{D}$ — *точка Торичелли*, т. е. точка, из которой все стороны треугольника видны под углом $120°$, и $\hat{D}'$ — образ $\hat{D}$ при повороте. Нетрудно понять, что тогда точки B, $\hat{D}$, $\hat{D}'$ и A' лежат на одной прямой и, значит, точка Торичелли и является решением задачи. ■

Мы познакомили читателя с несколькими экстремальными задачами геометрического содержания, поставленными и решенными в разные времена. Постановку их лишь частично можно оправдать практическими потребностями — основным стимулом здесь, пожалуй, было желание показать красоту самой геометрии.

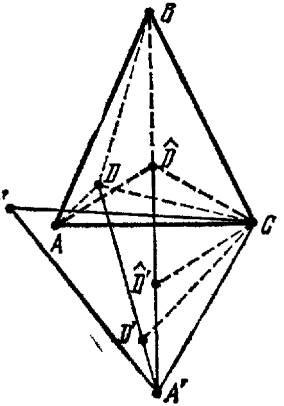


Рис. 9.

1.1.3. Вариационный принцип Ферма и принцип Гюйгенса. Задача о преломлении света. С именем Пьера Ферма связана формулировка первого вариационного принципа для физической проблемы. Речь идет о *вариационном принципе Ферма в геометрической оптике*. Экспериментально закон преломления света был установлен Снеллиусом. Вскоре Декарт дал теоретическое объяснение этого закона. Однако при этом у Декарта получалось, что в более плотной среде (скажем, в воде) скорость распространения света больше, чем в менее плотной (например, в воздухе). Этот факт многим показался сомнительным.

Ферма дал другое объяснение явления. Его основная идея состояла в том, что луч света «избирает» такую траекторию, вдоль которой время, затрачиваемое на преодоление пути от одной точки до другой было бы минимальным (в сравнении с любыми другими траекториями, соединяющими те же точки). В однородной среде, в которой скорость распространения света во всех точках и во всех направлениях одинакова, время, затрачиваемое светом на прохождение некоторой траектории, пропорционально ее длине. Поэтому траектория минимального времени, соединяющая точки A и B — это просто отрезок $[AB]$: в однородной среде свет распространяется прямолинейно.

Вывод закона преломления Снеллиуса из вариационного принципа Ферма сейчас содержится в школьных учебниках (см. также п. 1.6.1).

В основу принципа Ферма положено допущение о том, что свет распространяется по некоторым линиям. X. Гюйгенсу (1629—1695) принадлежит другое объяснение законов распространения и преломления света, основанное на представлении о свете как о волне, фронт которой движется со временем.

Отвлекаясь от обсуждения физических предпосылок этой идеи, и, в частности, от вопроса о том, что такое свет — волна или поток частиц, дадим следующее, скорее, наглядное, чем строгое, определение.

Волновым фронтом S_t называется множество точек, которых достигает за время t свет, распространяемый некоторым источником S_0.

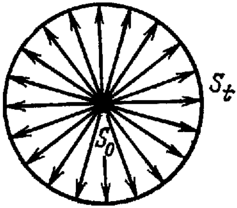


Рис. 10.

Если, например, источник S_0 — это точка, а среда однородна, то S_t — это сфера («сферическая волна») радиуса vt с центром в S_0 (рис. 10). С ростом t волновой фронт расширяется равномерно во все стороны со скоростью v. Линии же распространения света образуют пучок лучей (радиусов), ортогональных S_t в каждый момент t. По мере удаления от источника сферическая волна становится все более и более плоской и, если мы вообразим источник бесконечно удаленным, то в пределе волновой фронт оказывается плоскостью, равномерно движущейся со скоростью v и остающейся перпендикулярной пучку лучей света, которые теперь параллельны между собой (рис. 11).

Для определения движения волнового фронта в более сложных ситуациях Гюйгенс пользуется следующим правилом («*принцип Гюйгенса*»). Каждая точка волнового фронта S_t сама становится вторичным источником, через время Δt мы получаем семейство волновых фронтов от всех этих вторичных источников и истинный волновой фронт $S_{t+\Delta t}$ в момент $t+\Delta t$ есть огибающая этого семейства (рис. 12).

Легко видеть, что распространение света от точечного источника в однородной среде и предельный случай плоской волны удовлетворяют принципу Гюйгенса. В качестве примера применения этого принципа в простейшей неоднородной среде приведем вывод закона Снеллиуса «по Гюйгенсу».

Пусть параллельный пучок световых лучей падает на

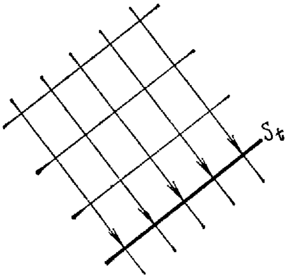


Рис. 11.

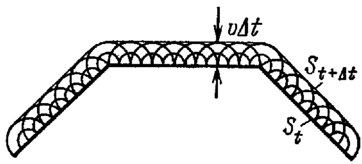


Рис. 12.

плоскую границу Σ раздела двух однородных сред; для простоты будем представлять себе, что Σ горизонтальна, а свет падает сверху (рис. 13). Через v_1 и v_2 обозначим скорости распространения света над и под Σ и через α_1

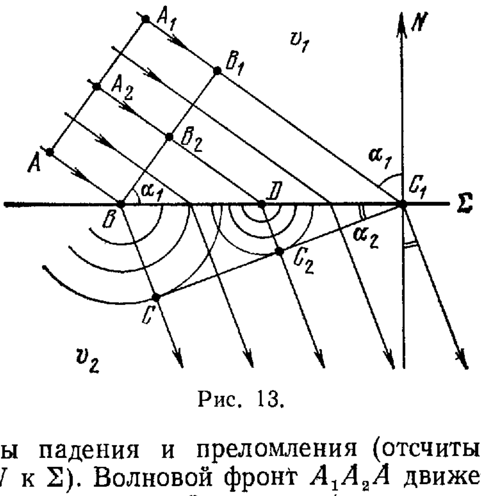


Рис. 13.

и α_2—углы падения и преломления (отсчитываемые от нормали N к Σ). Волновой фронт A_1A_2A движется со скоростью v_1 и в некоторый момент t свет, вышедший из точки A достигает границы Σ в точке B. После этого точка B становится вторичным источником сферических волн, распространяющихся в нижней среде со скоростью v_2. В точку C_1 свет придет в момент $t_1 = t + \frac{|B_1C_1|}{v_1} = t + |BC_1|\frac{\sin\alpha_1}{v_1}$, а в промежуточную точку $D \in [BC_1]$ — в момент $t_2 = t + |BD|\frac{\sin\alpha_1}{v_1}$. К моменту t_1 сферическая волна от вторичного источника B будет иметь радиус $r_1 = v_2(t_1 - t) = |BC_1|\frac{v_2}{v_1}\sin\alpha_1$, а волна от D — радиус $r_2 = v_2(t_1 - t_2) = |DC_1|\frac{v_2}{v_1}\sin\alpha_1$. Касательные C_1C и C_1C_2 к этим сферам совпадают, поскольку

$$\sin\widehat{BC_1C} = \frac{r_1}{|BC_1|} = \frac{v_2}{v_1}\sin\alpha_1 = \frac{r_2}{|DC_1|} = \sin\widehat{DC_1C_2}.$$

Точка D была взята на BC_1 произвольно, и, следовательно, вторичные волны в момент t_1 имеют огибающей

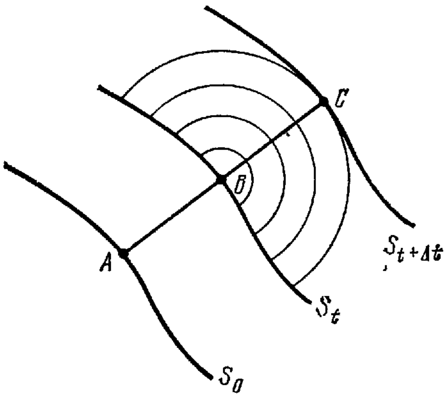


Рис. 14.

прямую CC_1, образующую с Σ угол α_2 такой, что $\sin\alpha_2 = \frac{v_2}{v_1}\sin\alpha_1$. А это и есть закон Снеллиуса

$$\frac{\sin\alpha_2}{\sin\alpha_1} = \frac{v_2}{v_1}.$$

Принципы Гюйгенса и Ферма тесно связаны между собой. В самом деле, пусть нам известно положение волнового фронта S_t в некоторый момент t. Где он будет находиться через некоторое время Δt? Возьмем точку $C \in S_{t+\Delta t}$ (рис. 14). По определению существует $A \in S_0$ и путь AC, который свет проходит за время $t+\Delta t$, и, в соответствии с принципом Ферма, любой другой путь из A в C требует большего времени. По непрерывности на дуге AC найдется точка B такая, что от A до B свет проходит за время t, а от B до C — за время Δt. Так как дуга AC обладает свойством минимальности, то этим же свойством должны обладать и дуги AB и BC. Действительно, если бы существовал, например, путь, по которому из A в B свет мог бы добраться меньше чем за время t, то, дополнив этот путь дугой BC, мы бы получили путь из A в C, по которому свет идет меньше чем за время $t+\Delta t$, а это невозможно. Отсюда следует, во-первых, что $B \in S_t$, а во-вторых, что точка C принадлежит волновому фронту, отвечающему точечному источнику света, находящемуся в B, и времени Δt. Это вполне согласуется с принципом Гюйгенса: точка B стала вторичным источником и распространяющаяся от нее волна оказалась через время Δt в точке C. Идея волнового фронта, принцип Гюйгенса и рассуждение, набросок которого мы только что привели, явились базой для будущей *теории Гамильтона — Якоби*, а в середине нашего века — для так называемой *теории динамического программирования*, которая является важным инструментом решения прикладных экстремальных задач.

* * *

Вслед за вариационным принципом Ферма было открыто множество других вариационных принципов — сначала в механике, а затем в физике. Со временем у большинства ученых созрела уверенность в том, что природа всегда «избирает» движение, как бы решая некоторую задачу на экстремум. Здесь уместно привести слова Эйлера: «В мире не происходит ничего, в чем не был бы виден смысл какого-нибудь максимума или минимума». В наши дни Карл Зигель позволил себе такое шутливое высказывание: «По Лейбницу наш мир является наилучшим из всех возможных миров, а поэтому законы можно описать экстремальными принципами».

1.1.4. Задача о брахистохроне. Зарождение вариационного исчисления. В 1696 г. появилась заметка И. Бернулли с интригующим заглавием: «Problema novum, ad cujus solutionem mathematici invitantur» («Новая задача, к решению которой приглашаются математики»). В ней была поставлена следующая задача: «В вертикальной плоскости даны две точки A и B (рис. 15). Определить путь AMB, спускаясь по которому под действием собственной тяжести тело M, начав двигаться из точки A, дойдет до точки B в кратчайшее время»[1]).

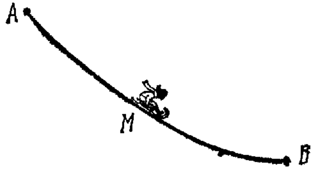


Рис. 15.

Решение этой задачи, по словам Лейбница, «столь прекрасной и поныне неизвестной», было дано самим И. Бернулли, а также Лейбницем, Я. Бернулли и еще одним анонимным автором, в котором знатоки «ex unge leonem» (по словам И. Бернулли)[2]) сразу же узнали Ньютона. Кривая наикратчайшего спуска или брахистохрона оказалась циклоидой. Решение Лейбница было основано на аппроксимации кривых ломаными. Развитая впоследствии в работах Эйлера, эта идея заложила основы прямых методов в вариационном исчислении. Замечательное решение Я. Бернулли использовало принцип Гюйгенса и идею «волнового фронта». Однако наибольшую популярность получило решение самого автора. Его мы и приведем.

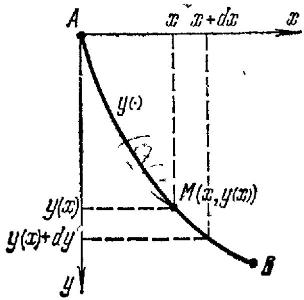


Рис. 16.

Введем в плоскости систему координат (x, y) так, чтобы ось x была горизонтальна, а ось y направлена вниз. В соответствии с законом Галилея скорость тела M в точке с координатами $(x, y(x))$ (если тело без трения спускается по кривой $y(\cdot)$ — см. рис. 16) не зависит от формы кривой $y(\cdot)$ между точками A и $(x, y(x))$, а зависит лишь от ординаты $y(x)$ и равна $\sqrt{2gy(x)}$, где g — ускорение силы тяжести. При этом требуется найти наименьшее время, которое потребуется на преодоление пути от A к B, т. е.

---

[1]) Может быть нелишне напомнить, что некоторый отдаленный намек на постановку задачи о брахистохроне содержится в «Беседах» Галилея. Там доказывается, что, двигаясь по хорде, тело придет в конечную точку позже, чем двигаясь по окружности, стягиваемой хордой.

[2]) По когтям (узнают) льва (*лат.*).

надо минимизировать интеграл

$$T=\int\limits_{\overset{\smile}{AB}} \frac{ds}{v}=\int\limits_{\overset{\smile}{AB}} \frac{ds}{\sqrt{2gy(x)}},$$

где ds — дифференциал длины дуги.

Но (в силу принципа Ферма из п. 1.1.3) мы получим в точности ту же задачу, если будем исследовать траектории света в неоднородной (двумерной) среде, где скорость в точке (x, y) равна $\sqrt{2gy}$. Далее, И. Бернулли «дробит» среду на параллельные слои, где считает скорость постоянной и равной v_i, $i=1, 2, \ldots$ (рис. 17). В силу закона Снеллиуса получаем

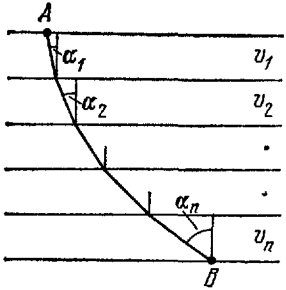


Рис. 17.

$$\frac{\sin\alpha_1}{v_1}=\frac{\sin\alpha_2}{v_2}=\ldots\Leftrightarrow\frac{\sin\alpha_i}{v_i}=\text{const},$$

где α_i — углы падения луча. Переходя к пределу при измельчении слоев (Бернулли, разумеется, не останавливался на обосновании законности такой процедуры), получаем, что

$$\frac{\sin\alpha(x)}{v(x)}=\text{const},$$

где $v(x)=\sqrt{2gy(x)}$, а $\alpha(x)$ — угол между касательной к кривой $y(\cdot)$ в точке $(x, y(x))$ и осью Oy, т. е. $\sin\alpha(x)= 1/\sqrt{1+(y'(x))^2}$. Итак, уравнение брахистохроны:

$$\sqrt{1+(y')^2}\sqrt{y}=C\Leftrightarrow y'=\sqrt{\frac{C-y}{y}}\Leftrightarrow\frac{dy\sqrt{y}}{\sqrt{C-y}}=dx.$$

Интегрируя его $\left(\text{подстановкой}\quad y=C\sin^2\frac{t}{2},\quad dx= C\sin^2\frac{t}{2}\,dt\right)$, приходим к уравнениям циклоиды:

$$x=C_1+\frac{C}{2}(t-\sin t),\quad y=\frac{C}{2}(1-\cos t).$$

Подчеркнем различие между задачей Евклида и Кеплера (о вписанном цилиндре) и, скажем, задачей о брахистохроне. Множество всех вписанных в треугольник параллелограммов и множество всех вписанных в шар цилиндров зависят от *одного* параметра. Таким образом, в этих задачах требуется найти экстремум функции *одного переменного*. В задаче же о брахистохроне множество всех кривых, соединяющих две точки, *бесконечномерно*. Здесь требуется найти экстремум функции *бесконечного числа переменных*. История математики проделала неожиданный скачок — от единицы сразу к бесконечности, от теории экстремумов функций одного переменного — к теории задач типа задачи о брахистохроне, т. е. к *вариационному исчислению*, как стали называть этот раздел в XVIII в.

Вскоре после работы И. Бернулли было решено много задач, подобных задаче о брахистохроне: о кратчайших линиях на поверхности, о равновесии тяжелой нити и другие.

Годом же рождения вариационного исчисления принято считать 1696 г. — год брахистохроны. Однако исторически это не совсем верно. Об этом мы расскажем в следующем пункте.

1.1.5. Аэродинамическая задача Ньютона. В 1687 г. вышли «Математические начала натуральной философии» Ньютона. В разделе седьмом, озаглавленном «О движении жидкостей и сопротивлении брошенных тел», Ньютон рассматривает задачу о сопротивлении шара и цилиндра в «редкой» среде[1]). Затем в «Поучении» Ньютон исследует вопрос о сопротивлении усеченного конуса, движущегося в той же «редкой» среде. В частности, он обнаруживает, что среди всех конусов, имеющих данную ширину и высоту, наименьшее сопротивление будет испытывать конус с углом 135°. Заметив мимоходом, что данный результат может быть «не бесполезен при построении судов», Ньютон пишет так: «Quod si figura *DNFG* ejusmodi sit ut, si ab ejus puncto quovis *N* ad axem *AB* demittatur perpendiculum *NM*, et dicatur recta *GP* quae parallela sit rectae figuram tangenti in *N*, et axem productam sicet in *P*, fuerit *MN* ad *GP* ut GP^{cub} ad $4BP \times GB^q$, solidum quod figurae hujus revolutione circa axem *AB* describitur resistetur minime omnium ejusdem longitudinis & latitudinis».

Сказанное Ньютоном можно перевести следующим образом: «Когда же кривая *DNFG* будет такова, что если из

[1]) См. предложение 34, теорема 28 в кн: Крылов А. Н. Собрание трудов, т. 7.— М.— Л.: Изд. АН СССР, 1936.

любой ее точки N опустить на ось AB перпендикуляр и [из заданной точки G] провести прямую GP, параллельную касательной к кривой в точке N, пересекающую ось в точке P, то [имеет место пропорция] $MN:GP=GP^3:(4BP\times GB^2)$, тогда тело, получающееся вращением этой кривой около оси AB, будет испытывать наименьшее сопротивление в вышеупомянутой редкой среде среди других тел той же длины и ширины» (см. рис. 18, принадлежащий Ньютону). Ньютон не дал объяснения тому, как он пришел к своему решению. Впоследствии он передал своим комментаторам наброски вывода, но они были опубликованы лишь в 1727—1729 гг., когда уже завершался первый этап вариационного исчисления. Подготовительные материалы Ньютона, опубликованные лишь в наше время, показывают, что он владел элементами многих конструкций, впоследствии осуществленных Эйлером и Лагранжем. Как мы увидим далее, задачу Ньютона следует отнести даже собственно не к вариационному исчислению, а к оптимальному управлению, теория которого начала разрабатываться в пятидесятые годы нашего века.

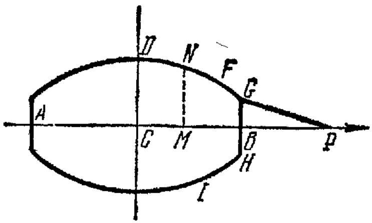


Рис. 18.

1.1.6. Задача о рационе и транспортная задача. Предположим, что запасы некоторого продукта распределены по нескольким базам и что этот продукт должен быть доставлен нескольким магазинам. Стоимость перевозки единицы продукта от каждой базы до каждого магазина известна, и известно, сколько продукта должно быть доставлено в каждый магазин. *Транспортная задача* заключается в составлении оптимального в этой ситуации плана перевозок, т. е. в указании того, какое количество данного продукта надо перевезти из каждой базы в каждый магазин, чтобы суммарная стоимость перевозки была минимальной. Похожая *задача о рационе* состоит в том, чтобы при определенном ассортименте продуктов, заданном содержании в каждом из них питательных веществ и известной стоимости единицы каждого продукта составить рацион, удовлетворяющий необходимым потребностям с минимальными денежными затратами.

Такого рода задачи в огромном количестве возникают в конкретной экономике. Обе упомянутые выше задачи относятся к разделу, называемому *линейным программированием*. Теория линейного программирования была построена лишь в сравнительно недавнее время — в сороковые-пятидесятые годы нашего века.

1.1.7. Задача о быстродействии. Приведем простейший пример экстремальной задачи с «техническим» содержанием. Пусть имеется тележка, движущаяся прямолинейно без трения по горизонтальным рельсам. Тележка управляется внешней силой, которую можно изменять в заданных пределах. Требуется остановить тележку в определенном положении в кратчайшее время. Эту задачу мы называем далее *простейшей задачей о быстродействии.*

Особенность постановок экстремальных задач техники состоит в том, что действующие силы делятся на две части. Одна из них — это силы природы (скажем, сила тяготения), другая (скажем, сила тяги) регулируется человеком. При этом, естественно, возникают ограничения на управляемые воздействия, связанные с техническими возможностями.

Теория решения подобных задач была построена еще позже — в конце пятидесятых годов. Ее называют *теорией оптимального управления.*

* * *

Итак, откуда берутся экстремальные задачи? Приведенными здесь примерами мы постарались показать, что ответов на этот вопрос много. Экстремальные задачи возникают как из естествознания, из экономики и техники, так и вызываются потребностями самой математики. Поэтому теория экстремальных задач и ее практический аспект — теория оптимизации — приобрели в наши дни большую популярность.

§ 1.2. Как формализуются экстремальные задачи?

1.2.1. Основные определения. Каждая из задач § 1.1 была сформулирована в содержательных терминах той частной области, где эта задача возникла. Обычно экстремальные задачи ставятся именно так, и, вообще говоря, не всякую задачу обязательно надо решать аналитически. К примеру, задачи Евклида и Штейнера мы решили чисто геометрическим способом. Однако если мы все-таки желаем воспользоваться преимуществами аналитического подхода, то первое, что необходимо, это осуществить перевод задачи с «содержательного» языка на формальный язык анализа. Такой перевод называется *формализацией*.

Точно поставленная экстремальная задача включает в себя следующие элементы: *функционал*[1]) $f\colon X \to \overline{\mathbf{R}}$, определенный на некотором множестве X, и *ограничение*, т. е. некоторое подмножество $C \subseteq X$. (Через $\overline{\mathbf{R}}$ обозначается «расширенная вещественная прямая», т. е. совокупность всех вещественных чисел, пополненная значениями $-\infty$ и $+\infty$.) Множество X называется иногда *классом допустимых элементов*, а точки $x \in C$ — *допустимыми по ограничению*. При этом сама задача формулируется так: *найти экстремум* (т. е. нижнюю или верхнюю грань) *функционала f при условии, что $x \in C$*. Для той же задачи будет употребляться стандартная запись:

$$f(x) \to \inf (\sup); \quad x \in C. \tag{1}$$

Таким образом, для точной постановки надо описать X, f и C.

Если $X = C$, то задача (1) называется *задачей без ограничений*. Точку $\hat{x}$ будем называть *решением* задачи (1), *минималью* (соответственно *максималью*) или *абсолютным минимумом (максимумом)*, если $f(x) \geqslant f(\hat{x})$ (соответственно $f(x) \leqslant f(\hat{x})$) для всех $x \in C$. Как правило, все задачи будем записывать как задачи минимизации, заменяя задачу $f(x) \to \sup$, $x \in C$, задачей $\tilde{f}(x) \to \inf$, $x \in C$, где $\tilde{f}(x) = -f(x)$. В тех случаях, когда хотим подчеркнуть, что для нас безразлично, рассматривается ли задача минимизации или максимизации, мы пишем $f(x) \to \mathrm{extr}$.

Далее, множество X у нас обычно бывает наделено *топологией*, т. е. в нем имеет смысл понятие близости элементов. Это можно сделать, например, задав в X набор окрестностей (как это стандартно делается в $\mathbf{R}^n$ или в нормированном пространстве). Если X — топологическое пространство, то $\hat{x}$ называется *локальным минимумом*, если существует такая окрестность U точки $\hat{x}$, что $\hat{x}$ — решение задачи $f(x) \to \inf$, $x \in C \cap U$. Аналогично определяется *локальный максимум*.

---

[1]) В теории экстремальных задач числовые функции часто называют *функционалами*.

1.2.2. Простейшие примеры формализации экстремальных задач. Приведем формализацию некоторых задач § 1.1. Начнем с задачи Евклида (см. п. 1.1.2, рис. 6). Из подобия треугольников DBE и ABC получаем: $h(x)/H = x/b$. Здесь x — сторона $|AF|$ параллелограмма $ADEF$, H — высота $\triangle ABC$, $h(x)$ — высота $\triangle BDE$, $b = |AC|$ — длина стороны AC. Площадь параллелограмма $ADEF$ равна $(H - h(x))\,x = H(b-x)\,x/b$. Теперь получаем следующую формализацию задачи Евклида:

$$\frac{H(b-x)\,x}{b} \to \sup,\ 0 \leqslant x \leqslant b, \Leftrightarrow x(x-b) \to \inf,\ x \in [0, b]. \quad (1)$$

Три элемента, из которых состоит всякая формализация, здесь суть:

$$X = \mathbf{R},\ f = H(b-x)\,x/b,\ C = [0,\ b].$$

Формализуем задачу Архимеда об изопифанных сегментах шаров (см. п. 1.1.2). Пусть h — высота шарового сегмента, R — радиус шара. Объем шарового сегмента, как известно из геометрии, равен $\pi h^2(R - h/3)$, а площадь поверхности $2\pi Rh$. Отсюда видно, что задачу Архимеда можно формализовать двояко:

$$\pi h^2\left(R - \frac{h}{3}\right) \to \sup,\ 2\pi Rh = a,\quad R \geqslant 0,\ 2R \geqslant h \geqslant 0, \quad (2)$$

или, исключая R из функционала в (2),

$$\frac{ha}{2} - \frac{\pi h^3}{3} \to \sup,\ 0 \leqslant h \leqslant \sqrt{\frac{a}{\pi}} \quad (2')$$

(последнее неравенство проистекает из-за того, что $h \leqslant 2R \Rightarrow a \geqslant \pi h^2$). В первом случае $X = \mathbf{R}^2_+$, $f = \pi h^2(R - h/3)$, $C = \{(R,\ h) \mid 2\pi Rh = a,\quad 2R \geqslant h\}$; во втором, полагая $X = [0, \sqrt{a/\pi}]$, имеем задачу без ограничений с функционалом $f = ha/2 - \pi h^3/3$.

Задача Кеплера о максимальном по объему цилиндре, вписанном в шар (см. п. 1.1.2), допускает такую очевидную формализацию:

$$2\pi x(1 - x^2) \to \sup;\ 0 \leqslant x \leqslant 1\ \ (X = \mathbf{R},\ C = [0,\ 1]). \quad (3)$$

Здесь шар имеет единичный радиус, а x—длина половины высоты цилиндра.

Вопрос о преломлении света на границе двух однородных сред, решаемый с помощью вариационного принципа Ферма (см. п. 1.1.3 рис. 19), сводится к такой задаче. Пусть две однородные среды разделены плоскостью $z=0$, причем скорость распространения света в верхнем полупространстве равна v_1, а в нижнем v_2. Мы ищем траекторию луча, идущего из точки $A=(0,\ 0,\ \alpha)$, $\alpha>0$, в точку $B=(\xi,\ 0,\ -\beta)$, $\beta>0$. По соображениям симметрии луч будет лежать в плоскости $y=0$. Пусть $C=(x,\ 0,\ 0)$—точка преломления луча. Тогда время распространения луча из A в B равно

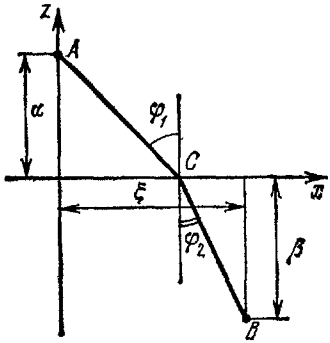


Рис 19.

$$T(x)=\frac{\sqrt{\alpha^2+x^2}}{v_1}+\frac{\sqrt{\beta^2+(\xi-x)^2}}{v_2}.$$

В соответствии с принципом Ферма координата искомой точки $\hat{x}$, где происходит преломление, находится из решения задачи

$$T(x)=\frac{\sqrt{\alpha^2+x^2}}{v_1}+\frac{\sqrt{\beta^2+(\xi-x)^2}}{v_2}\to\inf \qquad (X=C=\mathbf{R}). \quad (4)$$

Заметим, что получилась задача без ограничений.

Аналогично следующая задача без ограничений

$$|x-\xi_1|+|x-\xi_2|+|x-\xi_3|\to\inf, \quad (5)$$

где $X=C=\mathbf{R}^2$, ξ_1, ξ_2, ξ_3—три заданные точки плоскости $\mathbf{R}^2$, $|x|=\sqrt{x_1^2+x_2^2}$, является формализацией задачи Штейнера (см. п. 1.1.2). Отметим важную особенность функционала задачи (5)—он является выпуклой, но не всюду дифференцируемой функцией.

Задача Аполлония о кратчайшем расстоянии от точки $\xi=(\xi_1,\ \xi_2)$ до эллипса, задаваемого уравнением

$x_1^2/a_1^2+x_2^2/a_2^2=1$, допускает, очевидно, такую формализацию:

$$(x_1-\xi_1)^2+(x_2-\xi_2)^2 \to \inf, \quad \frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}=1 \qquad (6)$$

$$\left(X=\mathbf{R}^2,\ C=\left\{x \,\middle|\, \frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}=1\right\}\right).$$

Итак, мы познакомились с формализацией простейших задач. Формализация более сложных задач, возникающих в естествознании, технике или экономике, обычно составляет специальную и не слишком тривиальную проблему. Там сама формализация зависит от физических или иных гипотез. В следующем пункте мы покажем это на примере задачи Ньютона.

1.2.3. Формализация задачи Ньютона. Формализация этой задачи зависит, естественно, от законов сопротивления среды. Ньютон представлял себе среду (он называл ее «редкой») состоящей из неподвижных частиц фиксированной массы m, являющихся абсолютно упругими шарами. Примем и мы эту гипотезу.

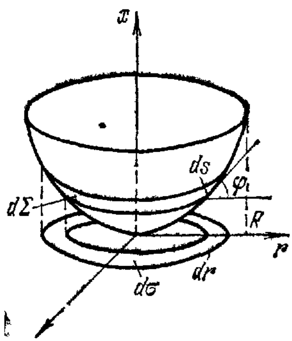


Рис. 20.

Пусть тело вращения вокруг оси x (рис. 20) движется в направлении, обратном оси x («вниз») в описанной нами «редкой» среде Ньютона со скоростью v. Элемент dr на оси r при вращении вокруг оси x описывает кольцо площади $d\sigma=2\pi r\,dr$, и этому кольцу соответствует пояс $d\Sigma$ на самом теле вращения. За время dt этот пояс «вытеснит» объем $dV=2\pi r\,dr\,v\,dt$. Пусть ρ — плотность среды. Тогда число частиц, ударившихся о слой, равно $N=\frac{\rho}{m}dV=\frac{\rho\, 2\pi r\,dr\,v}{m}dt$, где m — масса частицы. Подсчитаем силу dF, действующую на слой $d\Sigma$ за время dt. Пусть участок ds наклонен к оси r под углом φ. Отражаясь от $d\Sigma$, одна частица получает приращение импульса, равное $m(\mathbf{v}_2-\mathbf{v}_1)=-2mv\cos\varphi\cdot\mathbf{n}$, где $v=|\mathbf{v}_1|=|\mathbf{v}_2|$, $\mathbf{n}$ — единичный вектор нормали к $d\Sigma$, а $\varphi=\operatorname{arctg}\frac{dx}{dr}$ — угол ds с горизонталью. В силу третьего закона Ньютона тело получает противоположное приращение импульса $m2v\cos\varphi\cdot\mathbf{n}$, а за время dt таких приращений будет N, причем в силу симметрии компоненты импульса, ортогональные оси вращения, в сумме дают нуль, а осевая компонента суммарного приращения импульса равна

$$Nm\,2v\cos\varphi\cos\varphi = \frac{2\rho\pi r\,dr\,v\,dt}{m}\,m2v\cos^2\varphi =$$
$$= 4\rho\pi v^2 r\,dr\,dt\cos^2\varphi.$$

В силу второго закона Ньютона это выражение равно $dF\,dt$, откуда $dF = kr\,dr\cos^2\varphi$, $k = 4\rho\pi v^2$, а общая сила сопротивления равна

$$F = k\int_0^R \frac{r\,dr}{1+(dx/dr)^2}. \qquad (1)$$

Таким образом, заменив r на t и R на T, мы приходим к экстремальной задаче

$$\int_0^T \frac{t\,dt}{1+\dot{x}^2} \to \inf, \quad x(0) = 0, \quad x(T) = \xi. \qquad (2)$$

Легко сообразить, не решая задачи (2) (впервые это отметил Лежандр в 1788 г.), что нижняя грань в задаче равна нулю. Действительно, если выбрать ломаную $x(\cdot)$ с очень большой по модулю производной (рис. 21), то интеграл (2) будет очень мал. С другой стороны, для любой функции $x(\cdot)$ интеграл в (2) неотрицателен. Таким образом, нижняя грань значений интеграла равна нулю.

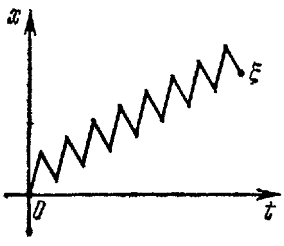


Рис. 21.

Отмеченное только что обстоятельство неоднократно вызывало критику, иногда и злорадную. Один из последних примеров — в книге Янга[1]) говорится: «Ньютон сформулировал вариационную задачу о теле вращения, испытывающем наименьшее сопротивление при движении в газе. Принятый им закон сопротивления физически абсурден, в результате чего поставленная им задача не имеет решения (чем более зазубрен профиль, тем меньше сопротивление)... Если выводы Ньютона хотя бы приблизительно

[1]) Янг Л. Лекции по вариационному исчислению и теории оптимального управления.— М.: Мир, 1974.

были верны, то мы не нуждались бы сегодня в дорогостоящих экспериментах в аэродинамических трубах». Задорно сказано! Многих людей утешает мысль, что и великие грубо ошибаются. Но так ли обстоит дело в данном случае? Прежде всего, заметим, что сам Ньютон не формализовал свою задачу—это за него сделали (и не вполне удачно) другие. Для правильной формализации надо учесть неявно подразумевавшуюся *монотонность* профиля (при зазубренном профиле частицы испытывают многократные отражения, что искажает всю картину). Требование монотонности делает задачу физически осмысленной. С учетом этого обстоятельства решение самого Ньютона не только «приблизительно» верно, но поразительно верно в деталях, о чем речь пойдет впереди. Более того, физические гипотезы, выдвинутые Ньютоном, и само его решение аэродинамической задачи оказались весьма актуальными в современной сверхзвуковой аэродинамике, когда на очередь дня встало построение сверхскоростных и высотных летательных аппаратов.

Допущение о монотонности приводит к следующей правильной формализации задачи Ньютона:

$$\int_0^T \frac{t\,dt}{1+\dot{x}^2} \to \inf, \quad x(0)=0, \quad x(T)=\xi, \quad \dot{x} \in \mathbf{R}_+. \qquad (3)$$

1.2.4. Различные формализации классической изопериметрической задачи и задачи о брахистохроне. Простейшая задача о быстродействии. Первые две из упомянутых в заглавии задач принадлежат к числу известнейших, но оказываются едва ли не самыми трудными для полного исследования. Мы дважды формализуем каждую из них—один раз традиционным, общеизвестным способом, другой—менее известным. Этим хотелось бы подчеркнуть принципиальную неединственность процедуры формализации. На самом деле, выбор удачной формализации составляет самостоятельную проблему и во многом успех при решении задачи зависит от искусства, которое здесь будет проявлено.

Начнем с классической изопериметрической задачи. Пусть длина кривой равна L, а сама кривая задана параметрически функциями $x(\cdot)$, $y(\cdot)$, причем в качестве параметра взята длина дуги s, отсчитываемая вдоль кривой от некоторой ее точки. Тогда в любой точке выполнено соотношение $\dot{x}^2(s)+\dot{y}^2(s)=1$, и, кроме того, $x(0)=x(L)$, $y(0)=y(L)$, так как кривая замкнута.

Для большей определенности в расположении искомой кривой можно потребовать также, чтобы ее центр тяжести попал в начало координат, т. е. чтобы имели место равенства $\int_0^L x(s)\,ds=\int_0^L y(s)\,ds=0$. Площадь S кривой $(x(\cdot),\ y(\cdot))$ равна $\int_0^L x\dot{y}\,ds$. Отсюда получается следующая формализация:

$$S=\int_0^L x\dot{y}\,ds\longrightarrow\sup;\quad \dot{x}^2(s)+\dot{y}^2(s)=1,$$
$$\int_0^L x(s)\,ds=\int_0^L y(s)\,ds=0,\quad x(0)=x(L),\quad y(0)=y(L). \tag{1}$$

Однако ту же задачу можно формализовать и по-другому. Представим себе самолет, пилот которого получил задачу облететь за заданное время возможно бо́льшую площадь и вернуться на свой аэродром. Если максимальная скорость самолета не зависит от направления полета, то данная задача приобретает следующую естественную формализацию.

Площадь должна быть максимальна $\Leftrightarrow \frac{1}{2}\int_0^T (x(t)\,v(t)-$ $-y(t)\,u(t))\,dt\longrightarrow\sup$, где $\dot{x}(t)=u(t)$, $\dot{y}(t)=v(t)$.

Максимальная скорость самолета равна $V\Leftrightarrow u^2+v^2\leqslant V^2$.

Самолет возвращается на свой аэродром $\Leftrightarrow x(0)=x(T)$, $y(0)=y(T)$.

Возможна и более общая постановка, когда максимальная скорость зависит от направления (например, при наличии ветра). Тогда мы получаем более общую задачу:

$$\frac{1}{2}\int_0^T (xv-yu)\,dt\longrightarrow\sup,\quad \dot{x}=u,\ \dot{y}=v,$$
$$x(0)=x(T),\ y(0)=y(T),\ (u,\ v)\in A, \tag{2}$$

где A — множество всех допустимых скоростей самолета.

Если A — круг, то мы, очевидно, приходим к классической изопериметрической задаче, если A — «сдвинуный» круг (что соответствует постоянному ветру), то получается известная *задача Чаплыгина.*

Приведем теперь самую традиционную формализацию задачи о брахистохроне. Введем, как и в п. 1.1.4 (рис. 16), в плоскости систему координат $(x,\ y)$ так, чтобы ось x была горизонтальна, а ось y направлена вниз. Не ограничивая себя в общности, можно считать, что точка A совпадает с началом координат. Пусть координаты точки $B-(x_1,\ y_1)$, $x_1>0$, $y_1>0$ (см. рис. 16) и $y(\cdot)$—функция, задающая уравнение кривой, соединяющей точки A и B. Напомним, что в соответствии с законом Галилея скорость тела M в точке $(x,\ y(x))$ зависит не от формы кривой $y(\cdot)$ в интервале $(0,\ x)$, а лишь от самой ординаты $y(x)$, причем эта скорость равна $\sqrt{2gy(x)}$, где g—ускорение силы тяжести. Следовательно, время T, требуемое для преодоления участка кривой длины $ds=\sqrt{dx^2+dy^2}$ на участке от $(x,\ y(x))$ до $(x+dx,\ y(x)+dy)$, равно $ds/\sqrt{2gy(x)}$. Отсюда получается следующая формализация задачи о брахистохроне:

$$\mathscr{J}(y(\cdot))=\int\limits_0^{x_1}\frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}\,dx\to\inf,\quad y(0)=0,\quad y(x_1)=y_1. \qquad (3)$$

Приведем другую формализацию задачи о брахистохроне, идейно близкую ко второй формализации классической изопериметрической задачи, следуя упоминавшейся статье И. Бернулли 1696 г., в которой он отталкивался от вариационного принципа Ферма.

Представим себе неоднородную среду, в которой скорость распространения света зависит лишь от «глубины» y по закону $v^2=2gy$. Тогда луч света в соответствии с вариационным принципом Ферма будет проходить путь от A до B в кратчайшее время. Так получается формализация задачи о брахистохроне в виде задачи о быстродействии:

$$\begin{gathered}T\to\inf,\ \dot{x}=\sqrt{y}\,u,\ \dot{y}=\sqrt{y}\,v,\ u^2+v^2=2g,\\ x(0)=y(0)=0,\quad x(T)=x_1,\quad y(T)=y_1.\end{gathered} \qquad (4)$$

Аналогично выглядит и формализация простейшей задачи о быстродействии (п. 1.1.7). Пусть масса тележки m, ее начальная координата x_0, а начальная скорость v_0. Внешнюю силу (силу тяги) обозначим через u, а текущую координату тележки — через $x(t)$. Тогда по закону Ньютона $m\ddot{x}=u$. Ограничение на тягу зададим в таком виде: $u\in[u_1, u_2]$. Отсюда

$$T\to\inf,\quad m\ddot{x}=u,\quad u\in[u_1\ u_2],\qquad (5)$$
$$x(0)=x_0,\quad \dot{x}(0)=v_0,\quad x(T)=\dot{x}(T)=0.$$

Получилась постановка, весьма похожая на (2) и (4).

Отметим одно важное обстоятельство. На самом деле мы «недоформализовали» еще наши задачи. Например, в формализации (3) не указана точно область определения функционала $\mathcal{J}$ и, следовательно, пока неизвестно, на каком классе кривых рассматривалась задача (т. е. не определено множество X п. 1.2.1), То же относится и к остальным формализациям этого пункта. Впрочем, «классикн» зачастую вообще не обращали внимания на аккуратную формализацию задач, а просто решали их «недоформализованными». Но мы хотим в дальнейшем быть педантичными и точными до конца, а потому придется заниматься этим несколько скучным делом — указывать всякий раз, в каком классе объектов ищется (илн было найдено) решение.

1.2.5. Формализация транспортной задачи и задачи о рационе. Начнем с транспортной задачи. Введем такие обозначения:

a_i — количество единиц продукта, находящегося на i-й базе, $1\leqslant i\leqslant m$,

b_j — потребность (в тех же единицах) в j-м магазине, $1\leqslant j\leqslant n$,

c_{ij} — стоимость перевозки единицы продукта из i-й базы в j-й магазин,

x_{ij} — планируемое количество единиц продукта для перевозки из i-й базы в j-й магазин.

Тогда стоимость перевозки равна $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}c_{ij}x_{ij}$, и ее нужно минимизировать. Ограничения при этом следующие:

а) $x_{ij}\in\mathbb{R}_+$ (очевидное ограничение на величину перевозки);

б) $\sum\limits_{j=1}^{n}x_{ij}\leqslant a_i$ (нельзя вывезти больше того, что есть);

в) $\sum\limits_{i=1}^{m}x_{ij}=b_j$ (нужно перевезти ровно столько, сколько необходимо).

В итоге получается следующая формализация:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \inf, \quad \sum_{j=1}^{n} x_{ij} \leqslant a_i, \quad \sum_{i=1}^{m} x_{ij} = b_j, \ x_{ij} \in \mathbf{R}_+. \qquad (1)$$

Задача о рационе формализуется так же просто. Пусть имеется n продуктов (зерно, молоко и т. п.) и m веществ (жиры, белки, углеводы и пр.). Допустим, что для полноценного питания необходимо b_j единиц j-го вещества. При этом a_{ij} есть содержание j-го вещества в единице i-го продукта, а c_i — цена единицы i-го продукта. Обозначив через x_i потребление i-го продукта, получаем задачу

$$\sum_{i=1}^{n} c_i x_i \to \inf, \quad \sum_{i=1}^{n} a_{ij} x_i \geqslant b_j, \quad x_i \geqslant 0. \qquad (2)$$

1.2.6. Основные классы экстремальных задач. Мы кратко упомянули уже в § 1.1, что в теории экстремальных задач выделилось несколько достаточно ясно очерченных классов задач. Прежде чем их описывать, проведем беглый обзор тех способов, какими задавались ограничения в задачах, формализованных выше.

Во-первых, нам встретились формализации, где ограничения отсутствовали вовсе (скажем, в задаче о преломлении света или задаче Штейнера). Во-вторых, случалось, что ограничения были заданы системой равенств (например, в задаче Аполлония, в задаче о брахистохроне в формализации (3) п. 1.2.4, где равенствами заданы краевые условия). В третьих, ограничения задавались неравенствами (например, в транспортной задаче). Наконец, в четвертых, некоторые ограничения записывались в виде включений (например, ограничение $\dot{x} \in \mathbf{R}_+$ в задаче Ньютона, ограничение $(u, v) \in A$, где $A = \{(u, v);\ u^2 + v^2 \leqslant 1\}$, в классической изопериметрической задаче в формализации (2) п. 1.2.4).

Подчеркнем некоторую условность такого разделения. Скажем, ограничение $\dot{x} \in \mathbf{R}_+$ в задаче Ньютона можно было бы записать в виде неравенства $\dot{x} \geqslant 0$, а ограничение $(u, v) \in A$ в классической изопериметрической задаче можно было бы записать в виде неравенства $u^2 + v^2 \leqslant 1$. Наоборот, всякое неравенство $f(x) \leqslant 0$ можно заменить на равенство $f(x) + u = 0$ и включение $u \in \mathbf{R}_+$ и т. д.

Тем не менее с точки зрения, принятой в этой книге, разделение ограничений на равенства и неравенства, с одной стороны, и включения — с другой, имеет свой смысл. В курсах анализа приводится (а в § 1.3 об этом будет сказано подробно) правило множителей Лагранжа для решения задач на «условный экстремум». Как известно, применение этого правила начинается с составления «функции Лагранжа», в которую входят как исследуемый функционал, так и функции, задающие ограничения. Может оказаться, что по разным причинам некоторые ограничения выгодно не включать в функцию Лагранжа. Так вот, в виде включений мы и выделяем именно те ограничения, которые при решении соответствующей задачи в функцию Лагранжа не войдут. При этом, как и при формализации задачи (где бывает много способов и выбор удачного зависит от искусства исследователя), в вопросе о разбиении ограничений нет однозначности. Перейдем к описанию основных классов экстремальных задач.

В дальнейшем с достаточно общих позиций будут рассмотрены следующие четыре класса.

I. Гладкие задачи с ограничениями типа равенств и неравенств. Здесь класс допустимых элементов X будет обычно нормированным пространством[1], и ограничение C задается равенством $F(x)=0$, где F — отображение X в другое нормированное пространство Y, и конечным числом неравенств $f_i(x)\leqslant 0$, $i=1, \ldots, m$. В итоге получается класс задач

$$f_0(x)\to \inf, \quad F(x)=0, \quad f_i(x)\leqslant 0, \quad i=1, \ldots, m. \tag{1}$$

При этом предполагается, что функции f_i, $i=1, 2, \ldots, m$, и отображение F обладают некоторыми свойствами гладкости. Гладкую задачу $f_0(x)\to\inf$ будем называть *элементарной гладкой задачей*[2].

II. Классическое вариационное исчисление. Здесь традиционным классом допустимых элементов является банахово пространство $X=C^1([t_0, t_1], \mathbf{R}^n)$ непрерывно дифференцируемых n-мерных вектор-функций $x(\cdot)=(x_1(\cdot), \ldots, x_n(\cdot))$, в котором норма задается

---

[1]) Термины «нормированное пространство» и «банахово пространство» используются здесь и ниже только для корректности постановки задачи. Точные определения читатель найдет в п. 2.1.1. В задаче (1) можно для простоты считать, что $x=(x_1, \ldots, x_n)$ — вектор в n-мерном арифметическом пространстве $\mathbf{R}^n$.

[2]) Задачи на максимум или с неравенствами другого знака ($\geqslant$) легко приводятся к виду (1), см. § 3.2.

формулами

$$\|x(\cdot)\|_1 = \max(\|x(\cdot)\|_0, \ \|\dot{x}(\cdot)\|_0),$$
$$\|x(\cdot)\|_0 = \max_{1 \leqslant i \leqslant n}\left(\max_{t \in [t_0, t_1]} |x_i(t)|\right).$$

Функционалы в задачах классического вариационного исчисления бывают обычно следующих типов:

— *интегральные*, т. е. функционалы вида

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x, \dot{x})\,dt =$$
$$= \int_{t_0}^{t_1} L(t, x_1(t), \ldots, x_n(t), \dot{x}_1(t), \ldots, \dot{x}_n(t))\,dt; \quad (2)$$

— *терминальные*, т. е. функционалы вида

$$\mathcal{T}(x(\cdot)) = l(x(t_0), x(t_1)) =$$
$$= l(x_1(t_0), \ldots, x_n(t_0), x_1(t_1), \ldots, x_n(t_1)); \quad (3)$$

— *смешанные* функционалы вида

$$\mathcal{B}(x(\cdot)) = \mathcal{J}(x(\cdot)) + \mathcal{T}(x(\cdot)). \quad (4)$$

(В дальнейшем функционалы (4) мы называем также *функционалами Больца.*)

Ограничения в задачах классического вариационного исчисления обычно распадаются на две части:

— *дифференциальные связи* вида

$$M(t, x(t), \dot{x}(t)) = 0 \Leftrightarrow M_i(t, x_1(t), \ldots, x_n(t), \dot{x}_1(t), \ldots$$
$$\ldots, \dot{x}_n(t)) = 0, \quad i = 1, 2, \ldots, p; \quad (5)$$

— *граничные условия* вида

$$\psi(x(t_0), x(t_1)) =$$
$$= 0 \Leftrightarrow \psi_j(x_1(t_0), \ldots, x_n(t_0), x_1(t_1), \ldots, x_n(t_1)). \quad (6)$$
$$j = 1, 2, \ldots, s.$$

В (2) и (5) L, M_i и в (3) и (6), l, ψ_j — гладкие функции $2n+1$ и $2n$ переменных соответственно. Задача

$$\mathcal{J}(x(\cdot)) \to \inf, \quad M(t, x, \dot{x}) = 0, \quad \psi(x(t_0), x(t_1)) = 0 \quad (7)$$

называется *задачей Лагранжа*. Задача

$$\mathcal{B}(x(\cdot)) \to \inf, \quad M(t, x, \dot{x}) = 0, \ \psi(x(t_0), x(t_1)) = 0$$

называется *задачей Больца.*

Задача

$$\mathscr{T}(x(\cdot)) \to \inf, \quad M(t, x, \dot{x}) = 0, \quad \psi(x(t_0), x(t_1)) = 0$$

называется *задачей Майера.*

Задачу без ограничений

$$\mathscr{B}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt + l(x(t_0), x(t_1)) \to \inf \quad (8)$$

будем называть *элементарной задачей Больца.*

Задача

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt \to \inf, \quad x(t_0) = x_0, \quad x(t_1) = x_1 \quad (9)$$

называется *простейшей векторной задачей классического вариационного исчисления*, а в случае, если $n = 1$, то — *простейшей задачей классического вариационного исчисления*. Для простоты здесь мы ограничились задачами с фиксированным временем.

Более общую постановку, в которой функционал и ограничения зависят также от переменных t_0 и t_1, читатель найдет в гл. IV.

III. Задачи выпуклого программирования. Здесь класс допустимых элементов X является линейным пространством, а ограничение C задается системой равенств $F(x) = 0$, (где $F\colon X \to Y$, Y — другое линейное пространство), неравенств $f_i(x) \leqslant 0$, $i = 1, 2, \ldots, m$, и включений $x \in A$. В итоге получается класс задач

$$f_0(x) \to \inf, \quad F(x) = 0, \quad f_i(x) \leqslant 0, \quad i = 1, \ldots, m, \quad x \in A. \quad (10)$$

При этом предполагается, что функции f_i, $i = 0, 1, \ldots, m$, выпуклы, отображение F аффинно (т. е. $F(x) = \Lambda x + \eta$, где η — фиксированный вектор, а Λ — линейный оператор из X в Y), а A — выпуклое множество. Если в (10) все функции f_i линейны, а A — некоторый стандартный конус, то задачу (10) называют *задачей линейного программирования*. Если в (10) ограничения отсутствуют, то задачу $f_0(x) \to \inf$ с выпуклой функцией f_0 мы называем *элементарной выпуклой задачей без ограничений.*

IV. Задачи оптимального управления. В этой книге будет рассмотрен следующий класс задач оптимального управления, где $x \in \mathbf{R}^n$, $u \in \mathbf{R}^r$:

$$\mathscr{Y}(x(\cdot), u(\cdot), t_0, t_1) = \int_{t_0}^{t_1} f(t, x(t), u(t))\,dt + g(t_0, x(t_0), t_1, x(t_1)) \to \inf, \quad (11)$$

$$\dot{x} = \varphi(t, x, u), \quad \psi(t_0, x(t_0), \quad t_1, x(t_1)) = 0, \quad u \in \mathfrak{U}.$$

При этом в (11) моменты t_0 и t_1, вообще говоря, не фиксированы, все функции f: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^r \to \mathbf{R}$, φ: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^r \to \mathbf{R}^n$, ψ: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^s$, g: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}$ предполагаем непрерывными по совокупности переменных и непрерывно дифференцируемыми по переменным t и x; множество $\mathfrak{U}$ — некоторое, вообще говоря, произвольное подмножество $\mathbf{R}^r$.

Для полного описания задачи осталось лишь объяснить, что же составляет здесь класс допустимых элементов. На первых порах будем рассматривать совокупность вектор-функций $(x(\cdot), u(\cdot))$, где $u(\cdot)$ определена и кусочно-непрерывна на $[t_0, t_1]$, причем для всех t выполнено включение $u(t) \in \mathfrak{U}$, а $x(\cdot)$ непрерывна на $[t_0, t_1]$ и дифференцируема во всех точках, кроме тех, где $u(\cdot)$ терпит разрыв; при этом во всех точках дифференцируемости $x(\cdot)$ выполнено равенство $\dot{x}(t) = \varphi(t, x(t), u(t))$.

Задачи, в которых функционалом является t_1, называются *задачами о быстродействии.*

* * *

Теперь можно посмотреть, в какие классы попадают задачи пп. 1.2.1—1.2.5.

Задачу Евклида (см. (1) п. 1.2.2) можно отнести и к гладким задачам и к задачам выпуклого программирования. Задача Архимеда в формализациях (2) и (2′) п. 1.2.2 и задача Кеплера (3) п. 1.2.2 относятся к числу гладких задач. Задача о преломлении света (см. (4) п. 1.2.2)—это и элементарная гладкая задача, и элементарная выпуклая задача. Задача Штейнера (см. (5) п. 1.2.2)—элементарная выпуклая задача. Задача Аполлония (см. (6) п. 1.2.2)—гладкая задача с ограничениями типа равенств. Задача Ньютона (см. (3) п. 1.2.3)—задача оптимального управления. Классическая изопериметрическая задача в формализации (1) п. 1.2.4 относится к классическому вариационному исчислению, а в формализации (2) — к оптимальному управлению. Задача о брахистохроне (см. (3) п. 1.2.4) — простейшая задача классического вариационного исчисления; эта же задача в формализации (4) п. 1.2.4 — задача о быстродействии оптимального управления. Транспортная задача и задача о рационе (см. (1) и (2) п. 1.2.5) — задачи линейного программирования; простейшая задача о быстродействии — задача оптимального управления.

Итак, задачи формализованы и классифицированы. Посмотрим теперь, что может дать для их решения аппарат анализа.

§ 1.3. Правило множителей Лагранжа и теорема Куна—Таккера

1.3.1. Теорема Ферма. Первый общий аналитический прием решения экстремальных задач был разработан Пьером Ферма. Открыт он был, по-видимому, в 1629 г., но впервые достаточно полно изложен в письме к Робервалю в 1638 г. Можно посоветовать читателю обратиться к книге Декарта[1]), где приведено это письмо, и самому вникнуть в первоначальную мысль Ферма. На современном языке (правда, у Ферма лишь для полиномов) прием Ферма сводится к тому, что в точке экстремума $\hat{x}$ в задаче без ограничений $f(x) \to \mathrm{extr}$ должно иметь место равенство $f'(\hat{x}) = 0$. Как мы помним, первый намек на этот результат содержится в словах Кеплера из «Стереометрии винных бочек».

Точный смысл рассуждения Ферма приобрели через 46 лет, когда в 1684 г. появилась работа Лейбница, в которой закладывались основы математического анализа. Уже само заглавие этой работы, которое начинается так: «Nova methodus pro maximis et minimis...» («Новый метод нахождения наибольших и наименьших значений...»), показывает, какую важную роль сыграла задача о нахождении экстремумов в становлении современной математики. В своей статье Лейбниц не только получает в качестве необходимого условия соотношение $f'(\hat{x}) = 0$

---

[1]) Р. Декарт. Геометрия./ С приложением избранных работ П. Ферма и переписки Декарта.— М. — Л.: ГОНТИ, 1938, с. 154.

(сейчас этот результат называют *теоремой Ферма*), но и употребляет второй дифференциал для различения максимума и минимума. Следует, конечно, напомнить, что большинство излагаемых Лейбницем фактов было к тому времени известно также и Ньютону. Но его работа «Метод флюксий», завершенная в основном к 1671 г., была опубликована лишь в 1736 г.

Перейдем теперь к самой теореме Ферма, напомнив предварительно некоторые факты из анализа. Начнем с одномерного случая, когда функции определены на вещественной прямой $\mathbf{R}$. Функция $f\colon \mathbf{R} \to \mathbf{R}$ одного переменного называется *дифференцируемой в точке* $\hat{x}$, если найдется такое число α, что

$$f(\hat{x}+\lambda) = f(\hat{x}) + \alpha\lambda + r(\lambda),$$

где $r(\lambda) = o(|\lambda|)$, т. е. для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что из $|\lambda| \leqslant \delta$ следует $|r(\lambda)| \leqslant \varepsilon |\lambda|$. Число α называется *производной* f *в точке* $\hat{x}$ и обозначается $f'(\hat{x})$. Таким образом, $f'(\hat{x}) = \lim\limits_{\lambda \to 0} (f(\hat{x}+\lambda) - f(\hat{x}))/\lambda$.

Теорема Ферма. *Пусть f — функция одного переменного, дифференцируемая в точке $\hat{x}$. Если $\hat{x}$ — точка локального экстремума, то*

$$f'(\hat{x}) = 0. \tag{1}$$

Точки $\hat{x}$, в которых выполнено соотношение (1), называются *стационарными*.

В соответствии с общими определениями п. 1.2.1 точка $\hat{x}$ доставляет функции f локальный минимум (максимум), если найдется такое $\varepsilon > 0$, что из неравенства $|x - \hat{x}| < \varepsilon$ следует неравенство $f(x) \geqslant f(\hat{x})$ $(\leqslant f(\hat{x}))$. Согласно теореме Ферма точки локального экстремума (максимума или минимума) являются стационарными. Обратное, разумеется, неверно: например, $f(x) = x^3$, $\hat{x} = 0$.

Доказательство. Допустим, что $\hat{x}$ — локальный минимум функции f, но $f'(\hat{x}) = \alpha \neq 0$. Пусть для определенности $\alpha < 0$. Задав $\varepsilon = |\alpha|/2$, найдем из определения производной такое $\delta > 0$, что из $|\lambda| < \delta$ следует $|r(\lambda)| < |\alpha||\lambda|/2$. Тогда для $0 < \lambda < \delta$ получаем

$$f(\hat{x}+\lambda) = f(\hat{x}) + \alpha\lambda + r(\lambda) \leqslant$$
$$\leqslant f(\hat{x}) + \alpha\lambda + \frac{|\alpha|\lambda}{2} = f(\hat{x}) - \frac{|\alpha|\lambda}{2} < f(\hat{x}),$$

т. е. $\hat{x}$ не является точкой локального минимума f. Противоречие доказывает теорему. ■

Для многих переменных (а тем более для «бесконечного числа переменных») имеется несколько разных определений производных. Более подробно говорится об этом в гл. II (см. п. 2.2.1). Здесь мы напомним лишь основное определение (в бесконечномерном случае принадлежащее Фреше).

Функция $f\colon \mathbf{R}^n \to \mathbf{R}$ n переменных называется *дифференцируемой в точке* $\hat{x}$, если найдутся числа $(\alpha_1, \ldots, \alpha_n)$ (кратко обозначим их через α) такие, что

$$f(\hat{x}+h) = f(\hat{x}) + \sum_{i=1}^{n} \alpha_i h_i + r(h),$$

где $|r(h)| = o(|h|)$, т. е. для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что из $|h| = \sqrt{h_1^2 + \ldots + h_n^2} < \delta$ следует $|r(h)| \leqslant \varepsilon |h|$. Набор $\alpha = (\alpha_1, \ldots, \alpha_n)$ называется *производной* f *в точке* $\hat{x}$ и обозначается также $f'(\hat{x})$. Подчеркнем, что $f'(\hat{x})$ — это набор из n чисел. При этом число $\alpha_i = \lim\limits_{\lambda \to 0} [f(\hat{x}+\lambda e_i) - f(\hat{x})]/\lambda$, где $e_i = (0, \ldots, 1, \ldots, 0)$ (единица стоит на i-м месте), называется *i-й частной производной* f и обозначается $f_{x_i}(\hat{x})$ или $\frac{\partial f}{\partial x_i}(\hat{x})$. Таким образом, $f'(\hat{x}) = (f_{x_1}(\hat{x}), \ldots, f_{x_n}(\hat{x}))$. Соотношение $f'(\hat{x}) = 0$ означает, что $f_{x_1}(\hat{x}) = \ldots = f_{x_n}(\hat{x}) = 0$.

Из «одномерной» теоремы Ферма вытекает очевидное следствие:

Следствие (теорема Ферма для функций n переменных). *Пусть f — функция n переменных, дифференцируемая в точке $\hat{x}$. Если $\hat{x}$ — точка локального минимума функции f, то*

$$f'(\hat{x}) = 0 \Leftrightarrow f_{x_i}(\hat{x}) = 0, \quad i = 1, \ldots, n. \tag{2}$$

Точки $\hat{x} \in \mathbf{R}^n$, в которых имеют место равенства (2), также называются *стационарными*.

Доказательство следствия. Если $\hat{x}$ доставляет экстремум функции f, то точка нуль должна быть экстремумом функции $\varphi_i(\lambda) = f(\hat{x}+\lambda e_i)$. По теореме Ферма $\varphi_i'(0) = 0$. Но $\varphi_i'(0) \overset{\text{def}}{=} f_{x_i}(\hat{x})$.

1.3.2. Правило множителей Лагранжа. Этот пункт посвящен гладким конечномерным задачам. Еще раз напомним о причудливости исторического развития нашей темы. Вслед за методом решения одномерных экстремальных задач наступила эра классического вариационного исчисления. Для задач вариационного исчисления с ограничениями, т. е. в бесконечномерной ситуации, Лагранж сформулировал в «Аналитической механике» (1788 г.)[1]) свое знаменитое правило множителей, и только спустя примерно десятилетие в «Теории аналитических функций» в 1797 г. он применил его к конечномерным задачам.

Конечномерной гладкой экстремальной задачей с ограничениями типа равенств (или задачей об условном экстремуме) называется задача

$$f_0(x) \to \mathrm{extr}, \quad f_1(x) = \ldots = f_m(x) = 0. \qquad (1)$$

Функции $f_k\colon \mathbf{R}^n \to \mathbf{R}$, $k = 0, 1, \ldots, m$, должны здесь обладать некоторыми свойствами гладкости (т. е. дифференцируемости). В этом пункте будем предполагать, что в некоторой окрестности U пространства $\mathbf{R}^n$ все функции f_k непрерывно дифференцируемы (в том смысле, что все частные производные $\partial f_k/\partial x_i$ существуют и непрерывны в U).

Точка $\hat{x} \in U$, $f_k(\hat{x}) = 0$, $k = 1, 2, \ldots, m$, доставляет локальный минимум (максимум) в задаче (1) в том случае, когда найдется такое $\varepsilon > 0$, что если точка $x \in U$ удовлетворяет всем ограничениям ($f_i(x) = 0$, $i = 1, \ldots, m$) и $|x - \hat{x}| < \varepsilon$, то $f_0(x) \geqslant f_0(\hat{x})$ (соответственно $f_0(x) \leqslant \leqslant f_0(\hat{x})$).

Функцию

$$\mathscr{L} = \mathscr{L}(x, \lambda, \lambda_0) = \sum_{k=0}^{m} \lambda_k f_k(x), \qquad (2)$$

где $\lambda = (\lambda_1, \ldots, \lambda_m)$, называют *функцией Лагранжа* задачи (1), числа $\lambda_0, \lambda_1, \ldots, \lambda_m$ — *множителями Лагранжа*.

Правило множителей Лагранжа. 1. *Пусть в задаче* (1) *все функции* $f_0, f_1, \ldots, f_m$ *непрерывно дифференцируемы в окрестности точки* $\hat{x}$. *Если* $\hat{x}$ — *точка локального экстремума в задаче* (1), *то найдутся мно-*

---

[1]) Лагранж Ж. Аналитическая механика.— М.— Л.: Гостехиздат, 1950.

жители Лагранжа $\hat{\lambda}=(\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$ *и* $\hat{\lambda}_0$, *не равные одновременно нулю и такие, что выполняются условия стационарности функции Лагранжа по* x

$$\mathscr{L}_x(\hat{x}, \hat{\lambda}, \hat{\lambda}_0)=0 \Leftrightarrow \mathscr{L}_{x_i}(\hat{x}, \hat{\lambda}_1, \ldots, \hat{\lambda}_m, \hat{\lambda}_0)=0, \quad (3)$$
$$i=1, 2, \ldots, n.$$

2. *Для того чтобы* $\hat{\lambda}_0 \neq 0$, *достаточно, чтобы векторы* $f_1'(\hat{x}), \ldots, f_m'(\hat{x})$ *были линейно независимы.*

Таким образом, для определения $\hat{x}$, $\hat{\lambda}$ и $\hat{\lambda}_0$ получается $n+m$ уравнений:

$$\frac{\partial}{\partial x_i}\left(\sum_{k=0}^{m}\lambda_k f_k(x)\right)=0, \quad i=1, \ldots, n, \quad f_1(x)=\ldots=f_m(x)=0 \qquad (4)$$

с $n+m+1$ неизвестными. Следует учесть, что множители Лагранжа определены при этом с точностью до пропорциональности. Если известно, что $\hat{\lambda}_0 \neq 0$ (а это—важнейший случай, ибо если $\hat{\lambda}_0=0$, то соотношения (3) отражают лишь вырожденность ограничений и не связаны с функционалом), то можно, умножив все $\hat{\lambda}_i$ на константу, добиться равенства $\hat{\lambda}_0=1$. Тогда число уравнений сравняется с числом неизвестных.

В более симметричной форме уравнения (4) записываются так:

$$\mathscr{L}_x=0, \quad \mathscr{L}_\lambda=0. \qquad (4')$$

Их решения называются *стационарными точками* задачи (1).

Со времен Лагранжа (слова Лагранжа приводятся чуть ниже) почти на протяжении целого века правило множителей формулировалось с $\hat{\lambda}_0=1$, хотя без дополнительных предположений, например линейной независимости, в таком виде оно неверно. Для подтверждения сказанного рассмотрим задачу

$$x_1 \to \inf, \quad x_1^2+x_2^2=0.$$

Очевидиым и единственным ее решением является $\hat{x}=(0, 0)$, ибо это—единственная допустимая точка. Попробуем теперь составить функцию Лагранжа с $\hat{\lambda}_0=1$ и применить далее алгоритм (3). Имеем $\mathscr{L}=x_1+\lambda(x_1^2+x_2^2)$, откуда

$$\frac{\partial \mathscr{L}}{\partial x_1}=0 \Rightarrow 2\lambda x_1+1=0, \quad \frac{\partial \mathscr{L}}{\partial x_2}=0 \Rightarrow 2\lambda x_2=0$$

и первое из этих уравнений несовместно с уравнением $x_1^2+x_2^2=0$.

В качестве условия регулярности, гарантирующего, что $\hat{\lambda}_0 \neq 0$, обычно используют линейную независимость производных $f_1'(\hat{x}), \ldots, f_m'(\hat{x})$ (см. выше утверждение 2 теоремы). Однако проверка этого условия обычно сложнее, чем проверка непосредственно из уравнений (3) того, что $\hat{\lambda}_0$ не может равняться нулю (см. решения задач в § 1.6). Поэтому приведенная выше формулировка теоремы, не содержащая никаких дополнительных предположений, кроме гладкости, очень удобна.

Доказательство теоремы опирается на одну из важнейших теорем конечномерного дифференциального исчисления—теорему об обратной функции [14, т. 1, с. 455], [9, т. 1, с. 259].

Теорема об обратной функции. *Пусть $\psi_1(x_1, \ldots, x_s), \ldots, \psi_s(x_1, \ldots, x_s)$—$s$ функций s переменных непрерывно дифференцируемых в некоторой окрестности точки $\hat{x}$. Пусть при этом якобиан*

$$\mathscr{J}=\det\left(\frac{\partial \psi_i(\hat{x})}{\partial x_j}\right)$$

отличен от нуля. Тогда существуют такие $\varepsilon_0>0$ и $\delta_0>0$, что для любого $\eta=(\eta_1, \ldots, \eta_s)$, $|\eta| \leqslant \varepsilon_0$ найдется ξ, $|\xi| \leqslant \delta_0$, такое, что $\psi(\hat{x}+\xi)=\psi(\hat{x})+\eta$, и при этом $\xi \to 0$, когда $\eta \to 0$.

Более общие варианты этой теоремы будут доказаны в § 2.3.

Доказательство правила множителей Лагранжа. Пусть точка $\hat{x}$ доставляет локальный минимум в задаче (1). Возможно одно из двух: или векторы $f_0'(\hat{x}), \ldots, f_m'(\hat{x})$ линейно зависимы, или эти векторы линейно независимы.

А) Пусть векторы линейно зависимы. Тогда ненулевые $\hat{\lambda}_k$, $k=0, 1, \ldots, m$, для которых имеют место условия (3), можно подобрать по определению.

Б) Пусть векторы линейно независимы. Приведем это к противоречию с тем, что $\hat{x}$—локальный минимум в задаче (1).

Рассмотрим отображение $\Phi(x)=(f_0(x)-f_0(\hat{x}), f_1(x), \ldots$ $\ldots, f_m(x))$. По условию векторы $f_0'(\hat{x}), \ldots, f_m'(\hat{x})$ линейно независимы, т. е. ранг матрицы

$$A=\begin{pmatrix} \frac{\partial f_0(\hat{x})}{\partial x_1}, & \ldots, & \frac{\partial f_0(\hat{x})}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m(\hat{x})}{\partial x_1}, & \ldots, & \frac{\partial f_m(\hat{x})}{\partial x_n} \end{pmatrix}$$

равен $m+1$. Допустим для определенности, что первые $m+1$ столбцов матрицы A линейно независимы, т. е.

$$\det\begin{pmatrix} \frac{\partial f_0(\hat{x})}{\partial x_1}, & \ldots, & \frac{\partial f_0(\hat{x})}{\partial x_{m+1}} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m(\hat{x})}{\partial x_1}, & \ldots, & \frac{\partial f_m(\hat{x})}{\partial x_{m+1}} \end{pmatrix}\neq 0.$$

Следовательно, функции

$$\psi_1(x_1, \ldots, x_{m+1})= \\ =f_0(x_1, \ldots, x_{m+1}, \hat{x}_{m+2}, \ldots, \hat{x}_n)-f_0(\hat{x}_1, \ldots, \hat{x}_n),$$

$$\psi_2(x_1, \ldots, x_{m+1})=f_1(x_1, \ldots, x_{m+1}, \hat{x}_{m+2}, \ldots, \hat{x}_n), \ldots,$$

$$\psi_{m+1}(x_1, \ldots, x_{m+1})=f_m(x_1, \ldots, x_{m+1}, \hat{x}_{m+2}, \ldots, \hat{x}_n)$$

удовлетворяют условиям теоремы об обратной функции. По этой теореме для любого ε, $0<\varepsilon<\varepsilon_0$, найдутся $x_1(\varepsilon), \ldots, x_{m+1}(\varepsilon)$, такие, что

$$\begin{gathered} f_0(x_1(\varepsilon), \ldots, x_{m+1}(\varepsilon), \hat{x}_{m+2}, \ldots, \hat{x}_n)-f_0(\hat{x})=-\varepsilon, \\ f_1(x_1(\varepsilon), \ldots, x_{m+1}(\varepsilon), \hat{x}_{m+2}, \ldots, \hat{x}_n)=0, \\ \cdots\cdots\cdots\cdots \\ f_m(x_1(\varepsilon), \ldots, x_{m+1}(\varepsilon), \hat{x}_{m+2}, \ldots, \hat{x}_n)=0, \end{gathered} \tag{5}$$

причем $x_i(\varepsilon)\to\hat{x}_i$ при $\varepsilon\to 0$, $i=1, 2, \ldots, m+1$. Но из (5) прямо следует, что точка $\hat{x}$ не является точкой локального минимума. Искомое противоречие получено. Утверждение 2 теоремы очевидно. ■

Обобщение правила множителей Лагранжа на бесконечномерные задачи с равенствами и неравенствами будет дано в гл. III.

А теперь—еще несколько слов об истории. Первое упоминание о правиле множителей содержится в работах

Эйлера по изопериметрическим задачам (1744 г.). Затем оно было выдвинуто Лагранжем в его «Аналитической механике» (1788 г.) для широкого класса задач вариационного исчисления (так называемых задач Лагранжа, о чем будет еще сказано). В 1797 г. в книге «Теория аналитических функций» Лагранж затрагивает вопрос и о конечномерных задачах[1]).

Он пишет: «On peut les réduire à ce principe générale. Lors qu'une fonction de plusieurs variables doit être un maximum ou minimum, et qu'il y a entre ces variables une ou plusieurs équations, il suffira d'ajouter à la fonction proposée les fonctions qui doivent être nulles, multipliées chacune par une quantité indéterminée, et là chercher ensuite le maximum ou minimum comme si les variables étaient indépendantes; les équations qu'on trouvées, serviront à déterminer toutes les inconnues».

«Можно высказать следующий общий принцип.

Если ищется максимум или минимум некоторой функции многих переменных при условии, что между этими переменными имеется связь, задаваемая одним или несколькими уравнениями, нужно прибавить к минимизируемой функции функции, задающие уравнения связи, умноженные на неопределенные множители, и искать затем максимум или минимум построенной суммы, как если бы переменные были независимы. Полученные уравнения, присоединенные к уравнениям связи, послужат для определения всех неизвестных».

Не будет преувеличением сказать, что основная часть этой книги посвящена раскрытию замысла Лагранжа в применении к задачам разной природы. Изложим здесь еще раз основную мысль Лагранжа. Пусть требуется найти экстремум в задаче (1). Тогда следует составить функцию Лагранжа (2) и рассмотреть задачу $\mathscr{L}(x, \hat{\lambda}, \hat{\lambda}_0) \to$ $\to \text{extr}$ без ограничений. Необходимое условие в этой задаче без ограничений в соответствии с теоремой Ферма для функций n переменных дает нужные уравнения $\mathscr{L}_x = 0$. Итак, в соответствии с принципом Лагранжа уравнения для экстремума в задаче с ограничениями совпадают с уравнениями для задачи $\mathscr{L}(x, \hat{\lambda}, \hat{\lambda}_0) \to \text{extr}$ без ограничений при надлежащем выборе множителей Лагранжа $\hat{\lambda}, \hat{\lambda}_0$.

*) Lagrange J. L. Théorie des fonctions analytiques, Paris, 1813.

Мы увидим далее, что для очень большого числа задач общий замысел Лагранжа оказывается правильным.

1.3.3. Теорема Куна — Таккера. В этом пункте мы рассмотрим задачи выпуклого программирования, для которых идея Лагранжа приобретает наиболее завершенную форму. Этот класс задач стали изучать сравнительно недавно. Основы теории линейного (а это частный случай выпуклого) программирования были заложены в работе Л. В. Канторовича в 1939 г. Теорема Куна—Таккера — основной результат этого параграфа—была доказана в 1951 г.

Рассмотрим следующую экстремальную задачу (*задачу выпуклого программирования*):

$$f_0(x) \to \inf, \quad f_i(x) \leqslant 0, \quad i=1, \ldots, m, \; x \in A, \tag{1}$$

где X — линейное пространство (не обязательно конечномерное), f_i — выпуклые функции на X, A — выпуклое подмножество X. Отметим, и это существенно, что (1) — это задача минимизации, а не максимизации.

Напомним, что множество C, лежащее в линейном пространстве называется *выпуклым*, если вместе с любыми двумя своими точками x и y оно содержит также и весь отрезок $[x, y] = \{z \mid z = \alpha x + (1-\alpha) y, \; 0 \leqslant \alpha \leqslant 1\}$. Функция $f\colon\; X \to \mathbf{R}$ называется *выпуклой*, если для любых x и $y \in X$ выполнено *неравенство Иенссена*

$$f(\alpha x + (1-\alpha) y) \leqslant \alpha f(x) + (1-\alpha) f(y), \quad \forall \alpha, \quad 0 \leqslant \alpha \leqslant 1,$$

или, что то же самое, если «надграфик» f, т. е. множество

$$\operatorname{epi} f = \{(\alpha, x) \in \mathbf{R} \times X \mid \alpha \geqslant f(x)\}$$

является выпуклым множеством в произведении $\mathbf{R} \times X$.

Функцию

$$\mathscr{L} = \mathscr{L}(x, \lambda, \lambda_0) = \sum_{k=0}^{m} \lambda_k f_k(x), \tag{2}$$

где $\lambda = (\lambda_1, \ldots, \lambda_m)$ называют *функцией Лагранжа* задачи (1), числа $\lambda_0, \ldots, \lambda_m$ — *множителями Лагранжа*. При этом в выражении (2) не нашло места ограничение $x \in A$.

Теорема Куна—Таккера. 1. *Пусть X — линейное пространство, $f_i\colon X \to \mathbf{R}$, $i = 0, 1, \ldots, m$, — выпуклые функции на X, A — выпуклое подмножество X.*

Если $\hat{x}$ является решением задачи (1), *то найдутся множители Лагранжа $\hat{\lambda}_0$, $\hat{\lambda}$, не равные одновременно нулю и такие, что:*

а) $\min\limits_{x \in A} \mathscr{L}(x, \hat{\lambda}, \hat{\lambda}_0) = \mathscr{L}(\hat{x}, \hat{\lambda}, \hat{\lambda}_0)$ (*принцип минимума*);

б) $\hat{\lambda}_i \geqslant 0$, $i = 0, 1, \ldots, m$ (*условие неотрицательности*);

в) $\hat{\lambda}_i f_i(\hat{x}) = 0$, $i = 1, 2, \ldots, m$ (*условие дополняющей нежесткости*).

2. *Если $\hat{\lambda}_0 \neq 0$, то условия* а)—в) *достаточны для того, чтобы допустимая точка $\hat{x}$ была решением задачи* (1).

3. *Для того, чтобы $\hat{\lambda}_0 \neq 0$, достаточно, чтобы нашлась точка $\bar{x} \in A$, для которой справедливо условие Слейтера $f_i(\bar{x}) < 0$, $i = 1, \ldots, m$.*

Таким образом, если выполнено условие Слейтера, то можно считать $\hat{\lambda}_0 = 1$.

Соотношение а) отражает мысль Лагранжа в наиболее завершенной форме: если $\hat{x}$ доставляет минимум в задаче с ограничениями (типа неравенств), то эта же точка доставляет минимум функции Лагранжа (в задаче без тех ограничений, которые включены в функцию Лагранжа). Соотношения б) и в) характерны для задач с неравенствами (подробнее см. § 3.2).

Доказательство теоремы Куна—Таккера опирается на одну из важнейших теорем выпуклого анализа—*теорему отделимости*. Правда, здесь нам достаточно ее простейшего конечномерного варианта, который мы сейчас сформулируем, а докажем в следующем пункте.

Конечномерная теорема отделимости. *Пусть C—выпуклое подмножество в $\mathbf{R}^N$, не содержащее точки 0. Тогда найдутся такие числа $a_1, \ldots, a_N$, что для любого $x = (x_1, \ldots, x_N) \in C$ выполняется неравенство*

$$\sum_{i=1}^{N} a_i x_i \geqslant 0 \tag{3}$$

(другими словами, гиперплоскость $\sum\limits_{i=1}^{N} a_i x_i = 0$, проходящая через 0, делит $\mathbf{R}^N$ на две части, в одной из которых целиком лежит C).

Доказательство теоремы Куна—Таккера. Пусть $\hat{x}$—решение задачи. Не ограничивая себя в общности, можно считать, что $f_0(\hat{x}) = 0$,—иначе введем новую функцию $\widetilde{f_0(x)}=f_0(x)-f_0(\hat{x})$. Положим

$$C=\{\mu\in\mathrm{R}^{m+1},\ \mu=(\mu_0,\ \ldots,\mu_m)\mid \exists x\in A:\ f_0(x)<\mu_0,\\ f_i(x)\leqslant\mu_i,\ i\geqslant 1\}. \quad (4)$$

Дальнейшую часть доказательства разбиваем на несколько этапов.

А) *Множество C непусто и выпукло.* Действительно, любой вектор $\mu\in\mathrm{R}^{m+1}$ с положительными компонентами принадлежит C, ибо в (4) можно положить $x=\hat{x}$. Следовательно, множество C непусто. Докажем его выпуклость. Пусть $\mu=(\mu_0,\ \ldots,\mu_m)$ и $\mu'=(\mu_0',\ \ldots,\ \mu_m')\in C$, $0\leqslant\alpha\leqslant 1$, x и x' — такие элементы из A, что (в соответствии с (4)) $f_0(x)<\mu_0$, $f_0(x')<\mu_0'$, $f_i(x)\leqslant\mu_i$, $f_i(x')\leqslant\mu_i'$, $i\geqslant 1$. Положим $x_\alpha=\alpha x+(1-\alpha)x'$. Тогда $x_\alpha\in A$, поскольку A выпукло, а ввиду выпуклости функций f_i

$$f_i(x_\alpha)=f_i(\alpha x+(1-\alpha)x')\leqslant\\ \leqslant\alpha f_i(x)+(1-\alpha)f_i(x')\begin{cases}<\alpha\mu_0+(1-\alpha)\mu_0', & i=0,\\ \leqslant\alpha\mu_i+(1-\alpha)\mu_i', & i\geqslant 1,\end{cases}$$

т. е. точка $\alpha\mu+(1-\alpha)\mu'\in C$.

Б) *Точка $0\in\mathrm{R}^{m+1}$ не принадлежит C.* Действительно, если бы точка 0 принадлежала C, то ввиду определения (4) отсюда следовало бы, что имеется элемент $\tilde{x}\in A$, для которого выполняются неравенства: $f_0(\tilde{x})<0$, $f_i(\tilde{x})\leqslant 0$, $i\geqslant 1$. Но из этих неравенств следует, что $\hat{x}$ не есть решение задачи. Значит, $0\notin C$.

Поскольку C выпукло и $0\notin C$, можно применить теорему отделимости, согласно которой найдутся такие $\hat{\lambda}_0,\ \ldots,\hat{\lambda}_m$, что

$$\sum_{i=0}^{m}\hat{\lambda}_i\mu_i\geqslant 0 \quad (5)$$

для любого $\mu\in C$.

В) *Множители $\hat{\lambda}_i$, $i\geqslant 0$ в (5) неотрицательны.* Действительно, в п. А) мы говорили уже, что любой вектор с положительными компонентами принадлежит C, в частности, C принадлежит вектор $(\varepsilon,\ \ldots,\varepsilon,\ 1,\ \varepsilon,\ \ldots,\varepsilon)$, где $\varepsilon>0$ и 1 стоит на i_0-м месте, $i_0\geqslant 0$. Подставив эту точку в (5), получаем, что $\hat{\lambda}_{i_0}\geqslant-\varepsilon\sum\limits_{i\neq i_0}\hat{\lambda}_i$, откуда в силу произвольности $\varepsilon>0$ следует, что $\hat{\lambda}_{i_0}\geqslant 0$.

Г) *Множители* $\hat{\lambda}_i$, $i \geqslant 1$, *удовлетворяют условиям дополняющей нежесткости.* Действительно, если $f_{j_0}(\hat{x})=0$, $j_0 \geqslant 1$, то равенство $\hat{\lambda}_{j_0} f_{j_0}(\hat{x})=0$ тривиально. Пусть $f_{j_0}(\hat{x})<0$. Тогда точка $(\delta, 0, \ldots, 0, f_{j_0}(\hat{x}), 0, \ldots, 0)$, где число $f_{j_0}(\hat{x})$ стоит на j_0-м месте, а $\delta>0$, принадлежит C — достаточно взять точку $\hat{x}$ в качестве точки x в (4).

Подставив эту точку в (5), получим, что $\hat{\lambda}_{j_0} f_{j_0}(\hat{x}) \geqslant -\hat{\lambda}_0 \delta$, откуда из-за произвольности $\delta>0$ получается неравенство $\hat{\lambda}_{j_0} \leqslant 0$. Но было доказано в п. В), что $\hat{\lambda}_{j_0} \geqslant 0$, значит, $\hat{\lambda}_{j_0}=0$ и $\hat{\lambda}_{j_0} f_{j_0}(\hat{x})=0$.

Д) *В точке* $\hat{x}$ *выполнен принцип минимума.* Действительно, пусть $x \in A$. Тогда точка $(f_0(x)+\delta, f_1(x), \ldots, f_m(x))$ принадлежит C для любого $\delta>0$ (см. определение (4)). Поэтому с учетом (5) $\hat{\lambda}_0 f_0(x)+\sum_{i=1}^{m} \hat{\lambda}_i f_i(x) \geqslant -\hat{\lambda}_0 \delta$, а отсюда (в силу произвольности $\delta>0$) следует неравенство $\mathscr{L}(x, \hat{\lambda}, \hat{\lambda}_0) \geqslant 0$.

Теперь же, если учесть равенство $f_0(\hat{x})=0$ и условия дополняющей нежесткости, получаем для любого $x \in A$

$$\mathscr{L}(x, \hat{\lambda}, \hat{\lambda}_0) \geqslant 0=\sum_{i=0}^{m} \hat{\lambda}_i f_i(\hat{x})=\mathscr{L}(\hat{x}, \hat{\lambda}, \hat{\lambda}_0).$$

Утверждение 1 теоремы доказано.

Е) Докажем *утверждение* 2. Если допустить, что $\hat{\lambda}_0 \neq 0$, то можно считать $\hat{\lambda}_0=1$ (ибо множители Лагранжа, удовлетворяющие соотношениям а) — в) теоремы, сохраняют свои свойства при умножении на любой положительный множитель). Но тогда для любого допустимого x $(x \in A,\ f_i(x) \leqslant 0,\ i \geqslant 1)$ получаем

$$f_0(x) \overset{\text{б)}}{\geqslant} f_0(x)+\sum_{i=1}^{m} \hat{\lambda}_i f_i(x)=\mathscr{L}(x, \hat{\lambda}, 1) \overset{\text{а)}}{\geqslant} \mathscr{L}(\hat{x}, \hat{\lambda}, 1)=$$

$$=f_0(\hat{x})+\sum_{i=1}^{m} \hat{\lambda}_i f_i(\hat{x}) \overset{\text{в)}}{=} f_0(\hat{x}),$$

т. е. $\hat{x}$ является решением задачи.

Ж) Докажем *утверждение* 3. Пусть выполнено условие Слейтера, (т. е. для некоторого $\bar{x} \in A$ имеют место неравенства $f_i(\bar{x})<0$, $i \geqslant 1$), но при этом в утверждении 1

$\hat{\lambda}_0 = 0$. Тогда сразу же получается противоречие (в выкладке используется, что не все $\hat{\lambda}_i = 0$, $i \geqslant 1$):

$$\mathscr{L}(\bar{x}, \hat{\lambda}, 0) = \sum_{i=0}^{m} \hat{\lambda}_i f_i(\bar{x}) < 0 = \mathscr{L}(\hat{x}, \hat{\lambda}, 0),$$

в то время как вследствие а) $\mathscr{L}(\bar{x}, \hat{\lambda}, 0) \geqslant \mathscr{L}(\hat{x}, \hat{\lambda}, 0)$. ■

Приведем еще один вариант теоремы Куна—Таккера. При выполнении условия Слейтера $\hat{\lambda}_0 > 0$ и поскольку множители Лагранжа определены с точностью до положительного множителя, можно считать, что $\hat{\lambda}_0 = 1$. Теперь функция Лагранжа

$$\mathscr{L}(x, \lambda, 1) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$$

определена на множестве

$$A \times \mathbf{R}_+^m = \{(x, \lambda) \mid x = (x_1, \ldots, x_m) \in A; \lambda = \\ = (\lambda_1, \ldots, \lambda_m), \lambda_i \geqslant 0\}$$

и соотношения а)—в) равносильны тому, что $(\hat{x}, \hat{\lambda})$ является ее *седловой точкой*, т. е.

$$\min_{x \in A} \mathscr{L}(x, \hat{\lambda}, 1) = \mathscr{L}(\hat{x}, \hat{\lambda}, 1) = \max_{\lambda \geqslant 0} \mathscr{L}(\hat{x}, \lambda, 1). \qquad (6)$$

Действительно, левое равенство (6) совпадает с а), а правое является следствием в) и б):

$$\mathscr{L}(\hat{x}, \hat{\lambda}, 1) = f_0(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i f_i(\hat{x}) = f_0(\hat{x}) \geqslant \\ \geqslant f_0(\hat{x}) + \sum_{i=1}^{m} \lambda_i f_i(\hat{x}) = \mathscr{L}(\hat{x}, \lambda, 1).$$

В §§ 2.6, 3.1, 3.3, 4.3 мы расскажем о других теоремах выпуклого анализа и выпуклого программирования.

Замечание. В «выпуклом анализе» оказывается удобным рассматривать функции со значениями в расширенной числовой прямой $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$, т. е. удобно разрешить функциям принимать бесконечные значения $-\infty$, $+\infty$. На расширенную числовую прямую распространяются с некоторыми ограничениями правила арифметики (об этом подробнее сказано в примечании в п. 2.6.1). Легко убедиться в том, что приведенное доказательство теоремы Куна—Таккера остается применимым без изменений и к таким функциям.

1.3.4. Доказательство конечномерной теоремы отделимости. Формулировка теоремы была приведена выше. Напомним, что C — выпуклое множество в $\mathbf{R}^N$ и $0 \notin C$.

А) Пусть $\operatorname{lin} C$ — *линейная оболочка* множества C, т. е. наименьшее линейное подпространство, содержащее C. Возможно одно из двух: либо $\operatorname{lin} C \neq \mathbf{R}^N$, либо $\operatorname{lin} C = \mathbf{R}^N$. В первом случае $\operatorname{lin} C$ — собственное подпространство в $\mathbf{R}^N$, и потому существует содержащая его гиперплоскость $\sum_{i=1}^{N} a_i x_i = 0$, проходящая через нуль. Она и является искомой.

Б) Если $\operatorname{lin} C = \mathbf{R}^N$, то из векторов, принадлежащих C, мы можем выбрать N линейно независимых и потому образующих базис в $\mathbf{R}^N$. Обозначим их $\tilde{e}_1, \ldots, \tilde{e}_N$; $\tilde{e}_i \in C$, $i = 1, \ldots, N$.

Рассмотрим в $\mathbf{R}^N$ два выпуклых конуса: «отрицательный ортант» $\mathcal{K}_1 = \left\{ x \in \mathbf{R}^N \,\middle|\, x = \sum_{i=1}^{N} \beta_i \tilde{e}_i,\ \beta_i < 0,\ i = 1, \ldots, N \right\}$ и коническую оболочку множества C:

$$\mathcal{K}_2 = \left\{ x \in \mathbf{R}^N \,\middle|\, x = \sum_{i=1}^{s} \alpha_i \xi_i,\ \xi_i \in C,\ \alpha_i \geqslant 0,\ i = 1, \ldots, s,\ s \text{ любое} \right\}.$$

Эти конусы не пересекаются. Действительно, если бы вектор $\tilde{x} = -\sum_{i=1}^{N} \tilde{\gamma}_i \tilde{e}_i$, $\tilde{\gamma}_i > 0$, принадлежал $\mathcal{K}_2$, то нашлись бы такие $s \in \mathbf{N}$, $\tilde{\alpha}_i \geqslant 0$ и $\tilde{\xi}_i \in C$, что $\tilde{x} = \sum_{i=1}^{s} \tilde{\alpha}_i \tilde{\xi}_i$. Но тогда точка 0 оказалась бы принадлежащей C, ибо эта точка могла бы быть представлена в виде выпуклой комбинации

$$0 = \sum_{i=1}^{s} \tilde{\alpha}_i \tilde{\xi}_i - \tilde{x} = \frac{\sum_{i=1}^{s} \tilde{\alpha}_i \tilde{\xi}_i + \sum_{i=1}^{N} \tilde{\gamma}_i \tilde{e}_i}{\sum_{i=1}^{s} \tilde{\alpha}_i + \sum_{i=1}^{N} \tilde{\gamma}_i}$$

точек из C.

В) Поскольку $\mathcal{K}_1$ открыто, из доказанного в Б) следует, что ни одна из точек $\mathcal{K}_1$ не может принадлежать замыканию $\overline{\mathcal{K}}_2$ множества $\mathcal{K}_2$. (Отметим, что $\overline{\mathcal{K}}_2$ замкнуто и выпукло. Почему?) Возьмем произвольную точку из $\mathcal{K}_1$, например $x_0 = -\sum_{i=1}^{N} \tilde{e}_i$, и найдем ближайшую к ней точку $\xi_0 \in \overline{\mathcal{K}}_2$. Такая точка существует, а именно, это та из точек компактного множества $\overline{\mathcal{K}}_2 \cap B(x_0, |x_0|)$, в которой непрерывная функция $f(x) = |x - x_0|$ достигает своего минимума.

Г) Проведем через ξ_0 гиперплоскость H, перпендикулярную $x_0-\xi_0$, и покажем, что она искомая, т. е. что $0\in H$ и множество C целиком лежит в одном из двух замкнутых полупространств, ограниченных этой гиперплоскостью. Мы докажем даже больше, а именно, если $\mathring{H}$ — внутренность того из полупространств, которое содержит точку x_0, то $\mathring{H}\cap\overline{\mathscr{K}}_2=\varnothing$. Поскольку множество $C\subset\overline{\mathscr{K}}_2$, оно целиком содержится в замкнутом полупространстве, дополнительном к $\mathring{H}$.

Предположим противное, и пусть $\xi_1\in\mathring{H}\cap\overline{\mathscr{K}}_2$. Тогда угол $\widehat{x_0\xi_0\xi_1}$ острый. Кроме того, $[\xi_0, \xi_1]\in\overline{\mathscr{K}}_2$, поскольку $\overline{\mathscr{K}}_2$ выпукло. Опустим из x_0 на прямую (ξ_0, ξ_1) перпендикуляр (x_0, ξ_2), $\xi_2\in(\xi_0, \xi_1)$ и покажем, что ξ_0 не является ближайшей к x_0 точкой $\overline{\mathscr{K}}_2$. Действительно, точки ξ_0, ξ_1 и ξ_2 лежат на одной прямой и $\xi_2\in\mathring{H}$ (почему?). Если $\xi_2\in[\xi_0, \xi_1]$, то $\xi_2\in\overline{\mathscr{K}}_2$ и $|x_0-\xi_2|<|x_0-\xi_0|$ (перпендикуляр меньше наклонной). Если же ξ_1 лежит между ξ_0 и ξ_2, то $|x_0-\xi_1|<|x_0-\xi_0|$ (из двух наклонных та меньше, основание которой лежит ближе к основанию перпендикуляра)

С другой стороны, H проходит через точку 0, ибо иначе луч, исходящий из 0 в направлении точки ξ_0 и целиком лежащий в $\overline{\mathscr{K}}_2$ (почему?), обязательно имел бы общие точки с $\mathring{H}$.

§ 1.4. Простейшая задача классического вариационного исчисления и ее обобщения

1.4.1. Уравнение Эйлера. Вскоре после работы И. Бернулли о брахистохроне стали появляться (и решаться) многие задачи того же типа. И. Бернулли поставил перед своим учеником Л. Эйлером проблему найти общий путь их решения.

В 1744 г. вышел труд Эйлера «Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latissimo sensu accepti», «Метод нахождения кривых линий, обладающих свойствами максимума или минимума, или решение изопериметрической задачи, взятой в самом широком смысле», в котором были заложены теоретические основы нового раздела математического анализа. В частности, аппроксимируя кривые ломаными, Эйлер вывел дифференциальное уравнение второго порядка, которому должны были удовлетворять экстремали. Впоследствии Лагранж назвал его *уравнением Эйлера*. В 1759 г. появляется первая работа Лагранжа и с ней новые методы исследования. Лагранж *«варьирует»* кривую, подозреваемую на экстремум, *выделяет из приращений функционалов главные линейные части*, которые называет *вариациями*, и пользуется тем, что в точке экстремума вариация должна обращаться в нуль. Метод Лагранжа становится впоследствии общепринятым. Этим методом и мы выведем далее уравнение Эйлера. Отметим еще, что после работ Лагранжа по предложению Эйлера весь раздел математики, к которому применялся метод Лагранжа, стали называть *вариационным исчислением*.

Перейдем к выводу уравнения Эйлера для простейшей задачи классического вариационного исчисления. В п. 1.2.6 этим именем была названа экстремальная задача

$$\mathcal{J}(x(\cdot))=\int_{t_0}^{t_1} L(t, x, \dot{x})\,dt \to \mathrm{extr}, \quad x(t_0)=x_0, \quad x(t_1)=x_1, \tag{1}$$

рассматриваемая в пространстве непрерывно дифференцируемых функций $C^1([t_0, t_1])$, $-\infty < t_0 < t_1 < \infty$. Пространство $C^1([t_0, t_1])$ является банаховым, т. е. полным нормированным, относительно нормы:

$$\|x(\cdot)\|_1=\max\left(\max_{t\in[t_0, t_1]}|x(t)|, \max_{t\in[t_0, t_1]}|\dot{x}(t)|\right).$$

Будем предполагать, что функция L (ее называют *интегрантом* или *лагранжианом* задачи) непрерывна по совокупности переменных вместе со своими частными производными L_x и $L_{\dot{x}}$.

Теорема. (Необходимое условие экстремума в простейшей задаче классического вариационного исчисления.) *Пусть функция* $\hat{x}(\cdot)\in C^1([t_0, t_1])$ *доставляет локальный экстремум в задаче* (1). *Тогда она удовлетворяет уравнению*

$$-\frac{d}{dt}L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))+L_x(t, \hat{x}(t), \dot{\hat{x}}(t))=0. \tag{2}$$

Уравнение (2) называется *уравнением Эйлера*. Допустимая функция $\hat{x}(\cdot)$, для которой оно выполнено, называется *стационарной точкой* задачи (1) или *экстремалью*. Таким образом, локальные экстремумы задачи являются экстремалями; обратное, вообще говоря, неверно.

Локальный экстремум в пространстве $C^1([t_0, t_1])$ в вариационном исчислении называют *слабым*. В соответствии с общими определениями функция $\hat{x}(\cdot)$ доставляет локальный минимум (максимум) в задаче (1) в пространстве $C^1([t_0, t_1])$, если найдется такое $\varepsilon>0$, что для всякой функции $x(\cdot)\in C^1([t_0, t_1])$, для которой $x(t_0)=x(t_1)=0$

и $\|x(\cdot)\|_1 \leqslant \varepsilon$ выполняется неравенство

$$\mathcal{J}(\hat{x}(\cdot)+x(\cdot)) \geqslant \mathcal{J}(\hat{x}(\cdot)) \qquad (\leqslant \mathcal{J}(\hat{x}(\cdot))).$$

Доказательство теоремы проведем дважды. Сначала воспользуемся рассуждением Лагранжа. Правда, при этом придется дополнительно предположить, что функция $t \mapsto L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))$ непрерывно дифференцируема. Затем докажем важную лемму Дюбуа-Реймона, из которой наша теорема вытекает и без дополнительного допущения. Обобщение конструкции Дюбуа-Реймона окажется существенным в гл. IV.

Доказательство теоремы складывается из трех этапов.

А) Определение первой вариации по Лагранжу. Пусть $x(\cdot) \in C^1([t_0, t_1])$. Рассмотрим функцию $\lambda \mapsto \varphi(\lambda) = \mathcal{J}(\hat{x}(\cdot)+\lambda x(\cdot))$. Имеем

$$\varphi(\lambda) = \int_{t_0}^{t_1} F(t, \lambda)\, dt = \int_{t_0}^{t_1} L(t, \hat{x}(t)+\lambda x(t), \dot{\hat{x}}(t)+\lambda \dot{x}(t))\, dt.$$

Допущения, наложенные на L, позволяют дифференцировать функцию F под знаком интеграла (для этого достаточно, чтобы функции $(t, \lambda) \to F(t, \lambda)$ и $(t, \lambda) \to \frac{\partial F}{\partial \lambda}(t, \lambda)$ были непрерывными [14, т. 2, с. 661], [9, т. 2, с. 107]). Дифференцируя и подставляя $\lambda = 0$, получаем

$$\varphi'(0) = \int_{t_0}^{t_1} (q(t)\, x(t) + p(t)\, \dot{x}(t))\, dt, \tag{3}$$

где

$$p(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)), \quad q(t) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t)).$$

Итак, $\lim\limits_{\lambda \to 0} (\mathcal{J}(\hat{x}(\cdot)+\lambda x(\cdot)) - \mathcal{J}(\hat{x}(\cdot))/\lambda)$ существует для любого $x(\cdot) \in C^1([t_0, t_1])$. Обозначим этот предел $\delta\mathcal{J}(\hat{x}(\cdot), x(\cdot))$. Функция $x(\cdot) \mapsto \delta\mathcal{J}(\hat{x}(\cdot), x(\cdot))$ называется *первой вариацией по Лагранжу функционала* $\mathcal{J}$.

Б) Преобразование первой вариации с помощью интегрирования по частям. Пусть $x(t_0) = x(t_1) = 0$ и функция $p(\cdot)$ непрерывно дифференцируема. Следуя Лагранжу, имеем

$$\delta\mathcal{J}(\hat{x}(\cdot),\ x(\cdot))=\int_{t_0}^{t_1}(q(t)\,x(t)+p(t)\,\dot{x}(t))\,dt=$$

$$=\int_{t_0}^{t_1}a(t)\,x(t)\,dt, \quad (4)$$

где $a(t)=-\dot{p}(t)+q(t)$ (произведение $p(t)\,x(t)$ обращается в нуль на концах промежутка интегрирования).

Нам известно, что $\hat{x}(\cdot)$ доставляет локальный экстремум, для определенности пусть это будет минимум, в задаче (1). Отсюда следует, что функция $\lambda\mapsto\varphi(\lambda)$ имеет локальный минимум в нуле. Вследствие теоремы Ферма (см. п. 1.3.1) $\varphi'(0)=\delta\mathcal{J}(\hat{x}(\cdot),\ x(\cdot))=0$. Сопоставляя это с (4) мы получаем, что для произвольной функции $x(\cdot)\in C^1([t_0,\ t_1])$ такой, что $x(t_0)=x(t_1)=0$, имеет место равенство $\int_{t_0}^{t_1}a(t)\,x(t)\,dt=0$.

В) Основная лемма классического вариационного исчисления. (Лемма Лагранжа.) *Пусть непрерывная функция* $t\mapsto a(t)$, $t\in[t_0,\ t_1]$ *обладает тем свойством, что* $\int_{t_0}^{t_1}a(t)\,x(t)\,dt=0$ *для любой непрерывно дифференцируемой функции* $x(\cdot)$, *у которой* $x(t_0)=x(t_1)=0$. *Тогда* $a(t)\equiv 0$.

Доказательство. Допустим, что, $a(\tau)\neq 0$ в некоторой точке $\tau\in[t_0,\ t_1]$. Тогда вследствие непрерывности $a(\cdot)$ найдется отрезок $\Delta=[\tau_1,\ \tau_2]\subset(t_0,\ t_1)$, на котором $a(\cdot)$ не обращается в нуль. Пусть для определенности $a(t)\geqslant m>0$, $t\in\Delta$. Построим функцию

$$\tilde{x}(t)=\begin{cases}(t-\tau_1)^2(t-\tau_2)^2, & t\in\Delta,\\ 0, & t\notin\Delta.\end{cases}$$

Нетрудно проверить, что $\tilde{x}(\cdot)\in C^1([t_0,\ t_1])$; кроме того, $\tilde{x}(t_0)=\tilde{x}(t_1)=0$. По условию леммы $\int_{t_0}^{t_1}a(t)\,\tilde{x}(t)\,dt=0$. С другой стороны, по теореме о среднем [14, т. 2, с. 115], [9, т. 1, с. 344] $\int\limits_{t_0}^{t_1} a(t)\,\bar{x}(t)\,dt = a(\xi)\int\limits_{t_0}^{t_1}\bar{x}(t)\,dt > 0$, и это противоречие доказывает лемму. ■

Сопоставляя полученное в А) и В), убеждаемся, что $-\dot{p}(t)+q(t)\equiv 0$, а это и требовалось доказать.

Г) Лемма Дюбуа-Реймона. *Пусть на* $[t_0, t_1]$ *функции* $a(\cdot)$ *и* $b(\cdot)$ *непрерывны, и пусть*

$$\int\limits_{t_0}^{t_1}(a(t)\,\dot{x}(t)+b(t)\,x(t))\,dt=0 \tag{5}$$

для любой $x(\cdot)\in C^1([t_0, t_1])$ *такой, что* $x(t_0)=x(t_1)=0$. *Тогда функция* $a(\cdot)$ *непрерывно дифференцируема и* $da(t)/dt=b(t)$.

Доказательство. Интегрируя в (5) второе слагаемое в подынтегральном выражении по частям, получаем

$$0=\int\limits_{t_0}^{t_1}\left(a(t)-\int\limits_{t_0}^{t} b(s)\,ds\right)\dot{x}(t)\,dt+\int\limits_{t_0}^{t} b(s)\,ds\,x(t)\Bigg|_{t_0}^{t_1}=$$

$$=\int\limits_{t_0}^{t_1}\left(a(t)-\int\limits_{t_0}^{t} b(s)\,ds\right)\dot{x}(t)\,dt. \tag{6}$$

Докажем, что $\varphi(t)=a(t)-\int\limits_{t_0}^{t} b(s)\,ds=\text{const}$. Если это не так, то найдутся такие τ_1 и τ_2, что $\varphi(\tau_1)\neq\varphi(\tau_2)$. Без ограничения общности можно считать, что τ_1 и τ_2 — внутренние точки отрезка $[t_0, t_1]$ и $\varphi(\tau_1)>\varphi(\tau_2)$. Выберем $\delta>0$ так, чтобы $[\tau_i-\delta,\ \tau_i+\delta]\subset[t_0, t_1]$, $i=1, 2$, и чтобы $\varphi(s+\tau_1)-\varphi(s+\tau_2)\geqslant\alpha>0$ при $|s|\leqslant\delta$.

Теперь возьмем функцию $x(\cdot)$, производная которой имеет специальный вид

$$\dot{x}(t)=\begin{cases} -(t-\tau_1+\delta)(t-\tau_1-\delta), & t\in[\tau_1-\delta,\ \tau_1+\delta],\\ +(t-\tau_2+\delta)(t-\tau_2-\delta), & t\in[\tau_2-\delta,\ \tau_2+\delta],\\ 0 \text{ для остальных } t. \end{cases} \tag{7}$$

Если положить $x(t_0)=0$, то в силу (7) $x(t_1)=\int\limits_{t_0}^{t_1}\dot{x}(t)\,dt=0$, так что для такой $x(\cdot)$ должно иметь место (6), но

$$\int\limits_{t_0}^{t_1}\left[a(t)-\int\limits_{t_0}^{t} b(s)\,ds\right]\dot{x}(t)\,dt=\int\limits_{\tau_1-\delta}^{\tau_1+\delta}\varphi(t)\,\dot{x}(t)\,dt+\int\limits_{\tau_2-\delta}^{\tau_2+\delta}\varphi(t)\,\dot{x}(t)\,dt=$$

$$=\int\limits_{-\delta}^{\delta}(\varphi(\tau_1+s)-\varphi(\tau_2+s))(\delta-s)(s+\delta)\,ds\geqslant\alpha\int\limits_{-\delta}^{\delta}(\delta-s)(s+\delta)\,ds>0.$$

Противоречие показывает, что $\varphi(t)=a(t)-\int\limits_{t_0}^{t} b(s)\,ds=\mathrm{const}$, а тогда $a(\cdot)$ дифференцируема и $da/dt=b$. ■

Применив лемму Дюбуа-Реймона к первой вариации (3), получаем, что $p(\cdot)\in C^1([t_0, t_1])$ и $\dot p(t)=q(t)$.

В только что проведенных рассуждениях заключен зародыш так называемого *метода вариаций*, с помощью которого выводятся различные необходимые условия экстремума. Суть его состоит в следующем. Пусть $\hat x$ — точка, подозреваемая на минимум в задаче $f(x)\to\inf$, $x\in C$. Тогда можно попытаться построить непрерывное отображение $\lambda\mapsto x(\lambda)$, $\lambda\in\mathbf{R}_+$ так, чтобы $x(0)=\hat x$ и $x(\lambda)\in C$, $0\leqslant\lambda\leqslant\lambda_0$. Эту кривую естественно назвать вариацией аргумента. Положим $\varphi(\lambda)=f(x(\lambda))$. Допустим, что функция φ оказывается дифференцируемой справа по λ. Если $\hat x$ — действительно точка минимума, то должно выполняться неравенство $\varphi'(+0)=\lim\limits_{\lambda\downarrow 0}(\varphi(\lambda)-\varphi(0))/\lambda\geqslant 0$, поскольку, очевидно, $\varphi(\lambda)\geqslant\varphi(0)=f(\hat x)$. Если удается построить достаточно массивное множество вариаций аргумента, то полученный набор неравенств $\varphi'(+0)\geqslant 0$, относящихся ко всем вариациям, образует некоторое необходимое условие минимума. При выводе уравнения Эйлера использовались «вариации по направлениям»: $x(\lambda)=\hat x+\lambda x$. При выводе необходимого условия Вейерштрасса и принципа максимума будут использоваться вариации другого вида — так называемые «игольчатые» (см. пп. 1.4.4 и 1.5.4).

В заключение отметим следующие два первых интеграла уравнения Эйлера:

1) Если лагранжиан L не зависит от x, то уравнение Эйлера имеет очевидный интеграл

$$p(t)=L_{\dot x}(t, \dot{\hat x}(t))=\mathrm{const}.$$

Этот интеграл называют *интегралом импульса*.

2) Если лагранжиан L не зависит от t, то уравнение Эйлера имеет первый интеграл

$$H(t)=L_{\dot x}(\hat x(t), \dot{\hat x}(t))\,\dot{\hat x}(t)-L(\hat x(t), \dot{\hat x}(t))=\mathrm{const}.$$

Он называется *интегралом энергии* (оба названия восходят к классической механике; мы еще будем иметь возможность к этому вернуться; см. п.п. 4.4.6 и 4.4.7). Для доказательства вычислим производную:

$$\frac{dH}{dt}=L_{\dot{x}}(\hat{x}(t),\ \dot{\hat{x}}(t))\ddot{\hat{x}}(t)+\frac{d}{dt}L_{\dot{x}}(\hat{x}(t),\ \dot{\hat{x}}(t))\dot{\hat{x}}(t)-$$
$$-L_x(\hat{x}(t),\ \dot{\hat{x}}(t))\dot{\hat{x}}(t)-L_{\dot{x}}(\hat{x}(t),\ \dot{\hat{x}}(t))\ddot{\hat{x}}(t)=$$
$$=\left(\frac{d}{dt}L_{\dot{x}}(\hat{x}(t),\ \dot{\hat{x}}(t))-L_x(\hat{x}(t),\ \dot{\hat{x}}(t))\right)\dot{\hat{x}}(t)=0$$

в силу уравнений Эйлера. Следовательно, $H(t)=\mathrm{const}$.

Замечание. По ходу дела мы использовали дополнительное предположение о существовании второй производной $\ddot{\hat{x}}(t)$. Это предположение можно ослабить, но совсем без него обойтись нельзя.

Отметим еще, что уравнение Эйлера — это дифференциальное уравнение второго порядка. Его общий интеграл зависит (вообще говоря) от двух произвольных постоянных. Их значения мы находим из граничных условий.

1.4.2. Необходимые условия в задаче Больца. Условия трансверсальности. Простейшая задача, рассмотренная в предыдущем пункте, — это задача с ограничениями: граничные условия $x(t_0)=x_0$, $x(t_1)=x_1$ образуют два ограничения типа равенств. Следующая задача — так называемая задача Больца — является задачей без ограничений. Пусть L удовлетворяет тем же условиям, что и в п. 1.4.1; $l=l(x_0,\ x_1)$ — непрерывно дифференцируемая функция двух переменных. Рассмотрим задачу на экстремум в пространстве $C^1([t_0,\ t_1])$:

$$\mathscr{B}(x(\cdot))=\int_{t_0}^{t_1}L(t,\ x,\ \dot{x})\,dx+l(x(t_0),\ x(t_1))\to\mathrm{extr}. \quad (1)$$

Ее и называют *задачей Больца.*

Теорема (необходимое условие экстремума в задаче Больца). *Пусть функция* $\hat{x}(\cdot)\in C^1([t_0,\ t_1])$ *и доставляет локальный минимум в задаче* (1). *Тогда выполняется уравнение Эйлера*

$$-\frac{d}{dt}L_{\dot{x}}(t,\ \hat{x}(t),\ \dot{\hat{x}}(t))+L_x(t,\ \hat{x}(t),\ \dot{\hat{x}}(t))=0 \quad (2)$$

и условия трансверсальности:

$$L_{\dot{x}}(t_0, \hat{x}(t_0), \dot{\hat{x}}(t_0)) = \frac{\partial l(\hat{x}(t_0), \hat{x}(t_1))}{\partial x_0},$$

$$L_{\dot{x}}(t_1, \hat{x}(t_1), \dot{\hat{x}}(t_1)) = -\frac{\partial l(\hat{x}(t_0), \hat{x}(t_1))}{\partial x_1}. \qquad (3)$$

Доказательство. Действуя в точности так, как на первом этапе доказательства теоремы предыдущего пункта, получим такое выражение для первой вариации функционала $\mathscr{B}$:

$$\delta\mathscr{B}(\hat{x}(\cdot), x(\cdot)) =$$
$$= \int_{t_0}^{t_1} (q(t)x(t) + p(t)\dot{x}(t))\,dt + \alpha_0 x(t_0) + \alpha_1 x(t_1), \qquad (4)$$

где $p(\cdot)$ и $q(\cdot)$ определены соотношениями (3) предыдущего пункта, а $\alpha_i = \frac{\partial l(\hat{x}(t_0), \hat{x}(t_1))}{\partial x_i}$, $i=0, 1$. Так как $\hat{x}(\cdot)$ — точка локального минимума, $\delta\mathscr{B}(\hat{x}(\cdot), x(\cdot)) = 0$ для любой точки $x(\cdot) \in C^1([t_0, t_1])$. В частности, $\delta\mathscr{B}(\hat{x}(\cdot), x(\cdot)) = 0$ для любой $x(\cdot) \in C^1([t_0, t_1])$, у которой $x(t_0) = x(t_1) = 0$. По лемме Дюбуа — Реймона функция $p(\cdot)$ непрерывно дифференцируема и при этом $\frac{dp(t)}{dt} = q(t)$. Интегрируя (4) по частям и используя то, что $\dot{p}(t) = q(t)$, приходим к следующему выражению для первой вариации по Лагранжу функционала Больца:

$$0 = \delta\mathscr{B}(\hat{x}(\cdot), x(\cdot)) =$$
$$= \int_{t_0}^{t_1} (q(t) - \dot{p}(t))\,x(t)\,dt + (\alpha_0 - p(t_0))\,x(t_0) +$$
$$+ (\alpha_1 + p(t_1))\,x(t_1) = (\alpha_0 - p(t_0))\,x(t_0) + (\alpha_1 + p(t_1))\,x(t_1). \qquad (5)$$

Положив в (5) последовательно $x(t) = (t - t_0)$, а затем $t - t_1$, получаем, что $\alpha_0 = p(t_0)$, $-\alpha_1 = p(t_1)$. ■

Итак, мы снова получили уравнение второго порядка и два краевых условия — условия трансверсальности.

Выше была рассмотрена «одномерная» задача Больца. Совершенно аналогично ставится векторная задача:

$$\mathscr{B}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x_1, \ldots, x_n, \dot{x}_1, \ldots, \dot{x}_n)\,dt +$$
$$+ l(x_1(t_0), \ldots, x_n(t_0), x_1(t_1), \ldots, x_n(t_1)) \to \text{extr}, \qquad (1')$$

где

$$L\colon \mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^n\to\mathbf{R},\qquad l\colon \mathbf{R}^n\times\mathbf{R}^n\to\mathbf{R}.$$

Необходимые условия экстремума здесь также имеют вид (2), (3), только соответствующие уравнения надо понимать векторно:

$$-\frac{d}{dt}L_{\dot{x}_j}(t_1,\hat{x}_1(t),\ldots,\hat{x}_n(t),\dot{\hat{x}}_1(t),\ldots,\dot{\hat{x}}_n(t))+$$
$$+L_{x_j}(t,\hat{x}_1(t),\ldots,\hat{x}_n(t),\dot{\hat{x}}_1(t),\ldots,\dot{\hat{x}}_n(t))=0,\quad (2')$$

$$L_{\dot{x}_j}(t_i,\hat{x}_1(t_i),\ldots,\hat{x}_n(t_i),\dot{\hat{x}}_1(t_i),\ldots,\dot{\hat{x}}_n(t_i))=$$
$$=(-1)^i\frac{\partial l(\hat{x}_1(t_0),\ \ ,\hat{x}_n(t_0),\hat{x}_1(t_1),\ \ ,\hat{x}_n(t_1))}{\partial x_{ij}},\quad (3')$$
$$i=0,1;\quad j=1,\ldots,n.$$

1.4.3. Расширения простейшей задачи. Простейшую задачу классического вариационного исчисления в п. 1.4.1 мы рассматривали в пространстве $C^1([t_0,t_1])$. В этом — определенная дань традиции. Пространство $C^1([t_0,t_1])$, конечно, удобно, но далеко не для каждой задачи оно является естественным. В частности, при тех предположениях относительно интегранта (непрерывность L, L_x, $L_{\dot{x}}$ по совокупности переменных $(t,x,\dot{x})$), существование решения в пространстве $C^1([t_0,t_1])$ не всегда можно гарантировать. Приведем один из простейших примеров.

Пример Гильберта. Рассмотрим задачу

$$\mathcal{J}(x(\cdot))=\int_0^1 t^{2/3}\dot{x}^2dt\to\inf,\quad x(0)=0,\quad x(1)=1.$$

Здесь уравнение Эйлера имеет интеграл импульса (см. п. 1 4.1) $t^{2/3}\dot{x}=C$, откуда следует, что единственной допустимой экстремалью является $\hat{x}(t)=t^{1/3}$. Эта экстремаль не принадлежит пространству $C^1([t_0,t_1])$ (почему?). Вместе с тем она доставляет абсолютный минимум в задаче. Действительно, пусть $x(\cdot)$ — любая абсолютно-непрерывная функция[1]), для которой интеграл $\mathcal{J}(x(\cdot))$ конечен

[1]) Об абсолютно непрерывных функциях говорится в п. 2.1.8. Читатели, не знакомые с этим понятием, могут считать, что $x(\ )$ непрерывно дифференцируема в полуинтервале $(0,1]$, интегрируема в несобственном смысле на отрезке $[0,1]$, имеет конечный интеграл $\mathcal{J}(x(\ ))$ и $x(0)=x(1)=0$.

и $x(0)=x(1)=0$. Тогда

$$\mathcal{J}(\hat{x}(\cdot)+x(\cdot))=\int_0^1 t^{2/3}(\dot{\hat{x}}^2(t)+2\dot{\hat{x}}(t)\dot{x}(t)+\dot{x}^2(t))\,dt=$$

$$=\mathcal{J}(\hat{x}(\cdot))+\frac{2}{3}\int_0^1 \dot{x}(t)\,dt+\mathcal{J}(x(\cdot))=$$

$$=\mathcal{J}(\hat{x}(\cdot))+\mathcal{J}(x(\cdot))\geqslant\mathcal{J}(\hat{x}(\cdot)).$$

Итак, решение задачи существует, но не принадлежит пространству $C^1([t_0,\ t_1])$.

Среди знаменитых «проблем Гильберта» некоторые посвящены вариационному исчислению. В частности, в двадцатой проблеме речь идет о существовании решения. Гильберт[1]) пишет так: «Я убежден, что эти доказательства существования можно будет провести с помощью некоторого общего основного положения, на которое указывает принцип Дирихле и который, вероятно, приблизит нас к вопросу о том, не допускает ли решение каждая регулярная вариационная задача..., если в случае необходимости самому понятию решения придать расширенное толкование». Оптимизм Гильберта подкрепляется огромным опытом классического анализа. Здесь мы вкратце обсудим проблему «расширенного толкования» решения на примере простейшей вариационной задачи.

Одна из конструкций расширения состоит в том, что к первоначальному классу допустимых элементов добавляются новые и функционал при этом доопределяется на этих новых элементах. В этом направлении мы сделаем скромный шаг — осуществим расширение простейшей задачи на класс кусочно-гладких функций.

Напомним, что кусочно-гладкой функцией на $[t_0,\ t_1]$ называется функция $x(\cdot)$, обладающая тем свойством, что сама она непрерывна, а ее производная кусочно-непрерывна, т. е. непрерывна всюду на $[t_0,\ t_1]$, за исключением конечного числа точек $t_0<\tau_1<\ldots<\tau_m<t_1$, и при этом в точках τ_i производная $\dot{x}(\cdot)$ имеет разрывы первого рода. Например, функция

$$x(t)=\begin{cases} 0, & -1\leqslant t\leqslant 0,\\ t^2\sin\frac{1}{t}, & 0<t,\end{cases}$$

[1]) Проблемы Гильберта/Под ред. П. С. Александрова.—М.: Наука, 1969, с. 55.

имеет производную, которая не является кусочно-непрерывной. Совокупность всех кусочно-гладких функций на $[t_0, t_1]$ обозначим $KC^1([t_0, t_1])$

Интегральный функционал

$$\mathcal{J}(x(\cdot))=\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt \tag{1}$$

(ср. п. 1.4.1) естественно продолжается на элементы $x(\cdot)\in KC^1([t_0, t_1])$, поскольку для таких $x(\cdot)$ подынтегральная функция кусочно-непрерывна и интеграл существует.

Рассмотрим теперь простейшую задачу

$$\mathcal{J}(x(\cdot))\to \mathrm{extr}, \quad x(t_0)=x_0, \quad x(t_1)=x_1 \tag{2}$$

в пространстве $KC^1([t_0, t_1])$. В конце этого пункта будет приведен пример функции L, удовлетворяющей стандартным условиям, для которой задача вида (2) в пространстве $C^1([t_0, t_1])$ решения не имеет, а в $KC^1([t_0, t_1])$ — имеет, так что проведенное расширение задачи оказывается разумным (хотя в примере Гильберта минимум не достигается и в KC^1).

Несколько аккуратней надо быть с понятием локального решения. Если в окрестность элемента $\hat{x}(\cdot)\in KC^1([t_0, t_1])$ мы будем по-прежнему зачислять такие $x(\cdot)$, что малы как сами разности $x(t)-\hat{x}(t)$, так и их производные (ср. определение слабого экстремума в п. 1.4.1), то вновь добавленные элементы будут располагаться далеко от старых. Действительно, если производная $\dot{\hat{x}}(\cdot)$ имеет в некоторой точке скачок величины δ, то ни одна из функций $x(\cdot)\in C^1([t_0, t_1])$ не может удовлетворять неравенству $|\dot{x}(t)-\dot{\hat{x}}(t)|<\delta/2$ для всех t (для которых $\dot{\hat{x}}(t)$ существует). Это, вообще говоря, неудобно: обычно хотят, чтобы старое пространство лежало в расширенном всюду плотно. Поэтому близость в $KC^1([t_0, t_1])$ мы будем определять, сравнивая только сами функции, но не их производные. Это приводит к замене понятия слабого экстремума (п. 1.4.1) понятием сильного экстремума.

Определение. *Функция* $\hat{\hat{x}}(\cdot)\in KC^1([t_0, t_1])$ *доставляет сильный минимум (максимум) в задаче* (2), *если существует такое* $\varepsilon>0$, *что для любой* $x(\cdot)\in KC^1([t_0, t_1])$, *для которой* $x(t_0)=x_0$, $x(t_1)=x_1$ *и*

$$\|x(\cdot)-\hat{x}(\cdot)\|_0=\sup_{t\in[t_0,t_1]}|x(t)-\hat{x}(t)|<\varepsilon, \qquad (3)$$

выполняется неравенство $\mathcal{J}(x(\cdot))\geqslant\mathcal{J}(\hat{x}(\cdot))$ $(\leqslant\mathcal{J}(\hat{x}(\cdot)))$.

Лемма о скруглении углов. 1) *Если функция* $L=L(t, x, \dot{x})$ *непрерывна по совокупности аргументов, то*

$$\inf_{\substack{x(\cdot)\in KC^1([t_0,t_1])\\ x(t_0)=x_0,\ x(t_1)=x_1}}\mathcal{J}(x(\cdot))=\inf_{\substack{x(\cdot)\in C^1([t_0,t_1])\\ x(t_0)=x_0,\ x(t_1)=x_1}}\mathcal{J}(x(\cdot)). \qquad (4)$$

2) *Равенство* (4) *сохранится, если брать только те* $x(\cdot)\in KC^1([t_0, t_1])$, *которые удовлетворяют неравенствам* (3) *для заданных* $\hat{x}(\cdot)$ *и* $\varepsilon>0$.

3) *Утверждение остается верным и при замене* inf *на* sup.

Доказательство. Поскольку $KC^1([t_0, t_1])\supset\supset C^1([t_0, t_1])$, левая часть в (4) не больше правой и нужно лишь доказать обратное неравенство. Предположим противное. Если $\inf\limits_{KC^1}\mathcal{J}(x(\cdot))<\inf\limits_{C^1}\mathcal{J}(x(\cdot))$, то найдутся такая кусочно-гладкая $\tilde{x}(\cdot)$ и такое $\eta>0$, что $\mathcal{J}(\tilde{x}(\cdot))<<\inf\limits_{C_1}\mathcal{J}(x(\cdot))-\eta$. Пусть τ_i, $i=1, 2, \ldots, m$, — точки разрыва производной $\dot{\tilde{x}}$ и $\Delta_i=\dot{\tilde{x}}(\tau_i+0)-\dot{\tilde{x}}(\tau_i-0)$ — ее скачки в этих точках.

На замкнутом ограниченном множестве

$$\mathcal{K}=\{(t, x, \dot{x})\,|\,t_0\leqslant t\leqslant t_1,\ |x-\tilde{x}(t)|\leqslant \leqslant\max|\Delta_i|\,\delta_0/4,\ |\dot{x}-\dot{\tilde{x}}(t)|\leqslant\max|\Delta_i|/2\}$$

непрерывная функция L ограничена: $|L(t, x, \dot{x})|\leqslant M$.

Функция

$$a(t)=\begin{cases}\dfrac{(1-|t|)^2}{4}, & |t|\leqslant 1,\\ 0, & |t|>1\end{cases} \qquad (5)$$

непрерывна, а ее производная при $t=0$ имеет скачок величины — 1. Поэтому функция $\delta a\left(\dfrac{t-\tau_i}{\delta}\right)$, график которой получается из графика $a(\cdot)$ подобным преобразованием и сдвигом (рис. 22) также непрерывна, и ее производная непрерывна, кроме точки τ_i, где она по-прежнему имеет скачок — 1.

Теперь нетрудно проверить, что функция

$$x_\delta(t)=\tilde{x}(t)+\sum_{i=1}^{m}\Delta_i\delta a\left(\frac{t-\tau_i}{\delta}\right)$$

непрерывна вместе со своей производной на $[t_0, t_1]$, причем $x_\delta(t)\equiv\tilde{x}(t)$ вне отрезков $[\tau_i-\delta,\ \tau_i+\delta]$. В частности, для достаточно малых δ эти отрезки не перекрываются, $x_\delta(t_i)=$

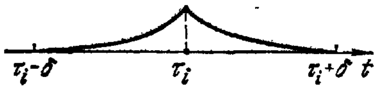


Рис. 22.

$=\tilde{x}(t_i)=x_i,\quad i=0,1,\quad$ и $\quad|x_\delta(t)-\tilde{x}(t)|\leqslant\max|\Delta_i|\delta/4$, $|\dot{x}_\delta(t)-\dot{\tilde{x}}(t)|\leqslant\max|\Delta_i|/2$, ибо согласно (5) $|a(t)|\leqslant 1/4$, $|\dot{a}(t)|\leqslant 1/2$. Поэтому при $\delta<\delta_0$

$$\mathcal{J}(x_\delta(\cdot))-\mathcal{J}(\tilde{x}(\cdot))=$$

$$=\int_{t_0}^{t_1}L(t,\ x_\delta(t),\ \dot{x}_\delta(t))\,dt-\int_{t_0}^{t_1}L(t,\ \tilde{x}(t),\ \dot{\tilde{x}}(t))\,dt=$$

$$=\sum_{i=1}^{m}\int_{\tau_i-\delta}^{\tau_i+\delta}(L(t,\ x_\delta(t),\ \dot{x}_\delta(t))-L(t,\ \tilde{x}(t),\ \dot{\tilde{x}}(t)))\,dt$$

и, следовательно,

$$\mathcal{J}(x_\delta(\cdot))\leqslant\mathcal{J}(\tilde{x}(\cdot))+4m\delta M<$$
$$<\inf_{C^1}\mathcal{J}(x(\cdot))-\eta+4m\delta M<\inf_{C^1}\mathcal{J}(x(\cdot)),$$

если δ достаточно мало. Так как $x_\delta(\cdot)\in C^1([t_0,\ t_1])$, мы пришли к противоречию. При малых δ функция $x_\delta(\cdot)$ будет удовлетворять неравенствам (3), если $\tilde{x}(\cdot)$ им удовлетворяет. Отсюда следует второе утверждение леммы. Третье утверждение очевидно. ■

Следствие. *Если функция $\hat{x}(\cdot)\in C^1([t_0,\ t_1])$ доставляет в задаче* (2) *абсолютный или сильный минимум (максимум) в пространстве $C^1([t_0,\ t_1])$, то она обладает тем же свойством и в пространстве $KC^1([t_0,\ t_1])$.*

Проведенное нами расширение на самом деле не всегда оказывается достаточным. Кроме того, пространство $KC^1([t_0,\ t_1])$ не является полным (относительно метрики, определяемой левой частью (3)). Более естественным было бы расширение задачи на класс $W^1_\infty([t_0,\ t_1])$ функций, удовлетворяющих условию Липшица[1]. Но это по-

[1]) Отметим, однако, что и при этом расширении мы не получим существования в примере Гильберта: функция $t\mapsto t^{1/3}$ условию Липшица не удовлетворяет.

требовало бы некоторых сведений из теории функций. Есть и еще одна важная конструкция расширения простейшей задачи, которую мы здесь обсудим на примере.

Пример Больца. Рассмотрим задачу

$$\mathcal{J}(x(\cdot))=\int_0^1((1-\dot{x}^2)^2+x^2)\,dt\to\inf,\quad x(0)=0,\quad x(1)=\xi.$$

Допустим, что $\xi=0$. Тогда ясно, что решение задачи не существует. Действительно, нижняя грань функционала равна здесь нулю. Для того чтобы понять это, достаточно рассмотреть минимизирующую последовательность функций из $KC^1([0,1])$:

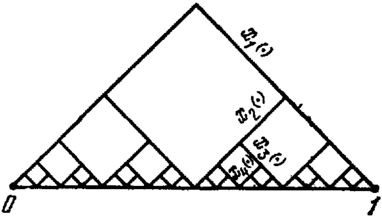


Рис. 23.

$$x_n(t)=\int_0^t \operatorname{sign}\sin 2\pi n\tau\,d\tau,\quad n=1,2,\ldots \text{ (рис. 23).}$$

Функции $x_n(\cdot)$ равномерно стремятся к нулю и $\dot{x}_n^2(t)\equiv 1$, за исключением конечного числа точек, т. е. $\mathcal{J}(x_n(\cdot))=\int_0^1 x_n^2(t)\,dt\to 0$. С другой стороны, если $x_0(\cdot)\equiv 0$, то

$$\mathcal{J}(x_0(\cdot))=1,\quad \text{а если}\quad x(\cdot)\not\equiv 0,\quad \text{то}\quad \mathcal{J}(x(\cdot))\geqslant\int_0^1 x^2dt>0.$$

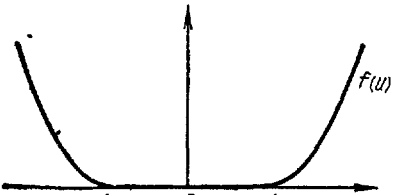


Рис. 24.

Дело в том, что функционал $\mathcal{J}$ не полунепрерывен снизу — мы установили, что в любой близости от функции $x_0(\cdot)\equiv 0$, где функционал принимает значение единица, есть функции, где значение функционала значительно меньше ($\mathcal{J}(x_n(\cdot))\to 0$). Так вот, возможно «полунепрерывное снизу» расширение задачи, когда функционал $\mathcal{J}$ заменяется функционалом

$$\mathcal{Y}(x(\cdot))=\varliminf_{y(\cdot)\to x(\cdot)}\mathcal{J}(x(\cdot)),$$

а сам запас первоначальных функций (скажем, $C^1([t_0, t_1])$ или $KC^1([t_0, t_1])$) остается прежним. Предельный переход $y(\cdot) \to x(\cdot)$ понимается здесь в пространстве $C([t_0, t_1])$ (т. е. в смысле равномерной сходимости). Оказывается, что функционал $\mathcal{Y}$ допускает простое описание. В примере Больца

$$\mathcal{Y}(x(\cdot)) = \int_0^1 (((\dot{x}^2 - 1)_+)^2 + x^2)\, dt,$$

где

$$(\dot{x}^2 - 1)_+ = \begin{cases} 0, & |\dot{x}| \leqslant 1, \\ \dot{x}^2 - 1, & |\dot{x}| > 1. \end{cases}$$

В общем случае, когда

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x, \dot{x})\, dt,$$

$$\mathcal{Y}(x(\cdot)) = \int_{t_0}^{t_1} \tilde{L}(t, x, \dot{x})\, dt,$$

где $\tilde{L}$ — «овыпукление» L по $\dot{x}$, т. е. функция $\dot{x} \to \tilde{L}(t, x, \dot{x})$ есть наибольшая выпуклая функция, не превосходящая функции $\dot{x} \to L(t, x, x)$. Это утверждение называется *теоремой Боголюбова*

Этот же пример Больца можно использовать для построения задачи (2), в которой минимум достигается на кривой с изломом. Однако по чисто техническим соображениям мы рассмотрим несколько иную задачу:

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} (f(\dot{x}) + x^2)\, dt \to \inf, \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad (6)$$

где функция

$$f(u) = ((|u| - 1)_+)^2 = \begin{cases} (u-1)^2, & u \geqslant 1, \\ 0, & |u| \leqslant 1, \\ (u+1)^2, & u \leqslant -1 \end{cases}$$

непрерывно дифференцируема и выпукла (рис. 24).

В частности, график $f(u)$ лежит не ниже любой своей касательной, т. е. всегда

$$f(v) \geqslant f(u) + f'(u)(v - u).$$

Поэтому для любых двух функций $x(\cdot), \hat{x}(\cdot) \in KC^1([t_0, t_1])$, удовлетворяющих заданным граничным условиям $x(t_i) = \hat{x}(t_i) = x_i$, $i = 0, 1$, имеем

$$\mathcal{J}(x(\cdot)) - \mathcal{J}(\hat{x}(\cdot)) =$$
$$= \int_{t_0}^{t_1} (f(\dot{x}(t)) - f(\dot{\hat{x}}(t)) + x^2(t) - \hat{x}^2(t))\, dt \geqslant$$
$$\geqslant \int_{t_0}^{t_1} (f'(\dot{\hat{x}}(t))(\dot{x}(t) - \dot{\hat{x}}(t)) +$$
$$+ 2\hat{x}(t)(x(t) - \hat{x}(t)) + (x(t) - \hat{x}(t))^2)\, dt \geqslant$$
$$\geqslant \int_{t_0}^{t_1} (f'(\dot{\hat{x}}(t))(\dot{x}(t) - \dot{\hat{x}}(t)) + 2\hat{x}(t)(x(t) - \hat{x}(t)))\, dt.$$

Предположим теперь, что $\hat{x}(\cdot)$ во всех точках диффенцируемости удовлетворяет уравнению Эйлера

$$\frac{d}{dt} f'(\dot{\hat{x}}(t)) = 2\hat{x}(t). \tag{7}$$

Интегрируя по частям на каждом отрезке непрерывности $\dot{\hat{x}}(\cdot)$, получаем

$$\mathcal{J}(x(\cdot)) - \mathcal{J}(\hat{x}(\cdot)) \geqslant$$
$$\geqslant \int_{t_0}^{t_1} \left(-\frac{d}{dt} f'(\dot{\hat{x}}(t)) + 2\hat{x}(t)\right)(x(t) - \hat{x}(t))\, dt +$$
$$+ \sum_i [f'(\dot{\hat{x}}(\tau_i + 0) - f'(\dot{\hat{x}}(\tau_i - 0))(x(\tau_i) - \hat{x}(\tau_i)) = 0,$$

если в дополнение к (7) функция $p(t) = f'(\dot{\hat{x}}(t))$ (импульс) непрерывна. Таким образом, функция $\hat{x}(\cdot)$, удовлетворяющая уравнению Эйлера, условию непрерывности $p(\cdot)$ и краевым условиям, является решением задачи (6). Например, $\hat{x}(t) = e^{|t|}$ с изломом в точке $t = 0$ является решением этой задачи для $t_0 = -1$, $t_1 = 1$, $x_0 = x_1 = e$. Действительно,

$$\hat{x}(t) = \begin{cases} e^t, & t \geqslant 0, \\ e^{-t}, & t < 0, \end{cases}$$

откуда $|\dot{\hat{x}}(t)| \geqslant 1$ и, значит,

$$p(t) = f'(\dot{\hat{x}}(t)) = \begin{cases} 2(e^t - 1), & t \geqslant 0, \\ 2(-e^{-t} + 1), & t < 0, \end{cases}$$

скачка в точке $t=0$ не имеет: $p(+0) = p(-0) = 0$. Кроме того, $\frac{d}{dt} p(t) = 2e^{|t|} = 2\hat{x}(t)$, так что уравнение Эйлера удовлетворяется.

1.4.4. Игольчатые вариации. Условие Вейерштрасса. Понятие сильного экстремума ввел в вариационное исчисление Вейерштрасс. Для доказательства необходимого

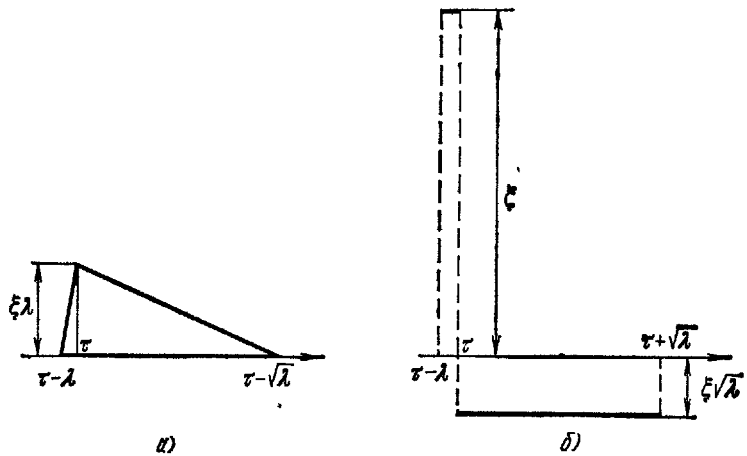


Рис. 25.

условия сильного минимума Вейерштрасс употребил специальные вариации такого вида (рис. 25, *а*)

$$h_\lambda(t) = h_\lambda(t;\, \tau,\, \xi) = \begin{cases} \xi\lambda + (t-\tau)\xi, & t \in [\tau-\lambda,\, \tau], \\ \xi\lambda - (t-\tau)\xi\sqrt{\lambda}, & t \in [\tau,\, \tau+\sqrt{\lambda}], \end{cases} \tag{1}$$

$$x_\lambda(t) = \hat{x}(t) + h_\lambda(t).$$

Производная вариации $h_\lambda(\cdot)$ имеет вид, изображенный на рис. 25, *б*. Эта производная несколько напоминает иголку, в связи с чем подобные вариации называют «игольчатыми». Игольчатые вариации приспособлены к исследованию задач на сильный экстремум.

Перейдем к выводу необходимого условия Вейерштрасса. Рассмотрим простейшую задачу классического вариационного исчисления

$$\mathscr{J}(x(\cdot))=\int_{t_0}^{t_1} L(t, x, \dot{x})\,dt \to \inf, \quad x(t_0)=x_0, \quad x(t_1)=x_1 \quad (2)$$

на классе $KC^1([t_0, t_1])$ кусочно-гладких функций. Пусть $\hat{x}(\cdot)$ — экстремаль, подозреваемая на сильный минимум; для простоты будем считать ее гладкой.

Следуя общему замыслу метода вариаций, рассмотрим функцию

$$\varphi(\lambda)=\mathscr{J}(x_\lambda(\cdot))=\mathscr{J}(\hat{x}(\cdot)+h_\lambda(\cdot)), \quad (3)$$

где h_λ определено формулами (1), τ — внутренняя точка $[t_0, t_1]$, ξ — произвольное число.

Для достаточно малых $\lambda \geqslant 0$ функция $x_\lambda(\cdot)$ допустима в задаче (2), т. е. $x_\lambda(t_i)=\hat{x}(t_i)+h_\lambda(t_i)=x_i$, $i=0, 1$.

Функция $\varphi(\ )$ определена при неотрицательных λ. Докажем, что она дифференцируема справа в точке нуль. Из определения (1) сразу видно, что

$$\begin{aligned} &\text{а)}\ \|h_\lambda(\cdot)\|_0=\max_{t\in[t_0, t_1]}|h_\lambda(t)|=O(\lambda),\\ &\text{б)}\ |\dot{h}_\lambda(t)| =\xi\sqrt{\lambda}=O(\sqrt{\lambda}), \quad t\in(\tau, \tau+\sqrt{\lambda}). \end{aligned} \quad (4)$$

Отсюда

$$\varphi(\lambda)-\varphi(0)=\int_{\tau-\lambda}^{\tau}(L(t, x_\lambda(t), \dot{\hat{x}}(t)+\xi)-L(t, \hat{x}(t), \dot{\hat{x}}(t)))\,dt+$$
$$+\int_{\tau}^{\tau+\sqrt{\lambda}}(L(t, x_\lambda(t), \dot{x}_\lambda(t))-L(t, \hat{x}(t), \dot{\hat{x}}(t)))\,dt=\mathscr{Y}_1+\mathscr{Y}_2. \quad (5)$$

Интеграл $\mathscr{Y}_1$ в (5) можно оценить так:

$$\mathscr{Y}_1=\lambda(L(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau)+\xi)-L(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau)))+o(\lambda) \quad (6)$$

(надо воспользоваться теоремой о среднем из дифференциального исчисления, см. п. 2.2.3, и оценкой (4а)).

Для оценки второго интеграла $\mathscr{Y}_2$ представим разность

$$\Delta=L(t, \hat{x}(t)+h_\lambda(t), \dot{\hat{x}}(t)+\dot{h}_\lambda(t))-L(t, \hat{x}(t), \dot{\hat{x}}(t))$$

в виде

$$\Delta=L_x(t, \hat{x}(t), \dot{\hat{x}}(t))\,h_\lambda(t)+L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))\,\dot{h}_\lambda(t)+o(\sqrt{\lambda})$$

(опять-таки надо применить теорему о среднем и оценку (4б)), проинтегрируем второй член по частям и воспользуемся тем, что

$$-\frac{d}{dt}L_{\dot{x}}+L_x\Big|_{\hat{x}(t)}=0$$

(ибо $\hat{x}(\ )$ — экстремаль). В итоге получаем

$$\mathscr{Y}_2=-\xi\lambda L_{\dot{x}}(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau))+\int\limits_{\tau}^{\tau+\sqrt{\lambda}} o(\sqrt{\lambda})\,dt=$$

$$=-\xi\lambda L_{\dot{x}}(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau))+o(\lambda). \qquad (7)$$

Сопоставляя (6) и (7), находим

$$\varphi(\lambda)-\varphi(0)=\lambda\,(L(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau)+\xi)-$$

$$-L(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau))-\xi L_{\dot{x}}(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau)))+o(\lambda). \qquad (8)$$

Таким образом, функция $\varphi(\ )$ имеет в точке $\lambda=0$ правую производную

$$\varphi'(+0)=\lim_{\lambda\downarrow 0}\frac{\varphi(\lambda)-\varphi(0)}{\lambda}=$$

$$=L(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau)+\xi)-L(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau))-\xi L_{\dot{x}}(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau)).$$

Но если $\hat{x}(\cdot)$ доставляет сильный минимум, то $\mathcal{J}(x_\lambda(\cdot))\geqslant\mathcal{J}(\hat{x}(\cdot))$ и, значит,

$$\varphi'(+0)=\lim_{\lambda\downarrow 0}(\varphi(\lambda)-\varphi(0))/\lambda\geqslant 0,$$

т. е. выполнено соотношение

$$\mathscr{E}(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau),\ \dot{\hat{x}}(\tau)+\xi)=$$

$$=L(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau)+\xi)-L(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau))-$$

$$-\xi L_{\dot{x}}(\tau,\ \hat{x}(\tau),\ \dot{\hat{x}}(\tau))\geqslant 0 \qquad (9)$$

для любого $\xi\in\mathbb{R}$. Функция

$$\mathscr{E}(t,\ x,\ y,\ z)=L(t,\ x,\ z)-L(t,\ x,\ y)-(z-y)\,L_y(t,\ x,\ y)$$

называется *функцией Вейерштрасса*.

Таким образом, доказана следующая

Теорема (необходимое условие Вейерштрасса для сильного минимума).

Для того чтобы экстремаль $\hat{x}(\cdot)\in C^1([t_0,\ t_1])$ простейшей задачи классического вариационного исчисления (2) *доставляла сильный минимум, необходимо, чтобы для* *любого* $\tau\in[t_0, t_1]$ *и любого* $\xi\in\mathbf{R}$ *было выполнено неравенство*

$$\mathscr{E}(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau), \dot{\hat{x}}(\tau)+\xi)=L(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau)+\xi)-$$
$$-L(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau))-\xi L_{\dot{x}}(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau))\geqslant 0.$$

1.4.5. Изопериметрическая задача и задача со старшими производными. *Изопериметрической задачей* (с закрепленными концами) в вариационном исчислении называют обычно такую задачу:

$$\mathscr{I}_0(x(\cdot))=\int_{t_0}^{t_1} f_0(t, x, \dot{x})dt\to\operatorname{extr},$$
$$\mathscr{I}_i(x(\cdot))=\int_{t_0}^{t_1} f_i(t, x, \dot{x})\,dt=\alpha_i,\ i=1, \ldots, m, \tag{1}$$
$$x(t_0)=x_0,\ x(t_1)=x_1.$$

Предполагается, что функции f_i удовлетворяют тем же условиям, что и L в п. 1.4 1 они сами и их частные производные по x и $\dot{x}$ непрерывны по совокупности переменных. Для простоты ограничимся случаем $m=1$.

Теорема (необходимое условие экстремума в изопериметрической задаче) *Пусть функция* $\hat{x}(\cdot)$ *доставляет локальный (относительно пространства* $C^1([t_0, t_1])$*) экстремум в задаче*

$$\mathscr{I}_0(x(\cdot))=\int_{t_0}^{t_1} f_0(t, x, \dot{x})\,dt\to\operatorname{extr},$$
$$\mathscr{I}_1(x(\cdot))=\int_{t_0}^{t_1} f_1(t, x, \dot{x})\,dt=\alpha_1,\quad x(t_i)=x_i,\ i=0, 1 \tag{1'}$$

и при этом функции $t\mapsto f_{0x}(t, \hat{x}(t), \dot{\hat{x}}(t))$ *и* $t\mapsto f_{1x}(t, \hat{x}(t), \dot{\hat{x}}(t))$ *принадлежат* $C^1([t_0, t_1])$. *Тогда найдутся числа* $\hat{\lambda}_0, \hat{\lambda}_1$, *не равные нулю одновременно и такие, что для интегранта* $L=\hat{\lambda}_0 f_0+\hat{\lambda}_1 f_1$ *выполнено уравнение Эйлера*

$$-\frac{d}{dt}L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))+L_x(t, \hat{x}(t), \dot{\hat{x}}(t))=0. \tag{2}$$

Доказательство. А) Аналогично тому, как это было проделано в начале доказательства теоремы п. 1.4.1, вычисляем первые вариации функционалов $\mathcal{J}_0$ и $\mathcal{J}_1$ по Лагранжу:

$$\delta\mathcal{J}_i(\hat{x}(\cdot), x(\cdot)) = \int\limits_{t_0}^{t_1} (p_i(t)\dot{x}(t) + q_i(t)x(t))\,dt,\ i=0,1, \tag{3}$$

где

$$p_i(t) = f_{i\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)),\ q_i(t) = f_{ix}(t, \hat{x}(t), \dot{\hat{x}}(t)).$$

Возможно одно из двух: или $\delta\mathcal{J}_1 \equiv 0$, $\forall x(\cdot) \in C^1([t_0, t_1])$, $x(t_0) = x(t_1) = 0$ или $\delta\mathcal{J}_1 \not\equiv 0$. В первом случае по теореме п. 1.4.1

$$-\frac{d}{dt} f_{1\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) + f_{1x}(t, \hat{x}(t), \dot{\hat{x}}t) = (0)$$

и, положив $\hat{\lambda}_0 = 0$, $\hat{\lambda}_1 = 1$, немедленно получаем (2).

Б) Пусть $\delta\mathcal{J}_1 \not\equiv 0$ и, значит, существует функция $y(\cdot) \in C^1([t_0, t_1])$, $y(t_0) = y(t_1) = 0$, для которой вариация $\delta\mathcal{J}_1(\hat{x}(\cdot), y(\cdot)) \neq 0$. Рассмотрим функции двух переменных

$$\varphi_i(\alpha, \beta) = \mathcal{J}_i(\hat{x}(\cdot) + \alpha x(\cdot) + \beta y(\cdot)),\quad i = 0, 1,$$

предоставив читателю проверить, что они непрерывно дифференцируемы в окрестности нуля, причем

$$\frac{\partial\varphi_i(0,0)}{\partial\alpha} = \delta\mathcal{J}_i(\hat{x}(\cdot), x(\cdot)),\quad \frac{\partial\varphi_i(0,0)}{\partial\beta} = \delta\mathcal{J}_i(\hat{x}(\cdot), y(\cdot)).$$

Лемма. *Для любой* $x(\cdot) \in C^1([t_0, t_1])$ *такой, что* $x(t_0) = x(t_1) = 0$

$$\frac{\partial(\varphi_0, \varphi_1)}{\partial(\alpha, \beta)}(0, 0) =$$

$$= \begin{vmatrix} \dfrac{\partial\varphi_0}{\partial\alpha}, & \dfrac{\partial\varphi_0}{\partial\beta} \\ \dfrac{\partial\varphi_1}{\partial\beta} & \dfrac{\partial\varphi_1}{\partial\beta} \end{vmatrix}_{(0,\ 0)} = \begin{vmatrix} \delta\mathcal{J}_0(\hat{x}(\cdot), x(\cdot)), & \delta\mathcal{J}_0(\hat{x}(\cdot), y(\cdot)) \\ \delta\mathcal{J}_1(\hat{x}(\cdot), x(\cdot)), & \delta\mathcal{J}_1(\hat{x}(\cdot), y(\cdot)) \end{vmatrix} = 0. \tag{4}$$

Доказательство. Если определитель $(4) \neq 0$, то отображение $(\alpha, \beta) \mapsto (\varphi_0(\alpha, \beta), \varphi_1(\alpha, \beta))$ переводит некоторую окрестность точки $(0, 0)$ на некоторую окрестность точки $(\varphi_0(0, 0), \varphi_1(0, 0))$ (соответствующую теорему об обратном отображении мы уже использовали в п. 1.3.2). В частности, найдутся такие α и β, а, следовательно, и допустимая функция $\hat{x}(\cdot) + \alpha x(\cdot) + \beta y(\cdot)$, что $\varphi_0(\alpha, \beta) = \mathcal{J}_0(\hat{x}(\cdot) + \alpha x(\cdot) + \beta y(\cdot)) = \mathcal{J}_0(\hat{x}(\cdot)) - \varepsilon$, где $\varepsilon > 0$, а

$\varphi_1(\alpha, \beta) = \varphi_1(0, 0) = \mathcal{J}_1(\hat{x}(\cdot)) = \alpha_1$. Это противоречит тому, что $\hat{x}(\cdot)$ доставляет локальный минимум в задаче (1). ■

В) Из равенства (4) имеем

$$\delta\mathcal{J}_0(\hat{x}(\cdot), x(\cdot)) - \frac{\delta\mathcal{J}_0(\hat{x}(\cdot), y(\cdot))}{\delta\mathcal{J}_1(\hat{x}(\cdot), y(\cdot))}\delta\mathcal{J}_1(\hat{x}(\cdot), x(\cdot)) = 0$$

(знаменатель здесь по предположению отличен от нуля). Полагая $\hat{\lambda}_0 = 1$, $\hat{\lambda}_1 = -\delta\mathcal{J}_0(\hat{x}(\cdot),\ y(\cdot))\,/\delta\mathcal{J}_1(\hat{x}(\cdot),\ y(\cdot))$, убеждаемся, что

$$\hat{\lambda}_0\delta\mathcal{J}_0(\hat{x}(\cdot), x(\cdot)) + \hat{\lambda}_1\delta\mathcal{J}_1(\hat{x}(\cdot), x(\cdot)) \equiv 0 \tag{5}$$

для любой $x(\cdot) \in C^1([t_0, t_1])$, у которой $x(t_0) = x(t_1) = 0$.

Легко видеть, что левая часть (5) имеет вид

$$\int_{t_0}^{t_1} ((\hat{\lambda}_0 p_0(t) + \hat{\lambda}_1 p_1(t))\dot{x}(t) + (\hat{\lambda}_0 q_0(t) + \hat{\lambda}_1 q_1(t))x(t))\,dt.$$

Применяя основную лемму вариационного исчисления, получаем уравнение

$$-\frac{d}{dt}(\hat{\lambda}_0 p_0(t) + \hat{\lambda}_1 p_1(t)) + (\hat{\lambda}_0 q_0(t) + \hat{\lambda}_1 q_1(t)) = 0,$$

совпадающее с (2). ■

Задачу (1′) впервые рассмотрел Эйлер в своей знаменитой работе 1744 г. Там же методом ломаных было получено соотношение (2). Собственно говоря, это и составляло основное содержание работы. Несомненно, в ней уже содержатся начáла метода множителей Лагранжа.

В заключение бегло рассмотрим задачу со старшими производными:

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} f(t, x, \dot{x}, \ddot{x}, \ldots, x^{(n)})\,dt \to \mathrm{extr},$$

$$x^{(j)}(t_i) = x_i^j, \quad i = 0, 1,\ j = 0, \ldots, n-1.$$

За подробностями отошлем читателя к [2] или [3]. Эту задачу будем исследовать в пространстве $C^n([t_0, t_1])$ функций, непрерывных вместе с производными до порядка n на отрезке $[t_0, t_1]$. Функцию f и ее производные по $x, \dot{x}, \ldots, x^{(n)}$ будем предполагать непрерывными по совокупности переменных. Пусть $\hat{x}(\cdot)$ — функция, подозреваемая на экстремум. Вычислим первую вариацию функционала по Лагранжу:

$$\delta \mathcal{J}(\hat{x}(\cdot),\ x(\cdot))=\int_{t_0}^{t_1}\left(\sum_{j=0}^{n} p_j(t)\,x^{(j)}(t)\right)dt, \tag{7}$$

где $p_j(t)=\dfrac{\partial f}{\partial x^{(j)}}(t,\ \hat{x}(t),\ \ldots,\ \hat{x}^{(n)}(t))$.

Из условия локальной экстремальности вытекает, что $\delta\mathcal{J}(\hat{x}(\ ),\ x(\cdot))=0$, если только $x^{(j)}(t_i)=0$, $i=0,\ 1$, $j=0,\ 1,\ \ldots,\ n-1$. Интегрируя (7) по частям (предоставив читателю сформулировать нужные требования на гладкость $p_j(\ )$), преобразуем первую вариацию к виду

$$\delta\mathcal{J}(\hat{x}(\cdot),\ x(\ ))=\int_{t_0}^{t_1}\left(\sum_{j=0}^{n}(-1)^j p_j^{(j)}(t)\right)x(t)\,dt$$

и, применив обобщение основной леммы классического вариационного исчисления на случай функций из $C^n([t_0,\ t_1])$ с условиями $x^{(j)}(t_i)=0$, $i=0,\ 1$, $j=0,\ 1,\ \ldots,\ n-1$ (это обобщение также предоставляем продумать читателю), получим, что необходимым условием экстремальности $\hat{x}(\ )$ является выполнение следующего уравнения, называемого *уравнением Эйлера—Пуассона*:

$$\sum_{j=0}^{n}(-1)^j\left(\frac{d}{dt}\right)^j\frac{\partial f}{\partial x^{(j)}}(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \ldots,\ \hat{x}^{(n)}(t))=0.$$

§ 1.5. Задача Лагранжа и основная задача оптимального управления

1.5.1. Постановки задач. Важным этапом в истории естествознания явилось сочинение Жозефа Луи Лагранжа «Аналитическая механика», опубликованное в 1788 г. В частности, трактат Лагранжа сыграл исключительную роль и в развитии вариационного исчисления. Именно там была поставлена следующая задача на условный экстремум:

$$\mathcal{J}(x(\cdot))=\int_{t_0}^{t_1} f(t,\ x(t),\ \dot{x}(t))\,dt\to\mathrm{extr}, \tag{1}$$

$$\Phi(t,\ x(t),\ \dot{x}(t))=0\Leftrightarrow\Phi_j(t,\ x(t),\ \dot{x}(t))=0,$$
$$j=1,\ \ldots,\ m, \tag{2}$$
$$x(t_0)=x_0,\ x(t_1)=x_1.$$

Здесь $x(\cdot)$: $[t_0, t_1] \to \mathbf{R}^n$, f: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, Φ: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^m$. Задачу (1) и различные ее модификации, связанные с дополнительными ограничениями (другими граничными условиями, дополнительными интегральными соотношениями и т. п.; см. также 1.2.6), стали называть впоследствии *задачей Лагранжа*. Для решения задачи (1) Лагранж использовал тот основной прием, о котором говорилось в п. 1.3.2,—правило множителей. Впрочем, оно не было им аккуратно обосновано, и потребовалось более ста лет для того, чтобы придать рассуждениям Лагранжа вид строго доказанной теоремы.

Отметим два наиболее важных частных типа ограничений, охватываемых общим выражением (2). Ограничение $\Phi(t, x) = 0$, когда функция в (2) не зависит от $\dot{x}$, называется в вариационном исчислении *фазовым*. В механике фазовые ограничения называют также *голономными связями*.

Другой случай—когда соотношения (2) можно разрешить относительно производных $\dot{x}$. Тогда это ограничение записывается в форме уравнений $\dot{x} = \varphi(t, x, u)$, $x(\cdot)$: $[t_0, t_1] \to \mathbf{R}^n$, $u(\cdot)$: $[t_0, t_1] \to \mathbf{R}^r$, φ: $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^r \to \mathbf{R}^n$. Переменные $x(\cdot)$ здесь называют *фазовыми*, переменные $u(\cdot)$—*управлениями*. Этому важнейшему случаю мы и уделим основное внимание. Точнее говоря, будем рассматривать далее задачу

$$\mathcal{J}(x(\cdot), u(\cdot)) = \\ = \int_{t_0}^{t_1} f(t, x(t), u(t))\, dt + \psi_0(x(t_0), x(t_1)) \to \mathrm{extr}, \quad (1')$$

$$\dot{x} - \varphi(t, x, u) = 0, \ \psi(x(t_0), x(t_1)) = \\ = 0 \ (\Leftrightarrow \psi_j(x(t_0), x(t_1)) = 0, \quad j = 1, \ldots, s). \quad (2')$$

При этом в (1'), (2') f: $V \to \mathbf{R}$, φ: $V \to \mathbf{R}^n$, ψ: $W \to \mathbf{R}^s$, где V и W—открытые множества в пространствах $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^r$ и $\mathbf{R}^n \times \mathbf{R}^n$ соответственно. Моменты времени t_0 и t_1 будем считать здесь фиксированными.

Ограничение $\dot{x} - \varphi(t, x, u) = 0$ называется *дифференциальной связью*, ограничения $\psi(x(t_0), x(t_1)) = 0$ называются *граничными* или *краевыми условиями*. Задачу (1'), (2') будем называть *задачей Лагранжа в понтрягинской форме*.

Все функции f, φ и ψ предположим непрерывно дифференцируемыми.

Далее в гл. IV будет рассмотрен еще более общий случай, когда t_0 и t_1 также могут меняться, допускаются изопериметрические ограничения типа равенств, неравенств и т. п.

Задачу (1′), (2′) будем рассматривать в банаховом пространстве $Z = C^1([t_0, t_1], \mathbf{R}^n) \times C([t_0, t_1], \mathbf{R}^r)$ Иначе говоря, будем рассматривать пары $(x(\cdot), u(\cdot))$, где $x(\cdot)$ — непрерывно дифференцируемая n-мерная вектор-функция, а $u(\cdot)$ — непрерывная r-мерная вектор-функция. Пару $(x(\cdot), u(\cdot))$ будем иногда сокращенно обозначать через z.

Элемент $z = (x(\cdot), u(\cdot)) \in Z$ будем называть *управляемым процессом* в задаче (1′), (2′), если $\dot{x}(t) = \varphi(t, x(t), u(t))$, $\forall t \in [t_0, t_1]$, и *допустимым управляемым процессом*, если, кроме того, удовлетворяются краевые условия. Допустимый элемент $\hat{z} = (\hat{x}(\cdot), \hat{u}(\cdot))$ будет называться *оптимальным в слабом смысле процессом* или *слабым минимумом* в задаче (1′), (2′), если он доставляет локальный минимум в задаче, т. е. если найдется такое $\varepsilon > 0$, что коль скоро $\|x - \hat{x}\|_1 < \varepsilon$ и $\|u - \hat{u}\|_0 < \varepsilon$, то $\mathcal{J}(z) \geqslant \mathcal{J}(\hat{z})$.

1.5.2. Необходимые условия в задаче Лагранжа. Попробуем применить к задаче (1′), (2′) предыдущего пункта общий прием Лагранжа, о котором шла речь в п. 1.3.2. По аналогии с конечномерным случаем функцию Лагранжа следует записать так:

$$\mathscr{L}(x(\cdot), u(\cdot); p(\cdot), \mu, \lambda_0) = \int_{t_0}^{t_1} L\,dt + l, \tag{1}$$

где

$$L(t, x, \dot{x}, u) = p(t)(\dot{x} - \varphi(t, x, u)) + \lambda_0 f(t, x, u), \tag{2}$$

$$l(x_0, x_1) = \sum_{j=0}^{s} \mu_j \psi_j(x_0, x_1), \quad \lambda_0 = \mu_0. \tag{3}$$

То, что «терминальная часть» или терминант функции Лагранжа l имеет вид (3), не вызывает сомнения — здесь имеется полное подобие с конечномерным случаем. Что же касается ограничения $\dot{x} = \varphi(t, x, u)$, то оно должно выполняться для всех $t \in [t_0, t_1]$, и соответствующий «множитель Лагранжа» $p(\cdot)$ по аналогии должен быть функцией t, а его вклад в функцию Лагранжа имеет вид интеграла, а не суммы.

Итак, функция Лагранжа составлена. Следуя рецепту Лагранжа, нужно теперь искать условия экстремума полученного выражения, «как если бы переменные были независимы». Иначе говоря, следует рассмотреть задачу

$$\mathscr{L}(x(\cdot),\ u(\cdot);\ \hat{p}(\cdot),\ \hat{\mu},\ \hat{\lambda}_0)\to \mathrm{extr}, \tag{4}$$

считая множители Лагранжа фиксированными. Задача (4)— это задача Больца, рассмотренная в п. 1.4.2. Выписанные там условия экстремума в применении к задаче (4) приведут к правильным уравнениям, называемым *уравнениям Эйлера—Лагранжа*. Точнее, имеет место следующая теорема.

Теорема Эйлера—Лагранжа. *Если* $\hat{z}=(\hat{x}(\cdot),\ \hat{u}(\cdot))$—*оптимальный в слабом смысле процесс для задачи* (1′), (2′) п. 1.5.1, *то найдутся множители Лагранжа* $\hat{\lambda}_0=\hat{\mu}_0\geqslant 0$ *в задаче на минимум и* $\leqslant 0$ *в задаче на максимум*, $\hat{p}(\cdot)\in C^1([t_0,\ t_1],\ \mathbf{R}^n)$, $\hat{\mu}=(\hat{\mu}_1,\ \ldots,\ \hat{\mu}_s)$, *не равные одновременно нулю*[1]) *и такие, что будут выполнены уравнения Эйлера*

$$-\frac{d}{dt}L_{\dot{x}}(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t))+L_x(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t))=0, \tag{5}$$

$$L_u(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t))=0, \tag{6}$$

и условия трансверсальности

$$L_{\dot{x}}(t_k,\ \hat{x}(t_k),\ \dot{\hat{x}}(t_k),\ \hat{u}(t_k))=(-1)^k\frac{\partial l}{\partial x_k}(\hat{x}(t_0),\ \hat{x}(t_1)),\quad k=0,1. \tag{7}$$

Вследствие того, что L не зависит от $\dot{u}$, а l от u, уравнение Эйлера по u имеет вырожденный вид (6), а условий трансверсальности «по u» нет вовсе.

Эту теорему (и даже в несколько более общем виде) мы докажем в § 4.1 как прямое следствие общей теоремы,

---

[1]) Математик-пурист отметит вопиющую неточность в этой фразе: $\hat{\lambda}_0$ и $\hat{\mu}_i$ это числа, а $\hat{p}(\ )$—элемент функционального пространства, так что они и не могли бы равняться одному и тому же нулю одновременно. Каждое из них может равняться или не равняться своему нулю в своем пространстве. Однако все мы так привыкли отождествлять нули всех пространств, что эта фраза уже не режет глаз.

касающейся правила множителей Лагранжа для гладких бесконечномерных задач.

Отметим, что из этой теоремы следует необходимое условие экстремума в изопериметрической задаче (с произвольным числом изопериметрических условий) и уравнение Эйлера—Пуассона для задач со старшими производными. Изопериметрические ограничения

$$\mathcal{I}_i(x(\cdot))=\int_{t_0}^{t_1} f_i(t,\ x,\ \dot{x})\,dt=\alpha_i,\quad i=1,\ \ldots,\ m;\quad x\in \mathbf{R}^n,$$

можно учесть, введя новые фазовые переменные, связанные со старыми дифференциальной связью $\dot{x}_{n+i}(t)= = f_i(t, x_1(t), \ldots, x_n(t), \dot{x}_1(t), \ldots, \dot{x}_n(t))$ и удовлетворяющие краевым условиям

$$x_{n+i}(t_1)-x_{n+i}(t_0)=\alpha_i,\qquad i=1,\ \ldots,\ m.$$

Если теперь применить теорему Эйлера—Лагранжа, то получатся нужные необходимые условия в изопериметрической задаче. При исследовании задачи со старшими производными ее можно свести к виду (1′), (2′), полагая $x=x_1$, $\dot{x}_1=x_2$, $\ldots$, $\dot{x}_n=u$ и применяя далее теорему Эйлера—Лагранжа.

1.5.3. Принцип максимума Понтрягина. В пятидесятых годах многочисленные потребности прикладных дисциплин (техники, экономики и др.) стимулировали постановку и рассмотреие нового класса экстремальных задач, получивших название *задач оптимального управления*. Необходимое условие экстремума для задач этого класса — «принцип максимума», — сформулированное Л. С. Понтрягиным в 1953 г., было доказано и развито впоследствии им, его учениками и сотрудниками (см. [12]). Важно отметить, что это условие имеет существенно иную форму в сравнении с классическими уравнениями Эйлера и Лагранжа: в качестве обязательиого условия в решение задачи оптимального управления входит решение вспомогательной задачи на максимум (отсюда название — «принцип максимума»).

Здесь мы будем рассматривать частный случай общей постаиовки задачи оптимального управления, когда в задаче Лагранжа в понтрягинской форме (см. (1′) и (2′) п. 1 5.1) появляется еще одно дополнительное условие на управление: $u\in\mathfrak{U}$. Точнее говоря, будет рассматриваться такая задача:

$$\mathcal{Z}(x(\cdot),\ u(\cdot))=$$
$$=\int_{t_0}^{t_1} f(t,\ x(t),\ u(t))\,dt+\psi_0(x(t_0),\ x(t_1))\to\inf, \quad (1)$$

$$\dot{x}(t)-\varphi(t,\ x(t),\ u(t))=0,\quad \psi_j(x(t_0),\ x(t_1))=0,$$
$$j=1,\ \ldots,\ s, \quad (2)$$
$$u\in\mathfrak{U}. \quad (3)$$

Функции f, φ, ψ_j — такие же, как в (1′), (2′) п. 1.5.1, а $\mathfrak{U}$ — фиксированное множество в $\mathbf{R}^r$. Более общая задача будет рассмотрена в гл. IV.

Задачу Лагранжа мы рассматривали в некотором банаховом пространстве. Здесь же, желая применять самые простые средства, изберем иной путь описания допустимых элементов, сходный с тем, о котором шла речь в п. 1.4.3, где простейшая задача классического вариационного исчисления была расширена до задачи в пространстве $KC^1([t_0,\ t_1])$ кусочно-дифференцируемых функций.

Пару $(x(\cdot),\ u(\cdot))$ будем называть *управляемым процессом* в задаче (1)—(3), если управление $u(\cdot)$: $[t_0,\ t_1]\to\to\mathfrak{U}$ — кусочно-непрерывная функция, фазовая траектория $x(\cdot)$ кусочно-непрерывно дифференцируема и при этом всюду, кроме точек разрыва управления $u(\cdot)$, функция $x(\cdot)$ удовлетворяет дифференциальному уравнению $\dot{x}(t)=\varphi(t,\ x(t),\ u(t))$. Управляемый процесс называется *допустимым*, если, кроме того, удовлетворяются краевые условия.

Допустимый процесс $(\hat{x}(\cdot),\ \hat{u}(\cdot))$ называется *оптимальным*, если найдется такое $\varepsilon>0$, что для всякого допустимого управляемого процесса $(x(\cdot),\ u(\cdot))$ такого, что $|x(t)-\hat{x}(t)|<\varepsilon$, $\forall t\in[t_0,\ t_1]$ выполняется неравенство $\mathcal{Z}(x(\cdot),\ u(\cdot))\geqslant\mathcal{Z}(\hat{x}(\cdot),\ \hat{u}(\cdot))$.

Попробуем снова применить к задаче (1)—(3) общий прием Лагранжа, о котором речь шла в п. 1.3.2. Функция Лагранжа задачи (1)—(3) будет иметь тот же вид, что и в задаче Лагранжа — ограничения типа принадлежности ($u\in\mathfrak{U}$) в функцию Лагранжа не включаются. Итак,

$$\mathscr{L}(x(\cdot),\ u(\cdot);\ p(\cdot),\ \mu,\ \lambda_0)=\int_{t_0}^{t_1} L\,dt+l, \quad (4)$$

где

$$L(t, x, \dot{x}, u) = p(t)(\dot{x} - \varphi(t, x, u)) + \lambda_0 f(t, x, u), \quad (5)$$

$$l(x_0, x_1) = \sum_{j=0}^{s} \mu_j \psi_j(x_0, x_1), \quad \mu_0 = \lambda_0. \quad (6)$$

Далее, как всегда, нужно рассмотреть задачу

$$\mathscr{L}(x(\cdot), u(\cdot); \hat{p}(\cdot), \hat{\mu}, \hat{\lambda}_0) \to \inf \quad (7)$$

(считая множители Лагранжа фиксированными) «как если бы переменные были независимы». Задача (7) естественным путем распадается на две:

$$\mathscr{L}(x(\cdot), \hat{u}(\cdot); \hat{p}(\cdot), \hat{\mu}, \hat{\lambda}_0) \to \inf \ (\text{по } x(\cdot)), \quad (8)$$

$$\mathscr{L}(\hat{x}(\cdot), u(\cdot); \hat{p}(\cdot), \hat{\mu}, \hat{\lambda}_0) \to \inf \ (\text{по } u(\cdot) \in \mathscr{U}), \quad (9)$$

где через $\mathscr{U}$ мы обозначили множество кусочно-непрерывных функций со значениями в $\mathfrak{U}$. Задача (8)—это снова задача Больца; задача же (9) имеет, как легко понять, следующий простой вид:

$$\int_{t_0}^{t_1} \chi(t, u(t))\, dt \to \inf, \ u(\cdot) \in \mathscr{U}, \quad (10)$$

где $\chi(t, u) = \hat{\lambda}_0 f(t, \hat{x}(t), u) - \hat{p}(t)\varphi(t, \hat{x}(t), u)$.

Необходимое (и достаточное) условие экстремума в задаче (10) совершенно очевидно: $\hat{u}(\cdot) \in \mathscr{U}$ доставляет минимум в задаче (10) тогда и только тогда, когда всюду, кроме точек разрыва $\hat{u}(\cdot)$ выполнено соотношение

$$\min_{u \in \mathfrak{U}} \chi(t, u) = \chi(t, \hat{u}(t)) \Leftrightarrow$$

$$\Leftrightarrow \min_{u \in \mathfrak{U}} L(t, \hat{x}(t), \dot{\hat{x}}(t), u) = L(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)) \Leftrightarrow$$

$$\Leftrightarrow \max_{u \in \mathfrak{U}} [\hat{p}(t)\varphi(t, \hat{x}(t), u) - \hat{\lambda}_0 f(t, \hat{x}(t), u)] =$$

$$= \hat{p}(t)\varphi(t, \hat{x}(t), \hat{u}(t)) - \hat{\lambda}_0 f(t, \hat{x}(t), \hat{u}(t)). \quad (11)$$

Необходимые условия экстремума в задаче (8), соединенные с (11), приводят к необходимым условиям экстремума в задаче (1)—(3), получившим название **принципа максимума Понтрягина** (из-за специфического вида условия (11)). Точнее, имеет место следующая теорема.

Теорема (принцип максимума Понтрягина). *Если* $(\hat{x}(\cdot), \hat{u}(\cdot))$ *есть оптимальный процесс для задачи* (1)—(3), *то найдутся множители Лагранжа* $\hat{\lambda}_0 = \hat{\mu}_0 \geqslant 0$, $\hat{p}(\cdot) \in KC^1([t_0, t_1], \mathbf{R}^{n*})$, $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_s)$, *не равные нулю одновременно и такие, что выполнено уравнение Эйлера*

$$\frac{d}{dt} L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)), \tag{12}$$

принцип максимума (11) *и условие трансверсальности*

$$L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t))\,|_{t=t_k} = (-1)^k \frac{\partial l}{\partial x_k}(\hat{x}(t_0), \hat{x}(t_1)), \quad k=0,1. \tag{13}$$

1.5.4. Доказательство принципа максимума в задаче со свободным концом. Здесь принцип максимума Понтрягина будет доказан в простейшей ситуации, когда терминальная часть функционала отсутствует, один из концов закреплен, а второй свободен, т. е. в (1) п. 1.5.3 $\psi_0 = 0$, а в (2) $\psi_j(x_0, x_1) = x_{0j} - \bar{x}_{0j}$, $j = 1, \ldots, n$.

Таким образом, мы имеем задачу

$$\mathcal{J}(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} f(t, x(t), u(t))\,dt \to \inf, \tag{1}$$

$$\dot{x}(t) - \varphi(t, x(t), u(t)) = 0, \quad x(t_0) = \bar{x}_0, \tag{2}$$

$$u \in \mathfrak{U}. \tag{3}$$

Функции f, f_x, φ, φ_x предполагаются (как и в предыдущем пункте) непрерывными по совокупности переменных.

Посмотрим, как выглядит в этом случае принцип максимума Понтрягина. Поскольку у нас функция $l(x_0, x_1) = \sum_{i=1}^{m} \hat{\mu}_i (x_{0i} - \bar{x}_{0i})$ не зависит от x_1, из условий трансверсальности (13) п. 1.5.3. получаем

$$L_{\dot{x}}(t_1, \hat{x}(t_1), \dot{\hat{x}}(t_1), \hat{u}(t_1)) = \hat{p}(t_1) = 0. \tag{4}$$

Далее, уравнение (12) п. 1.5.3 приобретает вид

$$-\frac{d}{dt} L_{\dot{x}} + L_x = 0 \Leftrightarrow -\dot{\hat{p}}(t) = \hat{p}(t)\,\hat{\varphi}_x(t) - \hat{\lambda}_0 \hat{f}_x(t), \tag{5}$$

где мы ввели сокращенные обозначения: $\hat{f}_x(t)=f_x(t, \hat{x}(t), \hat{u}(t))$, $\hat{\varphi}_x(t)=\varphi_x(t, \hat{x}(t), \hat{u}(t))$. Если допустить, что $\hat{\lambda}_0=0$, то вследствие единственности решеиия задачи Коши для однородного уравнения (5) должно быть $\hat{p}(\cdot)\equiv 0$, а значит, в силу условия трансверсальности на левом конце (см. (13) п. 1.5.3) и $\hat{\mu}=0$. Но это противоречит условию теоремы, согласно которому не все множители Лагранжа могут быть одновременно нулями. Значит, $\hat{\lambda}_0\neq 0$ и можно считать $\hat{\lambda}_0=1$. Но тогда $\hat{p}(\cdot)$ однозначно (в силу единственности решения задачи Коши для линейной неоднородной системы) определяется уравнением (5) и краевым условием (4). Суммируя сказанное, можно сформулировать принцип максимума так:

Теорема (принцип максимума Понтрягина для задачи со свободным концом). *Если процесс* $(\hat{x}(\cdot), \hat{u}(\cdot))$ *является оптимальным в задаче* (1)—(3), *то для решения* $\hat{p}(\cdot)$ *системы*

$$-\dot{p}(t)=p(t)\hat{\varphi}_x(t)-\hat{f}_x(t) \tag{6}$$

с краевым условием

$$p(t_1)=0 \tag{7}$$

в точках непрерывности управления $\hat{u}(\cdot)$ *выполнен принцип максимума*

$$\max_{u\in\mathfrak{U}}(\hat{p}(t)\varphi(t, \hat{x}(t), u)-f(t, \hat{x}(t), u))=$$
$$=\hat{p}(t)\varphi(t, \hat{x}(t), \hat{u}(t))-f(t, \hat{x}(t), \hat{u}(t)).$$

Для доказательства этой теоремы, как и в предыдущих случаях, воспользуемся методом вариаций. Начнем с определения элементарной — вейерштрассовской — игольчатой вариации, аналогичной той, что была применена в п. 1.4.4. Обозначим через T_0 множество тех точек из (t_0, t_1), в которых функция $\hat{u}(\cdot)$ непрерывна. Зафиксируем точку $\tau\in T_0$, элемент $v\in\mathfrak{U}$ и число $\lambda>0$ настолько малое, что $[\tau-\lambda, \tau]\subset[t_0, t_1]$.

Управление

$$u_\lambda(t)=u_\lambda(t;\ \tau,\ v)=\begin{cases}\hat{u}(t), & \text{если}\quad t\notin[\tau-\lambda, \tau),\\ v, & \text{если}\quad t\in[\tau-\lambda, \tau),\end{cases} \tag{8}$$

назовем *элементарной игольчатой вариацией управления* $\hat{u}(\cdot)$. Пусть $x_\lambda(t)=x_\lambda(t;\ \tau,\ v)$ — решение уравнения $\dot{x}=\varphi(t,\ x,\ u_\lambda(t))$ с начальным условием $x_\lambda(t_0)=\hat{x}(t_0)=x_0$. Назовем $x_\lambda(t)$ *элементарной игольчатой вариацией траектории*, а пару $(x_\lambda(t),\ u_\lambda(t))$ — *элементарной вариацией процесса* $(\hat{x}(\cdot),\ \hat{u}(\cdot))$. Пару $(\tau,\ v)$, определяющую эту вариацию, будем называть *элементарной иголкой.*

Доказательство теоремы как обычно разбиваем на этапы. Первые два этапа целиком относятся к теории обыкновенных дифференциальных уравнений.

А) Лемма о свойствах элементарной вариации. 1. *Пусть элементарная иголка* $(\tau,\ v)$ *фиксирована. Тогда существует такое* $\lambda_0>0$, *что при* $0\leqslant\lambda\leqslant\lambda_0$:

1) *траектория* $x_\lambda(\cdot)$ *определена на всем отрезке* $[t_0,\ t_1]$ *и при* $\lambda\downarrow 0$ $x_\lambda(t)\to\hat{x}(t)$ *равномерно на* $[t_0,\ \tau]$;

2) *при* $t\geqslant\tau$, $0\leqslant\lambda\leqslant\lambda_0$, *существует и непрерывна по* λ *производная*

$$\frac{d}{d\lambda}x_\lambda(t;\ \tau,\ v)=z_\lambda(t;\ \tau,\ v),$$

которая при $\lambda=0$ *определяется как производная справа;*

3) *функция* $t\mapsto y(t)=z_0(t,\ \tau,\ v)$ *на отрезке* $[\tau,\ t_1]$ *удовлетворяет дифференциальному уравнению*

$$\dot{y}=\hat{\varphi}_x(t)\,y \tag{9}$$

и начальному условию

$$y(\tau)=\varphi(\tau,\ \hat{x}(\tau),\ v)-\varphi(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau)). \tag{10}$$

Доказательство леммы основано на двух известных фактах теории обыкновенных дифференциальных уравнений: локальной теореме существования и единственности и теореме о непрерывно дифференцируемой зависимости решения от начальных данных. Эти теоремы в нужной для нас форме содержатся в стандартных учебниках по обыкновенным дифференциальным уравнениям [10, 11, 15]. Кроме того, читатель может обратиться к тексту § 2.5.

Докажем лемму сначала для случая, когда функция $\hat{u}(\ )$ непрерывна.

Рассмотрим дифференциальные уравнения

$$\dot{x}=\varphi(t,\ x,\ u_\lambda(t)), \tag{11}$$

$$\dot{x}=\varphi(t,\ x,\ \hat{u}(t)). \tag{12}$$

Согласно (8) правые части этих уравнений совпадают при $t < \tau-\lambda$, а так как $x_\lambda(t_0)=\overline{x}_0=\widehat{x}(t_0)$, то по теореме единствености решения задачи Коши $x_\lambda(t)\equiv\widehat{x}(t)$ при $t < \tau-\lambda$ и по непрерывности

$$\xi(\lambda)=x_\lambda(\tau-\lambda)=\widehat{x}(\tau-\lambda). \tag{13}$$

В частности, $\xi(\lambda)$ непрерывио дифференцируема по λ и

$$\xi(0)=\widehat{x}(\tau),\ \xi'(0)=-\dot{\widehat{x}}(\tau)=-\varphi(\tau,\widehat{x}(\tau),\ u\ (\tau)). \tag{14}$$

Обозначим через $\Xi(t,\ s,\ \xi)$ решение задачи Коши для дифференциального уравнения с фиксированным управлеиием v:

$$\dot{x}=\varphi(t,\ x,\ v),\quad x(s)=\xi. \tag{15}$$

В соответствии с локальной теоремой существования и едиственности можно подобрать такие $\varepsilon_1 > 0$ и $\delta_1 > 0$, что $\Xi(t,\ s,\ \xi)$ определено при

$$|t-\tau| < \delta_1,\ |s-\tau| < \delta_1,\ |\xi-\widehat{x}(\tau)| < \varepsilon_1,$$

а в силу теоремы о зависимости решеиия от начальных даниых Ξ — иепрерывно дифференцируемая функция.

Согласно (8) и (13) для определения $x_\lambda(t)$ на отрезке $[\tau-\lambda,\ \tau]$ мы должны в (15) положить $\xi=\xi(\lambda)=\widehat{x}(\tau-\lambda)$ и если $\lambda_1 < \delta$ выбрано так, что $|\xi(\lambda)-\widehat{x}(\tau)| < \varepsilon_1$ при $0\leqslant\lambda\leqslant\lambda_1$, то

$$x_\lambda(t)=\Xi(t,\ \tau-\lambda,\ \xi(\lambda)),\ \tau-\lambda\leqslant t\leqslant\tau.$$

В частности,

$$\eta(\lambda)=x_\lambda(\tau)=\Xi(\tau,\ \tau-\lambda,\ \xi(\lambda)), \tag{16}$$

будучи суперпозицией непрерывио диффереицируемых функций, сама непрерывио диффереицируема по λ и

$$\begin{aligned}\eta(0)=\widehat{x}(\tau),\ \eta'(0)&=-\Xi_s(\tau,\ \tau,\ \xi(0))+\Xi_\xi(\tau,\ \tau,\ \xi(0))\,\xi'(0)=\\&=-\Xi_s(\tau,\ \tau,\ \xi(0))-\Xi_\xi(\tau,\ \tau,\ \xi(0))\,\varphi(\tau,\widehat{x}(\tau),\ \widehat{u}(\tau))\end{aligned} \tag{17}$$

в силу (14). Решение $\Xi(t,\ s,\ \xi)$ задачи Коши (15) удовлетворяет эквивалентному интегральному уравнеиию

$$\Xi(t,\ s,\ \xi)=\xi+\int_s^t\varphi(\sigma,\ \Xi(\sigma,\ s,\ \xi),\ v)\,d\sigma. \tag{18}$$

Дифференцируя его по s, имеем

$$\Xi_s(t,\ s,\ \xi)=$$
$$=-\varphi(s,\ \Xi(s,\ s,\ \xi),\ v)+\int_s^t\varphi_x(\sigma,\ \Xi(\sigma,\ s,\ \xi),v)\,\Xi_s(\sigma,s,\xi)\ d\sigma$$

и, полагая $t=s=\tau$, $\xi=\widehat{x}(\tau)$, получаем (с учетом очевидного тождества $\Xi(t,\ t,\ \xi)\equiv\xi$)

$$\Xi_s(\tau,\ \tau,\ \widehat{x}(\tau))=-\varphi(\tau,\ \widehat{x}(\tau),\ v). \tag{19}$$

Аналогично, дифференцируя (18) по ξ и подставляя те же значения аргументов, получаем

$$\Xi_{\xi}(\tau,\ \tau,\ \hat{x}(\tau))=E \tag{20}$$

(здесь $E=(\partial\xi_i/\partial\xi_k)=(\delta_{ik})$ — единичная матрица). Подставляя (19) и (20) в (17), имеем

$$\eta(0)=\hat{x}(\tau),\ \eta'(0)=\varphi(\tau,\ \hat{x}(\tau),\ v)-\varphi(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau)). \tag{21}$$

Далее, функция Ξ непрерывна в точке $(\tau,\ \tau,\ \hat{x}(\tau))$, причем $\Xi(\tau,\ \tau,\ \hat{x}(\tau))=\hat{x}(\tau)$, а $\hat{x}(\cdot)$ непрерывна в точке τ. Поэтому для любого $\varepsilon>0$ существует такое $\delta>0$, что при

$$|t-\tau|<\delta,\quad |s-\tau|<\delta,\quad |\xi-\hat{x}(\tau)|<\delta$$

выполняются неравенства

$$|\Xi(t,\ s,\ \xi)-\hat{x}(\tau)|<\varepsilon/2,\quad |\hat{x}(t)-\hat{x}(\tau)|<\varepsilon/2. \tag{22}$$

Возьмем положительное $\lambda_1\leqslant\delta$ столь малым, чтобы при $0\leqslant\lambda\leqslant\lambda_1$ выполнялось неравенство $|\xi(\lambda)-\hat{x}(\tau)|<\delta$. Тогда для $\tau-\lambda\leqslant t\leqslant\tau$, $s=\tau-\lambda$ и $\xi=\xi(\lambda)$ будут иметь место неравенства (22), откуда

$$|x_{\lambda}(t)-\hat{x}(t)|=|\Xi(t,\ \tau-\lambda,\ \xi(\lambda))-\hat{x}(t)|\leqslant$$
$$\leqslant|\Xi(t,\ \tau-\lambda,\ \xi(\lambda))-\hat{x}(\tau)|+|\hat{x}(\tau)-\hat{x}(t)|<\varepsilon/2+\varepsilon/2=\varepsilon.$$

Поскольку $x_{\lambda}(t)\equiv\hat{x}(t)$ при $t_0\leqslant t\leqslant\tau-\lambda$,

$$0\leqslant\lambda\leqslant\lambda_1\Rightarrow|x_{\lambda}(t)-\hat{x}(t)|<\varepsilon,\quad t_0\leqslant t\leqslant\tau,$$

чем доказано первое утверждение леммы.

Теперь обозначим через $X(\ ,\eta)$ решение задачи Коши для уравнения (12) с начальным условием $x(\tau)=\eta$. По теореме о зависимости решений от начальных данных существует такое ε_2, что $X(t,\ \eta)$ определено при $|\eta-\hat{x}(\tau)|<\varepsilon_2$, $\tau\leqslant t\leqslant t_1$ и является непрерывно дифференцируемой функцией. Согласно (8) и (16), а также в силу теоремы единственности $x_{\lambda}(t)\equiv X(t,\ \eta(\lambda))$. Снова, как суперпозиция непрерывно дифференцируемых функций, $x_{\lambda}(t)$ непрерывно дифференцируема по $(t,\ \lambda)$ при $\tau\leqslant t\leqslant t_1$ и $0\leqslant\lambda\leqslant\lambda_2$, где λ_2 выбрано так, чтобы при $0\leqslant\lambda\leqslant\lambda_2$ выполнялось неравенство $|\eta(\lambda)-\hat{x}(\tau)|<\varepsilon_2$. Полагая $\lambda_0=\min(\lambda_1,\ \lambda_2)$, мы видим, что верно второе утверждение леммы.

Переходя от уравнения (12) к эквивалентному интегральному уравнению, имеем с учетом (16)

$$x_{\lambda}(t)=\eta(\lambda)+\int_{\tau}^{t}\varphi(\sigma,\ x_{\lambda}(\sigma),\ \hat{u}(\sigma))\,d\sigma.$$

Дифференцируя это уравнение по λ и полагая затем $\lambda=0$ и обозначая, как и в условии леммы, $y(t)=\left.\frac{d}{d\lambda}x_{\lambda}(t)\right|_{\lambda=0}$, находим

$$y(t)=\eta'(0)+\int_{\tau}^{t}\varphi_x(\sigma,\ \hat{x}(\sigma),\ \hat{u}(\sigma))\,y(\sigma)\,d\sigma.$$

Это интегральное уравнение эквивалентно уравнению (9) с начальным условием $y(\tau)=\eta'(0)$, совпадающим с (10) ввиду (21).

Если управление $\hat{u}(\cdot)$ — кусочно-непрерывная функция, то поступаем следующим образом. Для простоты пусть точек разрыва две, скажем, α_1 и α_2, и τ (в которой $\hat{u}(\cdot)$ должно быть непрерывным) расположена между ними: $t_0<\alpha_1<\tau<\alpha_2<t_1$. В полосе $t_0\leqslant t\leqslant\alpha_1$ системы (11) и (12) (в которых при $t=\alpha_1$ нужно считать управление равным его предельному значению $\hat{u}(\alpha_1-0)=\lim\limits_{t\to\alpha_1-0}\hat{u}(t)$) совпадают и по теореме единственности $x_\lambda(t)\equiv\hat{x}(t)$.

Теперь переходим в полосу $\alpha_1\leqslant t\leqslant\alpha_2$ (снова полагая на ее границах управление равным предельным значениям $\hat{u}(\alpha_1+0)$ при $t=\alpha_1$ и $\hat{u}(\alpha_2-0)$ при $t=\alpha_2$). Здесь мы решаем уравнения (11) и (12) с начальным условием $x=\hat{x}(\alpha_1)$. Наши предыдущие утверждения применимы, и мы убеждаемся в непрерывной дифференцируемости $x_\lambda(t)$ по λ.

Наконец, в полосе $\alpha_2\leqslant t\leqslant t_1$ (с тем же соглашением о значении управления при $t=\alpha_2$) решаем наши дифференциальные уравнения с начальными условиями $x_\lambda(\alpha_2)$ и $\hat{x}(\alpha_2)$. Еще раз ссылаясь на теорему о зависимости решения от начальных данных, доказываем непрерывную дифференцируемость $x_\lambda(t)$ по λ при $\alpha_2\leqslant t\leqslant t_1$ и вычисляем $y(t)=\left.\dfrac{d}{d\lambda}x_\lambda(t)\right|_{\lambda=0}$ тем же способом, что и раньше. ■

Б) Лемма о приращении функционала. Положим $\chi(\lambda)=\mathcal{J}(x_\lambda(\cdot),\ u_\lambda(\cdot))$ и докажем, что эта функция дифференцируема справа в точке $\lambda=0$.

Пусть $\hat{p}(\cdot)$ — решение системы (6) *с краевым условием* (7). *Тогда*

$$\chi'(+0)=\left.\frac{d}{d\lambda}\mathcal{J}(x_\lambda(\cdot),\ u_\lambda(\cdot))\right|_{\lambda=+0}=a(\tau,\ v),$$

где

$$a(\tau,\ v)=f(\tau,\ \hat{x}(\tau),\ v)-f(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau))-$$
$$-\hat{p}(\tau)[\varphi(\tau,\ \hat{x}(\tau),\ v)-\varphi(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau))]. \quad (23)$$

Доказательство. Поскольку

$$\chi(\lambda)-\chi(0)=\int_{t_0}^{t_1}f(t,\ x_\lambda(t),\ u_\lambda(t))\,dt-\int_{t_0}^{t_1}f(t,\ \hat{x}(t),\ \hat{u}(t))\,dt=$$
$$=\int_{\tau}^{t_1}[f(t,\ x_\lambda(t),\ \hat{u}(t))-f(t,\ \hat{x}(t),\ \hat{u}(t))]\,dt+$$
$$+\int_{\tau-\lambda}^{\tau}[f(t,\ x_\lambda(t),\ v)-f(t,\ \hat{x}(t),\ \hat{u}(t)]\,dt,$$

имеем

$$\chi'(0)=\lim_{\lambda\downarrow 0}\frac{\chi(\lambda)-\chi(0)}{\lambda}=\frac{d}{d\lambda}\int_{\tau}^{t_1} f\,(t,\ x_\lambda(t),\ \hat{u}(t))\,dt\,\Big|_{\lambda=+0}+$$

$$+\lim_{\lambda\downarrow 0}\frac{1}{\lambda}\int_{\tau-\lambda}^{\tau}[f(t,\ x_\lambda(t),\ v)-f(t,\ \hat{x}(t),\ \hat{u}(t)]\,dt.$$

Поскольку по лемме п. А) $x_\lambda(t)$ непрерывно дифференцируема по λ, в первом члене применимо обычное правило дифференцирования под знаком интеграла, а во втором воспользуемся теоремой о среднем, после чего

$$\chi'(+0)=\int_{\tau}^{t_1} f_x(t,\ \hat{x}(t),\ \hat{u}(t))\frac{d}{d\lambda}x_\lambda(t)\Big|_{\lambda=0}dt+$$

$$+\lim_{\lambda\downarrow 0}[f(c,\ x_\lambda(c),\ v)-f(c,\ \hat{x}(c),\ \hat{u}(c)]=$$

$$=\int_{\tau}^{t_1} f_x(t,\ \hat{x}(t),\ \hat{u}(t))\,y(t)\,dt+f(\tau,\ \hat{x}(\tau),\ v)-f(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau)) \tag{24}$$

(здесь мы воспользовались тем, что $\tau-\lambda\leqslant c\leqslant\tau$ и потому $c\to\tau$; $x_\lambda(c)\to\hat{x}(\tau)$ в силу первого утверждения леммы п. А); $y(t)$ обозначает то же, что и в этой лемме.

Далее, $\hat{p}(\cdot)$ удовлетворяет системе (6), а $y(\cdot)$ удовлетворяет системе (9). Поэтому

$$\frac{d}{dt}\hat{p}(t)\,y(t)=\dot{\hat{p}}(t)\,y(t)+\hat{p}(t)\,\dot{y}(t)=$$

$$=-\hat{p}(t)\hat{\varphi}_x(t)\,y(t)+\hat{f}_x(t)\,y(t)+\hat{p}(t)\,\hat{\varphi}_x(t)\,y(t)=\hat{f}_x(t)\,y(t).$$

Интегрируя это равенство в пределах от τ до t_1 и учитывая краевое условие (7) для $\hat{p}(\cdot)$ и условие (10) для $y(\cdot)$, получаем

$$\int_{\tau}^{t_1}\hat{f}_x(t)\,y(t)\,dt=\int_{\tau}^{t_1}\frac{d}{dt}[\hat{p}(t)\,y(t)]\,dt=\hat{p}(t)\,y(t)\Big|_{\tau}^{t_1}=$$

$$=-\hat{p}(\tau)\,[\varphi(\tau,\ \hat{x}(\tau),\ v)-\varphi(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau))]. \tag{25}$$

Сопоставив (24), (25) и (23), получаем, что $\chi'(+0)=a(\tau,\ v)$. ■

В) Завершение доказательства. Если правый конец свободен, то в силу леммы п. А) всякая элементарная вариация допустима (при достаточно малом λ). Значит, если $(\hat{x}(\cdot),\ \hat{u}(\cdot))$ — оптимальный процесс, то при малых λ

$$\mathcal{J}(x_\lambda(\cdot),\ u_\lambda(\cdot))\geqslant\mathcal{J}(\hat{x}(\cdot),\ \hat{u}(\cdot))\Leftrightarrow\chi(\lambda)\geqslant\chi(0)\Rightarrow\chi'(+0)\geqslant 0.$$

Применив лемму п. Б) получим, что условие $a(\tau, v) \geqslant 0$ необходимо для оптимальности $(\hat{x}(\cdot), \hat{u}(\cdot))$. Но $\tau \in T_0$ и $v \in \mathfrak{U}$ были произвольны. Мы доказали, следовательно, что для любого t, принадлежащего множеству точек непрерывности управления $\hat{u}(\cdot)$, и для любого $u \in \mathfrak{U}$ выполнено неравенство

$$\hat{p}(t)\varphi(t, \hat{x}(t), u) - f(t, \hat{x}(t), u) \leqslant \\ \leqslant \hat{p}(t)\varphi(t, \hat{x}(t), \hat{u}(t)) - f(t, \hat{x}(t), \hat{u}(t)),$$

равносильное неравенству $a(t, u) \geqslant 0$ и (11). Теорема полностью доказана.

§ 1.6. Решение задач

Задачи, о которых говорилось в начале главы,— задачи, поставленные с разными целями и в разные времена,— здесь будут рассмотрены единообразно, по одной стандартной схеме, диктуемой принципом Лагранжа.

Эта схема состоит из шести этапов:

1. Запись формализованной задачи и обсуждение проблемы существования и единственности решения.
2. Составление функции Лагранжа.
3. Применение принципа Лагранжа.
4. Исследование возможности $\hat{\lambda}_0 = 0$.
5. Нахождение стационарных точек, т. е. решение уравнений, вытекающих из принципа Лагранжа.
6. Исследование стационарных точек, выбор решения и запись ответа.

Во всех задачах, о которых пойдет речь, принцип Лагранжа является строго обоснованной теоремой либо уже доказанной, либо доказываемой в последующих главах. Применимость единой схемы к задачам столь разного содержания подчеркивает универсальность этого принципа. Разумеется, рассматривая другие задачи, исследователи могут столкнуться с ситуациями, которые не подходят ни под одну из известных конкретных схем (классическое вариационное исчисление, оптимальное управление, линейное программирование и т. п.). Тем не менее и здесь принцип Лагранжа в том или ином его понимании может оказаться верным или, по крайней мере, может быть полезным ориентиром. Понимание общих идей и ситуаций, в которых он применим,— о них речь пойдет в гл. III и IV — может помочь найти необходимые условия экстремума и в измененной обстановке. Впрочем, не менее важно отдавать себе отчет и в том, что принцип Лагранжа верен не всегда, и потому опасно применять его бездумно.

1.6.1. Геометрические экстремальные задачи. Здесь решены все задачи, поставленные в п. 1.1.2 и формализованные в п. 1.2.2. Первый этап предложенной схемы предполагает обсуждение вопроса о существовании решения. Геометрические задачи этого пункта конечномерны, и существование решения в них обеспечивает следующая теорема.

Теорема Вейерштрасса. *Пусть функция* $f\colon \mathbf{R}^N \to \mathbf{R}$ *непрерывна и для некоторого* α *множество* $\mathscr{L}_\alpha f = \{x \mid f(x) \leqslant \alpha\}$ *непусто и ограничено. Тогда решение задачи* $f(x) \to \inf$ *существует.*

Доказательство этой теоремы очевидным образом следует из классической теоремы Вейерштрасса о существовании минимума непрерывной функции на ограниченном и замкнутом подмножестве в $\mathbf{R}^N$ [14, т. 1, с. 176, 370], [9, т. 1; с. 234], ибо множество $\mathscr{L}_\alpha f$, очевидно, замкнуто.

Переходим к решению задач.

Задача Евклида о вписанном параллелограмме. Эта задача была формализована так (см. (1) п. 1.2.2):

1. $f_0(x) = x(x-b) \to \inf,\ 0 \leqslant x \leqslant b.$

Мы опустили несущественный множитель H/b и свели задачу к задаче минимизации. Функция f_0 непрерывна, отрезок $[0, b]$ ограничен и замкнут. По теореме Вейерштрасса решение задачи существует. Пусть это решение $\hat{x}$. Ясно, что $\hat{x} \neq 0$ и $\hat{x} \neq b$, ибо $f_0(0) = f_0(b) = 0$, а функция принимает и отрицательные значения. Значит, $\hat{x} \in (0, b)$. Функция f_0 — гладкая. Поэтому надо искать стационарные точки в задаче $f_0(x) \to \inf$.

2—5. $f_0'(\hat{x}) = 0 \Rightarrow \hat{x} = b/2.$

6. В силу единственности стационарной точки $\hat{x} = b/2 \in [0, b]$ есть решение задачи, т. е. искомый параллелограмм $A\hat{D}\hat{E}\hat{F}$ характеризуется тем, что $|A\hat{F}| = |AC|/2$, т. е. $\hat{F}$ *есть середина отрезка* $[A, C]$. Этот факт и был установлен в п. 1.1.2 геометрическим путем.

Примечание. Наша задача оказалась элементарной гладкой задачей. Поэтому функция Лагранжа не выписывалась и пп. 2—5 «слились». Была использована теорема Ферма (п. 1 3.1).

Задача Архимеда об изопифанных сегментах шаров (п. 1.2.1). Решение здесь совершенно аналогично задаче Евклида и поэтому приводится без комментариев.

1. $f_0(h)=ha/2-\pi h^3/3\to\sup;\ 0\leqslant h\leqslant\sqrt{a/\pi}$.

2—5. $f_0'(\hat h)=0\Rightarrow\hat h=\sqrt{a/2\pi}$.

6. Значение f_0 в нуле равно нулю, а в точке $\sqrt{a/\pi}$ меньше, чем значение в стационарной точке $\sqrt{a/2\pi}$. Значит, $\hat h=\sqrt{a/2\pi}$ есть решение задачи. Вспомнив, что $a=2\pi R\hat h$, получаем $\hat h=R$, т. е. *искомый шаровой сегмент является полушаром* (высота равна радиусу).

Задача Аполлония о кратчайшем расстоянии от точки до эллипса. Она была формализована так (см. (6) п. 1.2.2):

1. $f_0(x_1,x_2)=(x_1-\xi_1)^2+(x_2-\xi_2)^2\to\inf;\quad f_1(x_1,x_2)=(x_1/a_1)^2+(x_2/a_2)^2=1$. Эллипс—ограниченное и замкнутое множество, функция f_0 непрерывна, значит, решение $\hat x=(\hat x_1,\hat x_2)$ по теореме Вейерштрасса существует. Функции f_0 и f_1—гладкие.

2. $\mathscr{L}=\lambda_0((x_1-\xi_1)^2+(x_2-\xi_2)^2)+\lambda_1((x_1/a_1)^2+(x_2/a_2)^2-1)$.

3. $$\mathscr{L}_{x_1}=0\Rightarrow\hat\lambda_0(\hat x_1-\xi_1)+\hat\lambda_1\hat x_1/a_1^2=0,$$
$$\mathscr{L}_{x_2}=0\Rightarrow\hat\lambda_0(\hat x_2-\xi_2)+\hat\lambda_1\hat x_2/a_2^2=0.$$

4. Пусть $\hat\lambda_0=0$. Тогда $\hat\lambda_1\neq0$ (множители Лагранжа не равны нулю одновременно!). Значит, из 3-го этапа следует, что $\hat x_1=\hat x_2=0\Rightarrow f_1(0,\ 0)=f_1(\hat x_1,\ \hat x_2)=0\neq1$. Итак, $\hat\lambda_0\neq0$, и можно положить $\hat\lambda_0=1$. Обозначим при этом $\hat\lambda_1=\lambda$.

5. (Для простоты ограничиваемся случаем, когда $\xi_1\xi_2\neq0$).

$$(\hat x_i-\xi_i)+\lambda\hat x_i/a_i^2=0,\quad i=1,2\Rightarrow\hat x_i=\frac{\xi_ia_i^2}{(a_i^2+\lambda)},\quad i=1,2.$$

Подставляя в уравнения эллипса, получаем

$$\varphi(\lambda)=\frac{\xi_1^2a_1^2}{(a_1^2+\lambda)^2}+\frac{\xi_2^2a_2^2}{(a_2^2+\lambda)^2}=1.\qquad(1)$$

6. Число стационарных точек задачи (т. е. точек, соответствующих тем λ, которые удовлетворяют уравнению (1)) не больше четырех (см. рис. 26, неравенство $\varphi(0)=\xi_1^2/a_1^2+\xi_2^2/a_2^2>1$ показывает, что $\varphi(\lambda)$ изображена для точки (ξ_1, ξ_2), лежащей вне эллипса). Для полного решения задачи надо решить уравнение (1), получить λ_i, найти соответствующие точки $x(\lambda_i)$, подставить эти точки в f_0 и найти наименьшее из полученных чисел.

Примечания. 1. Задача 1 — это гладкая задача с ограничением типа равенства. В пп. 2—5 применялось правило множителей Лагранжа (п. 1.3.2).

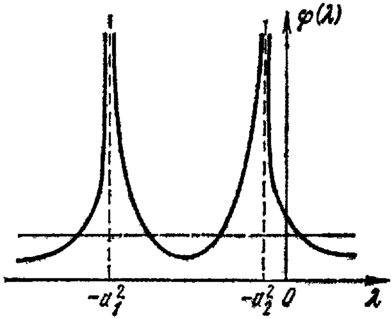


Рис. 26.

2. Соотношения $(\hat{x}_i-\xi_i)+\lambda\hat{x}_i/a_i^2=0$ имеют очевидный геометрический смысл: вектор $\xi-\hat{x}$, соединяющий точку ξ с минимальной точкой $\hat{x}$, пропорционален вектору-градиенту функции f_1 в точке $\hat{x}$, т. е. вектор $\xi-\hat{x}$ лежит на нормали к эллипсу. Этот факт был установлен впервые Аполлонием.

3. Выведем из полученных нами соотношений уравнение кривой, «разделяющей» те точки ξ, к которым можно провести две нормали, от точек, к которым можно провести четыре нормали. Очевидно, что это разделение происходит для λ, удовлетворяющих соотношению (1), для которых

$$\varphi'(\lambda)=-\frac{\xi_1^2a_1^2}{(a_1^2+\lambda)^3}-\frac{\xi_2^2a_2^2}{(a_2^2+\lambda)^3}=0, \quad \lambda\in(-a_1^2,\ -a_2^2). \qquad (2)$$

Из (2) имеем

$$a_1^2+\lambda=A(\xi_1a_1)^{2/3}, \quad a_2^2+\lambda=-A(\xi_2a_2)^{2/3},$$

где

$$A=(a_1^2-a_2^2)/((\xi_1a_1)^{2/3}+(\xi_2a_2)^{2/3}).$$

Подставляя в (1), получаем уравнение разделяющей кривой

$$(\xi_1a_1)^{2/3}+(\xi_2a_2)^{2/3}=(a_2^2-a_1^2)^{2/3}.$$

Это — уравнение *астроиды* (см. рис. 7 в § 1.1). Вне астроиды каждая точка имеет две нормали, внутри нее — четыре, на самой астроиде — три (за исключением вершин, где имеется две нормали). Описанный здесь результат также был получен Аполлонием в его «Конике».

Задача Кеплера о вписанном цилиндре (см. (3) в п. 1.2.2). Решение этой задачи подобно решению задачи Евклида, и мы его не комментируем.

$$1. \qquad f_0(x) = x(x^2-1) \to \inf, \quad 0 \leqslant x \leqslant 1.$$

$$2—5. \qquad f_0'(\hat{x}) = 0 \Rightarrow 3\hat{x}^2 = 1 \Rightarrow \hat{x} = \sqrt{3}/3.$$

6. В силу единственности стационарной точки в $(0, 1)$, $\hat{x} = \sqrt{3}/3$ есть решение задачи: *искомый цилиндр характеризуется тем, что отношение его высоты* $2\hat{x}$ *к радиусу* $\sqrt{1-\hat{x}^2}$ *равно* $\sqrt{2}$.

Задача о преломлении света. Эта задача была поставлена и решена методом Гюйгенса в п. 1.1.3. Здесь мы даем стандартное решение ее, восходящее к Лейбницу (см. (4) в п. 1.2.2).

$$1. \quad f_0(x) = \sqrt{\alpha^2+x^2}/v_1 + \sqrt{\beta^2+(\xi-x)^2}/v_2 \to \inf.$$

Все лебеговы множества $\mathscr{L}_a f_0$ непрерывной функции f_0 компактны, и, значит, по теореме Вейерштрасса решение задачи существует.

$$2—5. \quad f'(\hat{x}) = 0 \Rightarrow \frac{\hat{x}}{v_1\sqrt{\alpha^2+\hat{x}^2}} = \frac{\xi-\hat{x}}{v_2\sqrt{\beta^2+(\xi-\hat{x})^2}} \Leftrightarrow$$

$$\Leftrightarrow \frac{\sin\varphi_1}{v_1} = \frac{\sin\varphi_2}{v_2}$$

(см. рис. 19 в п. 1.2.2).

6. Точка $\hat{x}$, удовлетворяющая последнему уравнению, единственна (проверьте!), значит, она и является решением задачи. Итак, точка преломления луча света на границе двух сред характеризуется тем, что *отношение синуса угла падения к синусу угла преломления равно отношению скоростей распространения света в соответствующих средах*. Это и есть *закон Снеллиуса*.

Задача Штейнера. Она была формализована так (см. (5) в п. 1.2.2):

$$1. \; f_0(x) = |x-\xi_1| + |x-\xi_2| + |x-\xi_3| \to \inf, \quad x \in \mathbf{R}^2,$$
$$\xi_i \in \mathbf{R}^2, \quad i = 1, 2, 3, \quad |x| = \sqrt{x_1^2+x_2^2}.$$

Решение задачи существует по теореме Вейерштрасса (проверьте!). Возможно одно из двух: либо решение совпадает с одной из точек ξ_i, $i=1, 2, 3$, либо не совпадает ни с одной. Будем решать задачу в последнем случае. Тогда функция будет гладкой в окрестности точки $\hat{x}$ (проверьте!).

2—5. $f_0'(\hat{x})=0 \Rightarrow \dfrac{\xi_1-\hat{x}}{|\hat{x}-\xi_1|}+\dfrac{\xi_2-\hat{x}}{|\hat{x}-\xi_2|}+\dfrac{\xi_3-\hat{x}}{|\hat{x}-\xi_3|}=0.$ (3)

6. Уравнение (3) означает, что три единичных вектора, смотрящие из $\hat{x}$ в сторону ξ_1, ξ_2, ξ_3 соответственно, в сумме равны нулю. Значит, эти векторы параллельны сторонам равностороннего треугольника, и, следовательно, величины углов $\xi_1\hat{x}\xi_2$, $\xi_2\hat{x}\xi_3$, $\xi_3\hat{x}\xi_1$, под которыми из точки $\hat{x}$ видны стороны треугольника, равны 120°. Таким образом, точка $\hat{x}$ является *точкой Торичелли*. Она может быть найдена как точка пересечения двух дуг окружностей, стягиваемых хордами $[\xi_1, \xi_2]$ и $[\xi_1, \xi_3]$ и вмещающих углы 120°. Это построение возможно, если только ни один из углов треугольника не больше 120°. В противном случае дуги не пересекаются, и, значит, невозможно, чтобы точка $\hat{x}$ не совпадала ни с одной из точек ξ_i, $i=1, 2, 3$. Но тогда она должна, очевидно, совпадать с вершиной тупого угла, ибо против тупого угла лежит самая большая сторона треугольника.

Ответ: *Искомая точка есть точка Торичелли, если все величины углов треугольника меньше* 120° *и есть вершина тупого угла в остальных случаях.*

Итак, все геометрические задачи, о которых говорилось в п. 1.1.2, а также задача о преломлении света из п. 1.1.3 решены. Заодно решим и задачу Тартальи:

1. $f_0(x)=x(8-x)(8-2x)\to\sup,\quad 0\leqslant x\leqslant 4.$

2—5. $f_0'(\hat{x})=0\Rightarrow 3\hat{x}^2-24\hat{x}+32=0\Rightarrow\hat{x}=4-4/\sqrt{3}.$

6. Ответ: *Одно число равно* $4-4/\sqrt{3}$, другое $4+4/\sqrt{3}$.

1.6.2. Аэродинамическая задача Ньютона. Эта задача была поставлена в п. 1.1.5 и формализована в п. 1.2.3.

1. $\displaystyle\int_0^T \frac{t\,dt}{1+u^2}\to\inf,\quad \dot{x}=u,\quad x(0)=0,\quad x(T)=\xi,\quad u\in\mathbb{R}_+$

Для такого рода задач доказать непосредственно теорему существования не совсем просто. Главная беда — невыпуклость интегранта $t/(1+u^2)$ по u при $u \geqslant 0$. Но тем не менее эту задачу мы решим до конца. Наш дальнейший план таков. Предположив, что решение задачи существует, мы применим к гипотетическому решению принцип Лагранжа. Выяснив, что существует единственная допустимая в задаче стационарная кривая (т. е. допустимая кривая, для которой выполняются все условия, диктуемые принципом Лагранжа), непосредственной выкладкой убедимся, что именно она доставляет абсолютный миимум. В заключение вспомним слова Ньютона и убедимся, что найденное решение в точности то, которое было им описано в 1687 г.

2. $$\mathscr{L} = \int_0^T L\,dt + \mu_0 x(0) + \mu_1 (x(T) - \xi),$$

$$L = \frac{\lambda_0 t}{1+u^2} + p(\dot{x} - u).$$

3. Уравнение Эйлера:

$$-\frac{d}{dt} L_{\dot{x}} + L_x = 0 \Rightarrow \hat{p}(t) = \text{const} = p_0. \qquad (1)$$

Условие трансверсальности:

$$p_0 = \hat{\mu}_0 = -\hat{\mu}_1. \qquad (2)$$

Условие минимальности по u:

$$\frac{\hat{\lambda}_0 t}{1+u^2} - p_0 u \geqslant \frac{\hat{\lambda}_0 t}{1+\hat{u}^2(t)} - p_0 \hat{u}(t), \quad \forall u \geqslant 0. \qquad (3)$$

4. Если допустить, что $\hat{\lambda}_0 = 0$, то необходимо, чтобы $p_0 \neq 0$ (ибо иначе из (2) вытекали бы равенства $\hat{\lambda}_0 = p_0 = \hat{\mu}_0 = \hat{\mu}_1 = 0$, но все множители Лагранжа не могут обратиться в нуль одновременно). Если же $\hat{\lambda}_0 = 0$ и $p_0 \neq 0$, то из (3) следует, что $\hat{u}(t) \equiv 0$, а значит, $\hat{x}(t) = \int_0^t \hat{u}(t)\,d\tau \equiv 0$.

Но тогда искомое тело не имеет «длины» — оно является плоской пластиной. Если же $\xi > 0$, то $\hat{\lambda}_0 \neq 0$ и можно считать, что $\hat{\lambda}_0 = 1$. Отметим еще, что случай $p_0 \geqslant 0$ также невозможен, ибо при этом функция

$t/(1+u^2)-p_0u$ монотонно убывает и (3) не выполняется при $u > \hat{u}(t)$.

5. Из (3) (с $\hat{\lambda}_0=1$) следует, что до некоторого момента оптимальное управление равно нулю (проверьте, что при $p_0<0$ и малых t функция $u \rightarrow (t/(1+u^2))-p_0u$ достигает минимума при $u=0$). Затем оптимальное управление $\hat{u}(\cdot)$ должно быть найдено из уравнения

$$-p_0=\frac{2ut}{(1+u^2)^2}, \tag{4}$$

получающегося из уравнения $L_u=0$. Момент излома τ определяется тем, что функция $u \longrightarrow (\tau/(1+u^2))-p_0u$ имеет два равных минимума — в нуле и в точке, определяемой из (4) при $t=\tau$. Иначе говоря, в момент излома должны удовлетворяться соотношения (далее через $\hat{u}(\tau)$ обозначено $\hat{u}(\tau+0)\neq 0$):

$$-p_0=\frac{2\hat{u}(\tau)\,\tau}{(1+\hat{u}^2(\tau))^2}, \quad \frac{\tau}{1+\hat{u}^2(\tau)}-p_0\hat{u}(\tau)=\tau. \tag{5}$$

Из второго уравнения получаем: $-\hat{u}^2(\tau)\,\tau/(1+\hat{u}^2(\tau))=$ $=p_0\hat{u}(\tau)$, откуда $p_0=-\tau\hat{u}(\tau)/(1+\hat{u}^2(\tau))$. Подставив это соотношение в первое уравнение (5), находим, что $\hat{u}^2(\tau)=1 \Rightarrow \hat{u}(\tau)=1$ (ибо $\hat{u}\geqslant 0$) и тогда снова из первого уравнения (5) получаем равенство $\tau=-2p_0$.

После излома оптимальное решение удовлетворяет соотношению (4), из которого следует, что

$$t=-\frac{p_0(1+u^2)^2}{2u}=-\frac{p_0}{2}\left(\frac{1}{u}+2u+u^3\right). \tag{6}$$

Но

$$\frac{dx}{dt}=u \Rightarrow \frac{dx}{du}=\frac{dx}{dt}\frac{dt}{du}=u\frac{dt}{du}=-\frac{p_0}{2}\left(-\frac{1}{u}+2u+3u^3\right).$$

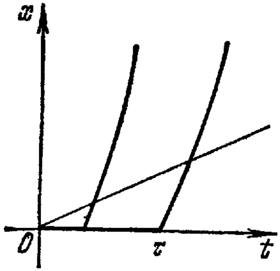
Интегрируя это соотношение с учетом равенства $\hat{x}(\tau)=0$, $\hat{u}(\tau)=1$, получаем параметрические уравнения искомой оптимальной кривой

$$\begin{gathered}\hat{x}(t,\ p_0)=-\frac{p_0}{2}\left(\ln\frac{1}{u}+u^2+\frac{3}{4}u^4\right)+\frac{7}{8}p_0,\\ t=-\frac{p_0}{2}\left(\frac{1}{u}+2u+u^3\right), \quad p_0<0.\end{gathered} \tag{7}$$

6. Кривую (7) называют *кривой Ньютона.* При этом в (7) $u \in [1, \infty)$. Нетрудно понять, что пересечение прямой $x = \alpha t$ с кривой Ньютона, соответствующей параметру $p_0 = -1$, единственно. Действительно, $x(\cdot, -1)$ непрерывна и выпукла, ибо

$$\frac{d^2x}{dt^2} = \frac{du}{dt} = \frac{1}{dt/du} \geqslant 0.$$

Но, с другой стороны, из формул (7) видно, что кривая Ньютона $\hat{x}(\cdot, p_0)$ получается из кривой $\hat{x}(\cdot, -1)$ гомотетией с центром в (0, 0) и коэффициентом $|p_0|$ (рис. 27). Значит, для того чтобы провести кривую семейства (7) через заданную точку (ξ, T), нужно найти точку пересечения прямой $x = \xi t/T$ с кривой $\hat{x}(\cdot, -1)$ и затем сделать соответствующую гомотетию кривой $\hat{x}(\cdot, -1)$. Получим допустимую кривую $\hat{x}(\cdot)$. Убедимся, что она дает абсолютный минимум в задаче. Для этого вернемся к соотношению (3) с $\hat{\lambda}_0 = 1$. Пусть $x(\cdot)$ — любая допустимая кривая (т. е. $x(\cdot) \in KC^1([0, T])$, $x(0) = 0$, $x(T) = \xi$). Тогда в силу (3)



Рис, 27.

$$\frac{t}{1 + \dot{x}^2(t)} - p_0\dot{x}(t) \geqslant \frac{t}{1 + \hat{u}(t)^2} - p_0\hat{u}(t).$$

Проинтегрировав это соотношение и учитывая, что $\hat{u}(t) \doteq \dot{\hat{x}}(t)$ и $\int\limits_0^T \dot{x}(t)\,dt = \int\limits_0^T \dot{\hat{x}}(t)\,dt = \xi$, получаем

$$\int\limits_0^T \frac{t\,dt}{1 + \dot{x}^2} \geqslant \int\limits_0^T \frac{t\,dt}{1 + \dot{\hat{x}}^2}.$$

Задача полностью решена.

Примечания. 1. В п. 2—5 применялся принцип Лагранжа для задачи оптимального уравнения, сводящийся к принципу максимума Понтрягина (п. 1.5.3).

2. Сопоставим теперь полученное решение с решением, описанным самим Ньютоном. Вспомним слова Ньютона, приведенные в п. 1.1.5 и посмотрим еще раз

на его чертеж (рис. 18). При этом наряду с буквами, расставленными самим Ньютоном, используем еще и некоторые наши обозначения. Имеем: $|MN|=t$, $|BM|=x$, $|BG|=\tau$, угол $BGP=\varphi$, и тогда из построения Ньютона получаем, что

$$\operatorname{tg}\varphi=\dot{x}(t), \quad |BP|/|BG|=\operatorname{tg}\varphi\Rightarrow|BP|=\tau\dot{x},$$
$$|GP|^2=|BG|^2+|BP|^2=(\dot{x}^2+1)\tau^2.$$

Теперь из пропорции Ньютона

$$|MN|:|GP|=|GP|^3:4|BP|\times|GB|^2,$$

подставляя наши обозначения, получаем

$$\frac{t}{(\dot{x}^2+1)^{1/2}\tau}=\frac{\tau^3(\dot{x}^2+1)^{3/2}}{4\tau\dot{x}\tau^2}\Leftrightarrow\frac{\dot{x}t}{(\dot{x}^2+1)^2}=\frac{\tau}{4}. \tag{8}$$

Но это — не что иное, как соотношение (4), в которое подставлено значение $p_0=-\tau/2$. Из (8), рассуждая так же, как и раньше, мы находим интегрированием выражение (7) для кривой Ньютона. Отметим еще, что «затупленность» кривой и условие на скачок в точке $G\Leftrightarrow\tau$ (угол там равен 135°) были по существу предусмотрены Ньютоном в его «Поучении» об усеченном конусе.

Таким образом, задача Ньютона была им решена полностью, но смысл его решения оказался недоступным ни для его современников, ни для многих его последователей — вплоть до нашего времени.

1.6.3. Простейшая задача о быстродействии. Поставленная в п. 1.1.7, эта задача была формализована следующим образом:

$$\begin{gathered}T\to\inf,\quad m\ddot{x}=u,\quad u\in[u_1,\ u_2],\\ x(0)=x_0,\quad \dot{x}(0)=v_0,\quad x(T)=\dot{x}(T)=0\end{gathered} \tag{1}$$

(см. (5) в п. 1.2.4). Случай $u_1=u_2$ интереса не представляет, поскольку при этом не для всякой пары $(x_0,\ v_0)$ существует функция $x(\cdot)$, удовлетворяющая всем ограничениям (1), а если такая функция существует и отлична от тождественного нуля, то значение T определяется однозначно, так что задачи минимизации в сущности и нет.

При $u_1<u_2$ мы можем уменьшить число параметров задачи при помощи замены $x(t)=A\xi(t)+B(t-T)^2$. В терминах функции $\xi(\cdot)$ общий вид задачи (1) не меняется, но параметры x_0, v_0, u_1, u_2 приобретают другие

значения. В частности, если положить

$$A=(u_1-u_2)/2m,\quad B=(u_1+u_2)/4m,$$

то $\ddot{\xi}\in[-1,\ 1]$. Имея это в виду, будем считать далее в (1) $m=1$, $u_1=-1$, $u_2=+1$. Кроме того, обозначим $x=x_1$, $\dot{x}=x_2$.

Теперь приступаем к реализации нашей стандартной схемы:

$$1.\ T=\int_0^T 1\cdot dt\to\inf,\quad \dot{x}_1=x_2,\quad \dot{x}_2=u,\quad u\in[-1,\ 1],$$
$$x_1(0)=x_0,\quad x_2(0)=v_0,\quad x_1(T)=x_2(T)=0. \tag{2}$$

С существованием решения мы поступим точно так же, как и в предыдущем пункте: найдя из принципа Лагранжа функцию $x(\cdot)$, «подозреваемую» на оптимальность, непосредственной проверкой убедимся в том, что она дает нам решение задачи.

$$2.\ \mathscr{L}=\int_0^T L\,dt+\mu_1(x_1(0)-x_0)+\mu_2(x_2(0)-v_0)+$$
$$+\nu_1x_1(T)+\nu_2x_2(T), \tag{3}$$

где $L=\lambda_0+p_1(\dot{x}_1-x_2)+p_2(\dot{x}_2-u)$.

3. Уравнения Эйлера—Лагранжа:

$$-\frac{d}{dt}L_{\dot{x}_i}+L_{x_i}=0,\quad i=1,\ 2\Rightarrow\frac{d\hat{p}_1}{dt}=0,\quad \frac{d\hat{p}_2}{dt}=-\hat{p}_1. \tag{4}$$

Условия трансверсальности:

$$\hat{p}_1(0)=\hat{\mu}_1,\quad \hat{p}_2(0)=\hat{\mu}_2,\quad \hat{p}_1(T)=-\hat{\nu}_1,\quad \hat{p}_2(T)=-\hat{\nu}_2. \tag{5}$$

Принцип максимума: опустив не зависящие от u слагаемые, можно записать это условие в виде

$$\hat{p}_2(t)\hat{u}(t)=\max_{-1\leqslant u\leqslant 1}\{\hat{p}_2(t)u\}=|\hat{p}_2(t)|,$$

или

$$\hat{u}(t)=\begin{cases}\operatorname{sign}\hat{p}_2(t), & \text{если } \hat{p}_2(t)\neq 0,\\ \text{любое из } [-1,\ 1], & \text{если } \hat{p}_2(t)=0.\end{cases} \tag{6}$$

Кроме того, в рассматриваемой задаче переменным является также конечный момент времени T (ведь именно его мы и минимизируем), так что формально эта задача не укладывается в рамки п. 1.5.3. Можно показать, и это будет сделано в гл. IV, что в такой ситуации к уравнениям (4)—(6) следует добавить еще и условие $\mathscr{L}_T=0$ (стационарность функции Лагранжа по T), вполне согласующееся с общей идеологией принципа Лагранжа. Дифференцируя (3) по T и учитывая равенства $\dot{x}_2(T)=u(T)$, $\dot{x}_1(T)=x_2(T)=0$, получаем

$$\mathscr{L}_T=\hat{\lambda}_0+\hat{\nu}_2\hat{u}(T)=0. \tag{7}$$

4. Обращается ли $\hat{\lambda}_0$ в нуль или нет—специальной роли в данной задаче не играет, поскольку $\hat{\lambda}_0$ не входит в (4)—(6).

5. Из уравнений (4) заключаем, что $\hat{p}_1(t)\equiv\text{const}$, а $\hat{p}_2(\cdot)$ — произвольная линейная функция. При этом $\hat{p}_2(t)\not\equiv 0$, ибо

$$\hat{p}_2(t)\equiv 0\overset{(4)}{\Longrightarrow}\hat{p}_1(t)\equiv 0\overset{(5)}{\Longrightarrow}\hat{\mu}_1=\hat{\mu}_2=\hat{\nu}_1=\hat{\nu}_2=0\overset{(7)}{\Longrightarrow}\hat{\lambda}_0=0$$

и все множители Лагранжа оказываются нулями.

Линейная функция, отличная от тождественного нуля, обращается в нуль на отрезке $[0,T]$ не более одного раза. Поэтому из (6) получаем следующие возможности для оптимального управления:

а) $$\hat{u}(t)\equiv 1;$$

б) $$\hat{u}(t)\equiv -1$$

($\hat{p}_2(\cdot)$ не обращается в нуль на $[0,T]$),

в) $$\hat{u}(t)=\begin{cases}1, & 0\leqslant t<\tau,\\ -1, & \tau<t\leqslant T,\end{cases}$$

г) $$\hat{u}(t)=\begin{cases}-1, & 0\leqslant t<\tau,\\ 1, & \tau<t\leqslant T.\end{cases}$$

($\hat{p}_2(\cdot)$ обращается в нуль в точке $t=\tau$, значение управления в точке τ несущественно, так как изменение его в одной точке не оказывает влияния на функцию $\hat{x}(\cdot)$ (почему?); по той же причине можно опустить случай $\tau=T$ и $\tau=0$.)

Дальнейшие рассмотрения удобно вести на фазовой плоскости (x_1, x_2) (рис. 28). Решая задачу Коши

$$\dot{x}_1=x_2,\quad \dot{x}_2=\hat{u}(t),\quad x_1(T)=x_2(T)=0 \tag{8}$$

для одного из управлений а)—г), получаем единственное решение, а с ним и единственную начальную точку $(x_0, v_0)=(x_1(0), x_2(0))$, которая этому решению соответствует. Нетрудно проверить, что при всевозможных τ и T эти начальные точки однозначно покрывают всю плоскость. Прежде всего,

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = 1 \Rightarrow x_1 = x_2^2/2 + C_1, \tag{9}$$

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -1 \Rightarrow x_1 = -x_2^2/2 + C_2, \tag{10}$$

и, таким образом, фазовые траектории на участках постоянства управления лежат на параболах одного из семейств (9) или (10).

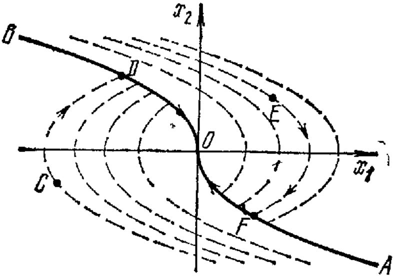


Рис. 28.

Управлению а) соответствуют начальные точки, лежащие на дуге OFA (рис. 28): $x_0 = v_0^2/2$, $v_0 < 0$; управлению б)— точки на дуге ODB: $x_0 = -v_0^2/2$, $v_0 > 0$.

Начальные точки C, лежащие слева от разделительной линии $BDOFA$ $(x_0 = -v_0|v_0|/2)$ отвечают управлениям в): на дуге CD семейства (9) $\hat{u}(t) \equiv 1$, в момент $t=\tau$ попадаем в точку D, происходит переключение управления, и далее движемся по дуге DO с $\hat{u}(t) \equiv -1$. Аналогично начальным точкам E справа от разделительной кривой соответствуют управления г).

6. Остается показать, что найденное единственное решение $\hat{x}(t) = \hat{x}_1(t)$, отвечающее заданной начальной точке (x_0, v_0), действительно доставляет решение задаче (2).

Предположим, что некоторая функция $x(\cdot)$ определена на отрезке $[0, \tilde{T}]$, имеет кусочно-непрерывную вторую производную и $x(0) = x_0$, $\dot{x}(0) = v_0$, $x(\tilde{T}) = \dot{x}(\tilde{T}) = 0$, причем $\tilde{T} \leqslant T$. При $\tilde{T} < T$ мы доопределим $x(\cdot)$, положив $x(t) \equiv 0$, $t \in [\tilde{T}, T]$. После этого обе функции $x(\cdot)$ и $\hat{x}(\cdot)$

будут определены на одном и том же отрезке $[0, T]$ и будут иметь одинаковые граничные значения

$$x(0)=\hat{x}(0)=x_0, \quad \dot{x}(0)=\dot{\hat{x}}(0)=v_0,$$
$$x(T)=\hat{x}(T)=\dot{x}(T)=\dot{\hat{x}}(T)=0. \tag{11}$$

Покажем, что если $|\ddot{x}|\leqslant 1$, то $x(\cdot)=\hat{x}(\cdot)$ и, в частности, неравенство $\tilde{T}<T$ невозможно. Тем самым будет доказана оптимальность функции $\hat{x}(\cdot)$. Ввиду симметрии задачи ограничимся управлением в) (или его предельным случаем а)). Если $|\ddot{x}|\leqslant 1$, то, интегрируя дважды неравенство $\ddot{x}(t)\leqslant 1$ и учитывая (11), получаем

$$\hat{x}(\tau)-x(\tau)=\int_0^\tau\int_0^t(1-\ddot{x}(s))\,ds\,dt\geqslant 0, \tag{12}$$

причем равенство здесь возможно, только если во всех точках непрерывности $\ddot{x}(s)\equiv 1$, а тогда $x(t)\equiv\hat{x}(t)$, $t\in\in[0, \tau]$.

Аналогично, интегрируя дважды неравенство $\ddot{x}(t)\geqslant -1$, получаем

$$\hat{x}(\tau)-x(\tau)=\int_\tau^T\int_t^T(-1-\ddot{x}(s))\,ds\,dt\leqslant 0, \tag{13}$$

причем и здесь равенство возможно лишь, если $\ddot{x}(s)\equiv -1$ и $x(t)\equiv\hat{x}(t)$, $t\in[\tau, T]$.

Сравнивая (12) и (13), находим, однако, что $\hat{x}(\tau)=x(\tau)$, а тогда, как уже было сказано, $\hat{x}(t)\equiv x(t)$, $t\in[0, T]$.

Примечание. Из соотношений (5) — (7) следует равенство $\hat{\lambda}_0=|p_2(T)|$, и, таким образом, равенство $\hat{\lambda}_0=0$ в этой задаче оказывается возможным, когда $p_2(\cdot)$ обращается в нуль при $t=T$. Тогда функция $p_2(\cdot)\not\equiv 0$ (ибо иначе все множители Лагранжа были бы нулями) и вследствие равенства $p_2(T)=0$, $p_2(\cdot)$ не меняет знака и, значит, переключений управления нет вовсе. Следовательно, случай $\hat{\lambda}_0=0$ соответствует движению по линиям переключения AFO и BDO.

1.6.4. Классическая изопериметрическая задача и задача Чаплыгина. Стариннейшая экстремальная задача — первая из двух названных в заголовке — была поставлена в п. 1.1.1 и формализована разными способами в п. 1.2.4. В частности, формализация (2) п. 1.2.4 сводит ее к более общей задаче, охватывающей также и задачу Чаплыгина. Исходя из этой формализации, приведем сейчас решение обеих задач, следуя нашей стандартной схеме. Итак:

1. $$S=\frac{1}{2}\int_0^T (xv-yu)\,dt\to\sup;$$
$$\dot{x}=u,\quad \dot{y}=v,\quad (u,\ v)\in A,\quad x(0)=x(T),\quad y(0)=y(T). \tag{1}$$

Множество допустимых скоростей A будем считать замкнутым выпуклым ограниченным множеством в $\mathbf{R}^2$.

Упражнение. Покажите, что для разрешимости поставленной задачи необходимо, чтобы $0\in A$. Указание: можно воспользоваться результатом упражнения 2 в п. 2.2.3.

2. $$\mathscr{L}=\int_0^T L\,dt+\mu\,(x(0)-x(T))+\nu\,(y(0)-y(T)), \tag{2}$$

где

$$L=-\frac{\lambda_0}{2}(xv-yu)+p\,(\dot{x}-u)+q\,(\dot{y}-v).$$

3. Уравнения Эйлера—Лагранжа:

$$\begin{aligned} -\frac{d}{dt}L_{\dot{x}}+L_x=0 &\Rightarrow -\dot{\hat{p}}-\frac{\hat{\lambda}_0}{2}\hat{v}=0,\\ -\frac{d}{dt}L_{\dot{y}}+L_y=0 &\Rightarrow -\dot{\hat{q}}+\frac{\hat{\lambda}_0}{2}\hat{u}=0. \end{aligned} \tag{3}$$

Условия трансверсальности:

$$\hat{p}(0)=\hat{p}(T)=\hat{\mu},\quad \hat{q}(0)=\hat{q}(T)=\hat{\nu}.$$

Принцип максимума:

$$\left(\hat{p}(t)-\frac{\hat{\lambda}_0}{2}\hat{y}(t)\right)\hat{u}(t)+\left(\hat{q}(t)+\frac{\hat{\lambda}_0}{2}\hat{x}(t)\right)\hat{v}(t)=$$
$$=\max_{(u,\ v)\in A}\left(\left(\hat{p}(t)-\frac{\hat{\lambda}_0}{2}\hat{y}(t)\right)u+\left(\hat{q}(t)+\frac{\hat{\lambda}_0}{2}\hat{x}(t)\right)v\right). \tag{4}$$

4. Если допустить, что $\hat{\lambda}_0=0$, то из уравнений Эйлера — Лагранжа $\hat{p}(t)\equiv\text{const}$, $\hat{q}(t)\equiv\text{const}$, причем $\hat{p}^2+\hat{q}^2>0$, ибо иначе ввиду условий трансверсальности все множители Лагранжа оказались бы нулями. Теперь

принцип максимума приобретает вид

$$\hat{p}\hat{u}(t)+\hat{q}\hat{v}(t)=\max_{(u,\,v)\in A}(\hat{p}u+\hat{q}v),$$

откуда видно, что $(\hat{u}(t),\ \hat{v}(t))$ все время принадлежит одной и той же прямой, а именно одной из двух опорных к множеству A прямых, перпендикулярных вектору $(\hat{p},\ \hat{q})$ (рис. 29). Поэтому

$$\dot{\hat{x}}(t)=\hat{u}(t)=\hat{u}(0)-\alpha(t)\hat{q},$$
$$\dot{\hat{y}}(t)=\hat{v}(t)=\hat{v}(0)+\alpha(t)\hat{p},$$

откуда, интегрируя, получаем

$$\begin{aligned}\hat{x}(t)&=\hat{x}(0)+\hat{u}(0)\,t-\int_0^t\alpha(t)\,dt\cdot\hat{q},\\ \hat{y}(t)&=\hat{y}(0)+\hat{v}(0)\,t+\int_0^t\alpha(t)\,dt\cdot\hat{p}.\end{aligned}\tag{5}$$

Полагая в (5) $t=T$ и используя краевые условия $x(0)=x(T)$, $y(0)=y(T)$, убеждаемся, что вектор $(\hat{q},-\hat{p})$ пропорционален $(\hat{u}(0),\ \hat{v}(0))$, после чего из (5) видно, что $(\hat{x}(t),\ \hat{y}(t))$ все время лежит на прямой, проходящей через $(\hat{x}(0),\ \hat{y}(0))$ параллельно вектору $(\hat{q},-\hat{p})$. Следовательно, замкнутая кривая $\{(\hat{x}(t),\ \hat{y}(t)),\, 0\leqslant t\leqslant T\}$ вырождается (лежит на прямой) и ограничиваемая ею площадь равна нулю (покажите это аналитически, исходя из указанного выше выражения для S!). Следовательно, эта кривая не может быть оптимальной.

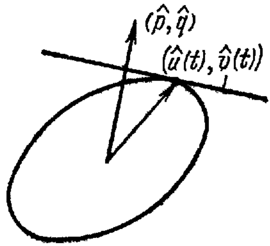


Рис. 29.

5. Из этапа 4 следует, что можно положить $\hat{\lambda}_0=1$. Тогда из (3)

$$\dot{\hat{p}}+\frac{\hat{v}}{2}=0\Rightarrow\hat{p}(t)+\frac{\hat{y}(t)}{2}=b=\text{const},$$
$$\dot{\hat{q}}-\frac{\hat{u}}{2}=0\Rightarrow\hat{q}(t)-\frac{\hat{x}(t)}{2}=a=\text{const}.$$

Подставляя эти выражения в (4), получаем

$$(\hat{x}(t)-a)\hat{v}(t)-(\hat{y}(t)-b)\hat{u}(t)=$$
$$=\max_{(u,\ v)\in A}\{(\hat{x}(t)-a)v-(\hat{y}(t)-b)u\}. \qquad (6)$$

Оставляя теперь в стороне общий случай, ограничимся двумя частными вариантами.

а) Классическая изопериметрическая задача. Здесь

$$A=\{(u,\ v)\mid u^2+v^2\leqslant 1\}$$

и из (6) находим

$$\dot{\hat{x}}(t)=\hat{u}(t)=-\frac{\hat{y}(t)-b}{\mathscr{H}},\quad \dot{\hat{y}}(t)=\hat{v}(t)=\frac{\hat{x}(t)-a}{\mathscr{H}}, \qquad (7)$$

где

$$\mathscr{H}=((\hat{x}(t)-a)^2+(\hat{y}(t)-b)^2)^{1/2}$$

(скалярное произведение векторов $(\hat{x}-a,\ \hat{y}-b)$ и $(v,\ -u)$ будет наибольшим, когда они одинаково направлены и второй имеет максимально возможную длину, т. е. 1). Из (7) имеем

$$\frac{d}{dt}(\mathscr{H}^2)=0,$$

т. е. на оптимальной траектории

$$\mathscr{H}(\hat{x}(t),\ \hat{y}(t))=((\hat{x}(t)-a)^2+(\hat{y}(t)-b)^2)^{1/2}=$$
$$=\max_{(u,\ v)\in A}((\hat{x}(t)-a)v-(\hat{y}(t)-b)u)=R=\text{const}. \qquad (8)$$

Отсюда заключаем, что *оптимальная траектория является окружностью с центром в точке* (a, b) *и радиусом* R Поскольку угловая скорость движения по этой окружности равна

$$\frac{d\varphi}{dt}=\frac{1}{R^2}((x-a)\dot{y}-(y-b)\dot{x})\equiv\frac{1}{R},$$

и через время T мы должны возвратиться в ту же точку (быть может, совершив несколько оборотов), $2\pi n=\frac{T}{R}$ и $R=\frac{T}{2\pi n}$. Площадь, ограниченная найденной кривой, будет при этом равна

$$S=\frac{1}{2}\int_0^T(xv-yu)\,dt=\frac{1}{2}\int_0^T(x\dot{y}-y\dot{x})\,dt=\frac{1}{2}\int_0^T((x-a)\,\dot{y}-$$

$$-(y-b)\,\dot{x})\,dt=\frac{R^2}{2}\frac{T}{R}=\frac{T^2}{4\pi n}.$$

Следовательно, $n=1$ (окружность обходится только один раз) и $S=T^2/4\pi$. Заметим, что здесь T — это длина кривой.

б) Задача Чаплыгина. Здесь

$$A=\{(u,\ v)\,|\,(u-u_0)^2+v^2\leqslant V^2\},$$

$(u_0,\ 0)$ — вектор скорости ветра, $V\geqslant|u_0|$ — максимальная скорость самолета. Для решения вспомогательной экстремальной задачи

$$(\hat{x}(t)-a)\,v-(\hat{y}(t)-b)\,u\to\sup,$$
$$(u-u_0)^2+v^2\leqslant V^2,$$

можно, например, сделать подстановку $u=u_0+\xi$, $v=\eta$ и воспользоваться тем же геометрическим соображением, что и в предыдущем случае. Это дает уравнения

$$\dot{\hat{x}}(t)=\hat{u}(t)=u_0-V\frac{\hat{y}(t)-b}{R(t)},\quad \dot{\hat{y}}(t)=\hat{v}(t)=V\frac{\hat{x}(t)-a}{R(t)},$$

где теперь уже величина

$$R(t)=((\hat{x}(t)-a)^2+(\hat{y}(t)-b)^2)^{1/2}$$

не будет постоянной. Однако

$$\frac{dR(t)}{dt}=\frac{1}{R(t)}((\hat{x}(t)-a)\,\dot{\hat{x}}(t)+(\hat{y}(t)-b)\,\dot{\hat{y}}(t))=$$

$$=\frac{\hat{x}(t)-a}{R(t)}\,u_0=\frac{u_0}{V}\frac{d\hat{y}(t)}{dt},$$

так что

$$\mathscr{H}(\hat{x}(t),\ \hat{y}(t))=R(t)-\frac{u_0}{V}\hat{y}(t)=((\hat{x}(t)-a)^2+(\hat{y}(t)+b)^2)^{1/2}-$$

$$-\frac{u_0}{V}\hat{y}(t)=V\max_{(u,\ v)\in A}\{(\hat{x}(t)-a)\,v-(\hat{y}(t)-b)\,u\}=\mathrm{const}. \qquad (9)$$

Следовательно, *оптимальная кривая — это эллипс с уравнением*

$$((x-a)^2+(y-b)^2)^{1/2}-\frac{u_0}{V}y=\mathrm{const}. \qquad (10)$$

Как и в предыдущем случае, можно показать, что он обходится траекторией $(\hat{x}(t), \hat{y}(t))$ ровно один раз. Заметим, что если бы скорость самолета была меньше скорости ветра ($V < |u_0|$), он не мог бы вернуться в исходную точку (докажите это и сравните с упражнением в начале этого пункта), а кривая (10) была бы гиперболой, т. е. незамкнутой.

Примечание. Вопреки принятой схеме, мы обошли молчанием вопрос о существовании решения. Это не случайно: простой ссылкой на теорему Вейерштрасса здесь не обойдешься. Наметим вкратце возможный путь рассуждений.

Прежде всего без ограничения общности можно считать, что $x(0)=y(0)=0$.

Используя теорему Арцела [КФ, с. 110—111], можно установить, что множество пар функций $\{(x(\cdot), y(\cdot))\}$, удовлетворяющих на отрезке $[0, T]$ обобщенному условию Липшица

$$\left(\frac{x(t_1)-x(t_2)}{t_1-t_2}, \frac{y(t_1)-y(t_2)}{t_1-t_2}\right) \in A, \ t_1 \neq t_2, \ 0 \leqslant t_1, \ t_2 \leqslant T \quad (11)$$

и краевым условиям

$$x(0)=x(T)=0, \ y(0)=y(T)=0$$

компактно в пространстве $C([0, T]) \times C([0, T])$.

Далее, на этом множестве функционал «площадь» определен и является полунепрерывной снизу функцией. Применяя соответствующее обобщение теоремы Вейерштрасса, убеждаемся в существовании решения $(\hat{x}(\cdot), \hat{y}(\cdot))$.

Наконец, функции, удовлетворяющие условию (11), дифференцируемы почти всюду, причем $(\dot{\hat{x}}(t), \dot{\hat{y}}(t)) \in A$. На эту ситуацию оказывается возможным распространить принцип максимума, и, следовательно, к решению $(\hat{x}(\cdot), \hat{y}(\cdot))$ применимы приведенные выше рассуждения.

1.6.5. Задача о брахистохроне и некоторые задачи геометрии. Рассмотрим следующую серию простейших задач классического вариационного исчисления:

1.

$$\mathcal{J}(y(\cdot)) = \int_{x_0}^{x_1} y^\alpha \sqrt{1+y'^2}\, dx \to \inf; \quad y(x_0)=y_0, \quad y(x_1)=y_1. \quad (1)$$

В нее попадает несколько интересных задач: при $\alpha=0$ получается задача о кратчайших линиях на плоскости, при $\alpha=-1/2$ — задача о брахистохроне (см. п. 1.2.4), при $\alpha=1$ — задача о минимальной поверхности вращения; при $\alpha=-1$ — задача о кратчайших линиях на плоскости Лобачевского в интерпретации Пуанкаре (полуплоскость Пуанкаре [13, с. 131—132]).

Интегрант $f_\alpha=y^\alpha\sqrt{1+y'^2}$ в задачах (1) не зависит от x. Поэтому уравнение Эйлера (п. 1.4.1) имеет интеграл энергии, из которого мы и находим все экстремали, лежащие в верхней полуплоскости.

2—4.

$$y'f_{\alpha y'}-f=\text{const}\Rightarrow y^{-2\alpha}(1+y'^2)=D^2. \qquad (2)$$

5. Сначала разберем случай отрицательных α: $\alpha=-\beta$, $\beta>0$. Тогда

$$\frac{y^\beta\,dy}{\sqrt{D^2-y^{2\beta}}}=dx\Rightarrow x-C_1=\int\frac{y^\beta\,dy}{\sqrt{D^2-y^{2\beta}}}. \qquad (3)$$

Делаем замену

$$y^{2\beta}=D^2\sin^2 t\Rightarrow y=D^{1/\beta}\sin^{1/\beta}t\Rightarrow dy=\frac{D^{1/\beta}}{\beta}\sin^{1/\beta-1}t\cos t\,dt.$$

Подставляя полученные выражения в (3), приходим к следующим соотношениям ($D^{1/\beta}=C$):

$$y=C\sin^{1/\beta}t,\quad x=C_1+\frac{C}{\beta}\int\limits_0^t\sin^{1/\beta}s\,ds.$$

В частности, если $\beta=1/2$ (брахистохрона), получается семейство циклоид, заданных в параметрической форме ($\tau=2t$):

$$\hat{y}(\tau,C,C_1)=\frac{C}{2}(1-\cos\tau),\ \hat{x}(\tau,C,C_1)=C_1+\frac{C}{2}(\tau-\sin\tau).$$

Если $\beta=1$ (полуплоскость Пуанкаре), получаем семейство полуокружностей

$$\hat{y}(t,C,C_1)=C\sin t,\ \hat{x}(t,C,C_1)=C_1-C\cos t\Rightarrow$$
$$\Rightarrow(\hat{x}-C_1)^2+\hat{y}^2=C^2.$$

Из серии $\alpha>0$ решим лишь пример о минимальной поверхности вращения. Из (2) имеем

$$\frac{dy}{\sqrt{D^2y^2-1}}=dx.$$

С помощью подстановки

$$Dy = \operatorname{ch} t \Rightarrow D\,dy = \operatorname{sh} t\,dt, \quad \sqrt{D^2y^2-1} = \operatorname{sh} t$$

находим двупараметрическое семейство цепных линий:

$$y = \frac{1}{D}\operatorname{ch}(Dx + D_1).$$

Получив во всех случаях двупараметрическое семейство экстремалей (решений уравнений Эйлера), нужно теперь подобрать постоянные интегрирования так, чтобы удовлетворялись краевые условия $y(x_i) = y_i$, и перейти к этапу 6 — исследованию найденных решений и записи ответа.

Это, однако, здесь сделано не будет. Вспомнив, что вдобавок мы снова уклонились от обсуждения вопроса о существовании решения, читатель с еще большим основанием, чем в предыдущем пункте, вправе заявить: «Неладно что-то в оптимальном королевстве!» И действительно:

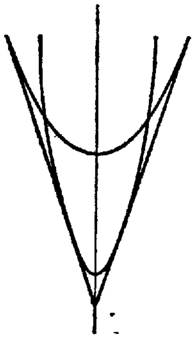


Рис. 30.

При $\alpha > 0$ задача (1) может не иметь обычного решения. Например, если в задаче о минимальной поверхности вращения ($\alpha = 1$) граничные условия симметричны ($y(-x_0) = y(x_0)$), то надлежит искать симметричную экстремаль, т. е. $y = \operatorname{ch} Dx/D$. Все экстремали этого вида получаются из экстремали $y = \operatorname{ch} x$ гомотетией, и в совокупности они заполняют только угол, а не всю полуплоскость (рис. 30). Поэтому при достаточно большом x_0 задача, скажем, с условиями $y(-x_0) = y(x_0) = 1$, неразрешима.

При $\alpha < 0$ интегрант в задаче (1) стремится к бесконечности при $y \to 0$. Поэтому, например, задача о брахистохроне $\alpha = -1/2$ с обычными краевыми условиями $y(x_0) = 0$, $y(x_1) > 0$ (п. 1.1.4) вообще не укладывается в стандартные рамки.

Для того чтобы получить полное решение задач серии (1), нужна дополнительная работа.

ГЛАВА II

АППАРАТ ТЕОРИИ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

Читатель познакомился с различными постановками экстремальных задач, с основными понятиями и общими идеями, а также получил представление о некоторых приемах решения этих задач. Теперь мы переходим к последовательному и достаточно формальному изложению соответствующей математической теории.

Общность развиваемой теории и ее сравнительная простота оказались возможными за счет свободного использования фактов и понятий, относящихся к смежным разделам математики, в первую очередь функционального анализа. Здесь они изложены в нужной нам форме. Значительную часть главы составляет то, что можно было бы назвать «Элементами дифференциального исчисления».

§ 2.1. Предварительные сведения из функционального анализа

В этом параграфе рассказывается о тех фактах функционального анализа, на которые мы далее опираемся при построении теории экстремальных задач. Доказательства, которые можно прочесть в учебнике А. Н. Колмогорова и С. В. Фомина [КФ], мы, как правило, опускаем.

2.1.1. Линейные нормированные и банаховы пространства. Напомним, что линейное пространство X называется *нормированным*, если на X определен функционал $\|\cdot\|: X \to \mathbf{R}$, называемый *нормой* и удовлетворяющий трем условиям:

а)
$$\|x\| \geqslant 0, \quad \forall x \in X \quad \text{и} \quad \|x\| = 0 \quad \Leftrightarrow x = \theta,$$
б)
$$\|\alpha x\| = |\alpha| \, \|x\|, \quad \forall x \in X, \quad \forall \alpha \in \mathbf{R}, \tag{1}$$
в)
$$\|x_1 + x_2\| \leqslant \|x_1\| + \|x_2\|, \quad \forall x_1, \quad x_2 \in X.$$

Иногда, чтобы подчеркнуть, что норма задана именно в X, мы пишем $\|\cdot\|_X$.

Всякое нормированное пространство становится *метрическим*, если ввести в нем расстояние
$$\rho(x_1, x_2) = \|x_1 - x_2\|.$$

Полное (относительно введенного расстояния) линейное нормированное пространство называется *банаховым* пространством.

Упражнения. Пусть $X = \mathbf{R}^2$. Какие функции из перечисленных ниже и при каких значениях параметров задают норму в X?

1. $N(x) = (|x_1|^p + |x_2|^p)^{1/p}, \quad 0 < p < \infty$;
2. $N(x) = |a_{11}x_1 + a_{12}x_2| + |a_{21}x_1 + a_{22}x_2|$;
3. $N(x) = \max\{|a_{11}x_1 + a_{12}x_2|, \quad |a_{21}x_1 + a_{22}x_2|\}$.
4. Описать все нормы в $\mathbf{R}^2$.
5. Доказать, что все нормы в $\mathbf{R}^2$ эквивалентны. (Нормы N_1 и N_2 называются *эквивалентными*, если существуют такие $c > 0$ и $C > 0$, что $cN_1(x) \leqslant N_2(x) \leqslant CN_1(x)$, $\forall x \in X$.)

Совокупность X^* всех линейных непрерывных функционалов на X *(сопряженное к X пространство)* является банаховым пространством, если задать в X^* норму
$$\|x^*\|_{X^*} = \sup_{\|x\|_X \leqslant 1} \langle x^*, x \rangle,$$
где $\langle x^*, x \rangle$ означает результат применения к x функционала x^*.

Подробнее об этом см. [КФ, гл. IV, § 2].

Упражнение 6. Найти нормы в пространствах, сопряженных к описанным в упражнении 1—4.

Для нас важнейшую роль будут играть следующие банаховы пространства.

Пример 1. Пространство $C(K, \mathbf{R}^n)$ непрерывных вектор-функций $x(\cdot)\colon K \to \mathbf{R}^n$, заданных на компакте K

с нормой

$$\|x(\cdot)\|_0=\max_{t\in K}|x(t)|.$$

Пространство $C(K, \mathbf{R})$ мы обозначаем просто $C(K)$.

Пример 2. Пространство $C^r([t_0, t_1], \mathbf{R}^n)$ r раз непрерывно дифференцируемых вектор-функций $x(\cdot)$: $[t_0, t_1]\to\mathbf{R}^n$, заданных на конечном отрезке $[t_0, t_1]\subset\mathbf{R}$, с нормой

$$\|x(\cdot)\|_r=\max(\|x(\cdot)\|_0, \ldots, \|x^{(r)}(\cdot)\|_0).$$

Пространство $C^r([t_0, t_1], \mathbf{R})$ мы обозначаем $C^r([t_0, t_1])$.

Упражнения.

7. Доказать, что всякое конечномерное нормированное пространство является банаховым.

8. Доказать, что единичный шар конечномерного нормированного пространства является выпуклым, замкнутым, ограниченным центрально-симметричным множеством, для которого начало координат является внутренней точкой, и наоборот, для любого выпуклого замкнутого ограниченного центрально-симметричного множества, для которого начало координат является внутренней точкой, существует такая норма, в которой это множество является единичным шаром.

9. Привести пример нормированного, но не банахова пространства.

2.1.2. Произведение пространств. Фактор-пространство. Пусть X и Y — линейные пространства. Их произведение $X\times Y$, т. е. множество всех пар (x, y), $x\in X$, $y\in Y$, превращается в линейное пространство, если операции сложения и умножения на число определить покоординатно:

$$(x_1, y_1)+(x_2, y_2)=(x_1+x_2, y_1+y_2), \quad \alpha(x, y)=(\alpha x, \alpha y).$$

Если X и Y — нормированные пространства, то и в произведении $X\times Y$ можно ввести норму, например, так:

$$\|(x, y)\|_{X\times Y}=\max\{\|x\|_X, \|y\|_Y\}. \tag{1}$$

Упражнение. Проверьте, что $\|x\|_X+\|y\|_Y$ и $\sqrt{\|x\|_X^2+\|y\|_Y^2}$ — также некоторые нормы в $X\times Y$, эквивалентные норме (1).

Имеют место следующие очевидные леммы.

Лемма 1. *Если X и Y — банаховы пространства, то и $X\times Y$ банахово.*

Доказательство предоставляется читателю.

Лемма 2. *Всякий линейный функционал* $\Lambda \in (X \times Y)^*$ *однозначно представим в виде*

$$\langle \Lambda, (x, y) \rangle = \langle x^*, x \rangle + \langle y^*, y \rangle, \tag{2}$$

где $x^* \in X^*$ *и* $y^* \in Y^*$.

Доказательство. Положим $\langle x^*, x \rangle = \langle \Lambda, (x, 0) \rangle$, $\langle y^*, y \rangle = \langle \Lambda, (0, y) \rangle$. Линейность этих функционалов очевидна, а непрерывность ($\Leftrightarrow$ ограниченность) вытекает из оценок:

$$|\langle x^*, x \rangle| \leqslant \|\Lambda\| \|(x, 0)\| = \|\Lambda\| \|x\|,$$
$$|\langle y^*, y \rangle| \leqslant \|\Lambda\| \|(0, y)\| = \|\Lambda\| \|y\|.$$

Однозначность представления (2) также не вызывает сомнений. ■

Поскольку при любых $x^* \in X^*$ и $y^* \in Y^*$ формула (2) определяет некоторый функционал $\Lambda \in (X \times Y)^*$, то мы получили полное описание пространства $(X \times Y)^*$. Кратко оно дается следующей формулой: $(X \times Y)^* = X^* \oplus Y^*$.

Пусть теперь X — линейное пространство, L — некоторое его подпространство Положим $x \sim x'$, если $x - x' \in L$. Введенное отношение будет отношением эквивалентности (ибо очевидно, что $x \sim x$, $x_1 \sim x_2 \Rightarrow x_2 \sim x_1$ и $x_1 \sim x_2$, $x_2 \sim x_3 \Rightarrow x_1 \sim x_3$), а значит [КФ, гл. I, § 2] определено разбиение X на классы. Класс эквивалентных элементов по введенному выше отношению называется *классом смежности* по подпространству L. В совокупности классов смежности естественным образом вводятся операции сложения и умножения на число (КФ, гл. III, § 1, п. 4); при этом выполняются аксиомы линейного пространства. Таким образом, совокупность классов смежности превращается в линейное пространство, называемое *фактор-пространством X по L* и обозначаемое X/L.

Классом смежности $\pi(x)$, которому принадлежит элемент $x \in X$, является класс $x + L$. Отображение π: $X \to X/L$ является линейным (докажите!). Его называют *каноническим отображением* X на X/L. (Отображение π является, разумеется, эпиморфизмом[1]).) Очевидно, что $L = \operatorname{Ker} \pi$.

Пусть теперь X — нормированное пространство и L — его подпространство. Положим

$$\|\xi\|_{X/L} = \inf_{x \in \xi} \|x\|_X, \tag{3}$$

[1]) То есть отображает X на все X/L.

или

$$\|\xi\|_{X/L}=\inf_{\pi x=\xi}\|x\|_X=\inf_{\substack{\pi x_0=\xi\\ x'\in L}}\|x_0+x'\|. \qquad (3')$$

Из определений (3), (3′) сразу следует, что оператор π непрерывен ($\|\pi x\|\leqslant\|x\|$) и для любого $\xi\in X/L$ найдется $x\in X$, для которого

$$\pi x=\xi,\quad \|x\|\leqslant 2\|\xi\|_{X/L}. \qquad (4)$$

Теорема о фактор-пространстве. *Пусть X — нормированное пространство и L — его замкнутое подпространство. Тогда функция $\|\cdot\|_{X/L}$, задаваемая соотношением* (3), *является нормой на X/L. Если пространство X является банаховым, то фактор-пространство X/L с нормой $\|\cdot\|_{X/L}$ также является банаховым.*

Доказательство. Необходимо доказать, во-первых, что функционал $\|\cdot\|_{X/L}$ удовлетворяет аксиомам нормы и, во-вторых, что полнота X влечет за собой полноту X/L. Докажем первое утверждение теоремы, т. е. проверим выполнение аксиом (1) п. 2.1.1.

Аксиома а): $\|\xi\|_{X/L}\geqslant 0$ $(\forall\xi)$ вследствие неотрицательности нормы в X. Если $\xi=0$, то в качестве $x\in\xi$ можно взять $x=0$, и потому $\|0\|_{X/L}=0$. Пусть $\|\xi\|_{X/L}=0$. Тогда из (3) следует, что существует последовательность $\{x_n\}$, $x_n\in\xi$ такая, что $\|x_n\|\to 0$, т. е. $x_n\to 0$. Вследствие замкнутости ξ, $0\in\xi$, т е. ξ есть нулевой элемент в X/L.

Аксиома б): Пусть $\alpha\in\mathbb{R}$. Тогда, если $\pi x=\xi$, то

$$\|\alpha\xi\|_{X/L}=\inf_{\pi y=\alpha\xi}\|y\|=\inf_{\pi x=\xi}\|\alpha x\|=|\alpha|\inf_{\pi x=\xi}\|x\|=|\alpha|\|\xi\|.$$

Аксиома в):

$$\|\xi_1+\xi_2\|=\inf_{\substack{\pi x_i=\xi_i\\ x'\in L}}\|x_1+x_2+x'\|=\inf_{x_i'\in L}\|x_1+x_1'+x_2+x_2'\|\leqslant$$

$$\leqslant\inf_{x_1'\in L}\|x_1+x_1'\|+\inf_{x_2'\in L}\|x_2+x_2'\|=\|\xi_1\|_{X/L}+\|\xi_2\|_{X/L}.$$

Докажем второе утверждение теоремы, т. е. полноту X/L. Пусть $\{\xi_n\}$ — фундаментальная последовательность в X/L, т. е.

$$\forall\varepsilon>0\ \exists N(\varepsilon).\ n>N(\varepsilon)\Rightarrow\|\xi_{n+m}-\xi_n\|_{X/L}<\varepsilon,\quad \forall m\geqslant 1.$$

Выберем $\varepsilon_k=2^{-k}$ и номера n_k такие, что $\|\xi_{n_k+m}-\xi_{n_k}\|\leqslant 2^{-k}$,

$k \geqslant 1$. Тогда $\|\xi_{n_2}-\xi_{n_1}\| \leqslant 1/2$, и вследствие (4) существуют представители $x_i \in \xi_{n_i}$ такие, что $\|x_2-x_1\| \leqslant 1$. Аналогичным путем построим элементы $\{x_n\}_{n\geqslant 3}$ так, что $\|x_k-x_{k-1}\| \leqslant 2^{-(k-1)}$, $x_k \in \xi_{n_k}$, $k=3, 4, \ldots$ Последовательность $\{x_k\}_{k=1}^{\infty}$ фундаментальна в X (проверьте), а X — полно по условию. Значит, существует предел $x_0 = \lim\limits_{k\to\infty} x_k$. Рассмотрим класс $\xi_0 = \pi x_0$. Тогда

$$\|\xi_{n_k}-\xi_0\|_{X/L} = \inf_{x'\in L} \|x_k-x_0-x'\|_X \leqslant \|x_k-x_0\|_X \underset{k\to\infty}{\longrightarrow} 0 .$$

Итак, $\xi_{n_k} \to \xi_0$, а тогда и $\xi_n \to \xi_0$ в силу фундаментальности $\{\xi_n\}$. Значит, X/L — полное. ∎

Упражнения.

1. Пусть $X=C([0,1])$, $0 \leqslant \tau_1 < \tau_2 < \ldots < \tau_m \leqslant 1$, подпространство L состоит из функций $x(\cdot) \in C([0,1])$ таких, что $x(\tau_i)=0$, $i=1, \ldots, m$. Найти фактор-пространство X/L и определить $\|\cdot\|_{X/L}$.

2. Обобщить результат упражнения 1 и доказать, что если F — замкнутое подмножество отрезка $[0, 1]$ и

$$L_F = \{x(\cdot) \in C([0,1]) \mid x(t) \equiv 0,\ t \in F\},$$

то пространство $C([0,1])/L_F$ изометрически изоморфно пространству $C(F)$.

2.1.3. Теорема Хана — Банаха и ее следствия. В теории экстремальных задач важную роль играют теоремы отделимости и некоторые другие факты выпуклого анализа. Большинство из них является следствием теоремы Хана—Банаха, которую часто называют первым основным принципом линейного анализа. Учитывая, что эта теорема входит во все стандартные курсы функционального анализа, ограничимся здесь лишь ее формулировкой и некоторыми важнейшими следствиями.

В этом пункте X — линейное пространство, $\overline{\mathbb{R}}$ — расширенная числовая прямая, т. е. $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Определение 1. Функция $p\colon X \to \overline{\mathbb{R}}$ называется *выпуклой и однородной*, если

$$p(x+y) \leqslant p(x)+p(y) \quad \text{для любых} \quad x, y \in X, \qquad (1)$$

$$p(\alpha x) = \alpha p(x) \quad \text{для любых} \quad x \in X \text{ и } \alpha > 0. \qquad (2)$$

Примеры. 1) X — нормированное пространство, $p(x)=\|x\|$.

2) X — линейное пространство, L — линейное подпространство в X,

$$p(x) = \begin{cases} 0, & x \in L, \\ +\infty, & x \notin L. \end{cases}$$

3) X — линейное пространство, $l_1, \ldots, l_m$ — набор линейных функционалов на X; $p(x) = \max(\langle l_1, x\rangle, \ldots, \langle l_m, x\rangle)$.

Важный пример доставляет следующее

Определение 2. Пусть A — выпуклое подмножество линейного пространства X, содержащее 0. Его *функция Минковского* $\mu A(\cdot)$ определяется равенством

$$\mu A(x) = \inf\{t > 0 \mid x/t \in A\} \tag{3}$$

(если таких $t > 0$, что $x/t \in A$, нет вовсе, то $\mu A(x) = +\infty$).

Предложение 1. *Функция Минковского неотрицательна, выпукла и однородна;*

$$\{x \mid \mu A(x) < 1\} \subseteq A \subseteq \{x \mid \mu A(x) \leqslant 1\}. \tag{4}$$

Если X — линейное топологическое пространство[1]*, то $\mu A(\cdot)$ непрерывна в точке* 0 *тогда и только тогда, когда* $0 \in \operatorname{int} A$.

Доказательство. Если $\mu A(x)$ или $\mu A(y)$ равно $+\infty$, то (1) верно. Поэтому пусть $\mu A(x) < +\infty$ и $\mu A(y) < +\infty$. По определению для любого $\varepsilon > 0$ найдутся такие t и s, что

$$0 < t < \mu A(x) + \varepsilon/2 \quad \text{и} \quad x/t \in A,$$
$$0 < s < \mu A(y) + \varepsilon/2 \quad \text{и} \quad y/s \in A.$$

Но тогда

$$\frac{x+y}{t+s} = \frac{x}{t}\frac{t}{t+s} + \frac{y}{s}\frac{s}{t+s} \in A,$$

поскольку A выпукло, и следовательно,

$$\mu A(x+y) \leqslant t + s < \mu A(x) + \mu A(y) + \varepsilon.$$

Ввиду произвольности ε верно (1). Далее, для $\alpha > 0$

$$\mu A(\alpha x) = \inf\{t > 0,\ \alpha x/t \in A\} = \inf\{\alpha s > 0,\ x/s \in A\} = \\ = \alpha \inf\{s > 0;\ x/s \in A\} = \alpha \mu A(x),$$

так что верно (2).

Неотрицательность $\mu A(x)$ и второе включение в (4) следуют непосредственно из определения. Если $\mu A(x) < 1$, то существует $t \in (0, 1)$, для которого $x/t \in A$, а поскольку

[1]) О линейных топологических пространствах см. [КФ, гл. III, § 5].

$0 \in A$ и A выпукло,

$$x = 0\cdot(1-t) + \frac{x}{t}\cdot t \in A,$$

так что первое включение в (4) верно.

Пусть теперь X — линейное топологическое пространство. Тогда имеет место следующая цепочка эквивалентностей: $\mu A(\cdot)$ непрерывна в $0 \Leftrightarrow$

$$\begin{gathered}\forall \varepsilon > 0,\ \exists U_\varepsilon \ni 0,\ \forall x \in U_\varepsilon,\quad \mu A(x) < \varepsilon \Leftrightarrow \\ \exists U_1 \ni 0,\ \forall x \in U_1,\quad \mu A(x) < 1 \Leftrightarrow \\ \exists U_1 \ni 0,\quad U_1 \subset A \Leftrightarrow 0 \in \operatorname{int} A\end{gathered}$$

(здесь U_ε и U_1 — окрестности точки 0; вторая эквивалентность верна ввиду однородности $\mu A(\cdot)$ и возможности положить $U_\varepsilon = \varepsilon U_1$). ■

Предложение 2. *Для того чтобы линейный функционал x^* на линейном топологическом пространстве X был непрерывен, необходимо и достаточно, чтобы для некоторой выпуклой однородной непрерывной в точке 0 функции $p(\cdot)$ для всех $x \in X$ выполнялось неравенство*

$$\langle x^*, x\rangle \leqslant p(x). \tag{5}$$

Доказательство. Необходимость устанавливаем сразу, полагая $p(x) = |\langle x^*, x\rangle|$.

Предположим теперь, что верно (5) и $p(x)$ непрерывна в точке 0. Тогда для любого $\varepsilon > 0$ найдется окрестность $U \ni 0$ такая, что $p(x) < \varepsilon$ для всех $x \in U$. Поскольку $0 \in U$ и $0 = (-0) \in (-U)$, существует окрестность W такая, что $0 \in W \subset U \cap (-U)$. Если $x \in W$, то x и $-x \in U$ и в силу (5)

$$\begin{gathered}\langle x^*, x\rangle \leqslant p(x) < \varepsilon, \\ \langle x^* - x\rangle \leqslant p(-x) < \varepsilon.\end{gathered}$$

Следовательно, $|\langle x^*, x\rangle| < \varepsilon$ для всех $x \in W$, т. е. x^* непрерывен в 0. Остается заметить, что линейный функционал, непрерывный в одной точке, непрерывен на X [КФ, стр. 174].

Теорема Хана—Банаха [КФ, стр. 134—137]. *Пусть $p: X \to \overline{\mathbf{R}}$ — выпуклая однородная функция на линейном пространстве X, и пусть $l: L \to \mathbf{R}$ — линейный функционал на подпространстве L пространства X такой, что*

$$\langle l, x\rangle \leqslant p(x) \quad \text{для всех} \quad x \in L. \tag{6}$$

Тогда существует линейный функционал Λ, определенный на всем X, являющийся продолжением l, т. е.

$$\langle \Lambda, x\rangle \equiv \langle l, x\rangle, \quad x \in L, \tag{7}$$

и удовлетворяющий неравенству

$$\langle \Lambda, x\rangle \leqslant p(x) \quad \text{для всех} \quad x \in X. \tag{8}$$

Следствие 1. *Пусть X — нормированное пространство и $x_0 \in X$, $x_0 \neq 0$. Тогда найдется элемент $\Lambda \in X^*$ такой, что $\|\Lambda\| = 1$ и $\langle \Lambda, x_0\rangle = \|x_0\|$.*

Доказательство. На подпространстве

$$L = \{x \mid \quad x = \alpha x_0, \ \alpha \in \mathbf{R}\}$$

зададим линейный функционал l, полагая

$$\langle l, \alpha x_0\rangle = \alpha \|x_0\|. \tag{9}$$

Функция $p(x) = \|x\|$ — выпуклая однородная и $\langle l, \alpha x_0\rangle = \alpha\|x_0\| \leqslant |\alpha|\,\|x_0\| = \|\alpha x_0\| = p(\alpha x_0)$, так что имеет место неравенство (6). По теореме Хана — Банаха l продолжается до линейного функционала Λ на всем X, причем выполняются соотношения (7) и (8). Вспоминая определение нормы функционала (п. 2.1.1), имеем из (8)

$$\|\Lambda\| = \sup_{\|x\|=1} \langle \Lambda, x\rangle \leqslant \sup_{\|x\|=1} p(x) = 1,$$

так что, в частности, $\Lambda \in X^*$.

С другой стороны, из (7) и (9)

$$\langle \Lambda, x_0\rangle = \langle l, x_0\rangle = \|x_0\|,$$

так что верно второе утверждение доказываемого следствия и, кроме того,

$$\|\Lambda\| = \sup_{\|x\|=1} \langle \Lambda, x\rangle \geqslant \left\langle \Lambda, \frac{x_0}{\|x_0\|}\right\rangle = 1,$$

а потому $\|\Lambda\| = 1$. ■

Из следствия 1 немедленно вытекает

Следствие 2. *Если нормированное пространство X нетривиально* (т. е. $X \neq \{0\}$), *то и сопряженное к нему пространство X^* нетривиально.*

2.1.4. Теоремы отделимости. В этом пункте X — линейное топологическое пространство, X^* — сопряженное к нему пространство, состоящее из всех линейных непрерывных функционалов на X.

Определение 1. Функционал $x^* \in X^*$ *разделяет множества* $A \subset X$ *и* $B \subset X$, если

$$\sup_{x \in A} \langle x^*, x \rangle \leqslant \inf_{x \in B} \langle x^*, x \rangle, \tag{1}$$

и *строго разделяет* A *и* B, если

$$\sup_{x \in A} \langle x^*, x \rangle < \inf_{x \in B} \langle x^*, x \rangle. \tag{2}$$

Геометрически неравенство (1) означает, что гиперплоскость

$$H(x^*, c) = \{x \mid \langle x^*, x \rangle = c\},$$

где $\sup_{x \in A} \langle x^*, x \rangle \leqslant c \leqslant \inf_{x \in B} \langle x^*, x \rangle$ отделяет множества A и B друг от друга в том смысле, что A лежит в одном полупространстве $(H_+(x^*, c) = \{x \mid \langle x^*, x \rangle \leqslant c\})$, порожденном $H(x^*, c)$, а B — в другом $(H_-(x^*, c) = \{x \mid \langle x^*, x \rangle \geqslant c\})$ (рис. 31), неравенство (2) означает, что при этом c можно

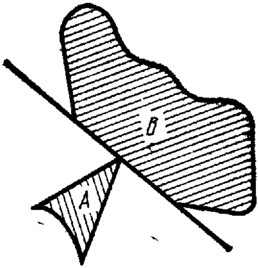


Рис. 31.

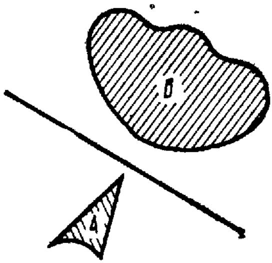


Рис. 32.

выбрать так, чтобы A и B лежали внутри соответствующих полупространств и не имели общих точек с $H(x^*, c)$ (рис. 32).

Первая теорема отделимости. *Если множества $A \subset X$ и $B \subset X$ выпуклы, непусты и не пересекаются между собой, причем A открыто, то существует ненулевой линейный непрерывный функционал, разделяющий A и B.*

Доказательство. А) Поскольку A и B не пусты, существуют точки $a_0 \in A$, $b_0 \in B$. Множество

$$C = (A - a_0) - (B - b_0) = \{x \mid x = a - a_0 - b + b_0,\ a \in A,\ b \in B\},$$

очевидно, выпукло (см. также предложение 1 и упражнение 2 в п. 2.6.1), содержит 0 и открыто (действительно, если $\hat{x}=\hat{a}-a_0-\hat{b}+b_0$ и $\hat{a}\in A$, то существует окрестность U, $\hat{a}\in U\subset A$, а тогда $\hat{x}\in U-a_0-\hat{b}+b_0\subset C$. Кроме того, $c=b_0-a_0\notin C$, так как в противном случае $b_0-a_0=\hat{a}-a_0-\hat{b}+b_0$ для некоторых $\hat{a}\in A$ и $\hat{b}\in B$, откуда $\hat{a}=\hat{b}\in A\cap B$, т. е. эти множества пересекаются, вопреки условию.

Б) Обозначим через $p(x)$ функцию Минковского множества C. Согласно предложению 1 п. 2.1.3 $p(x)$—неотрицательная выпуклая однородная и непрерывная в точке 0. Кроме того, $p(x)\leqslant 1$ для всех $x\in C$.

В) На подпространстве

$$L=\{x \mid x=\alpha c=\alpha(b_0-a_0),\ \alpha\in\mathbf{R}\}$$

определим линейный функционал l, полагая $\langle l, \alpha c\rangle=\alpha p(c)$. Тогда для $\alpha>0$ имеем $\langle l, \alpha c\rangle=\alpha p(c)=p(\alpha c)$, а для $\alpha\leqslant 0$ $\langle l, \alpha c\rangle=\alpha p(c)\leqslant 0\leqslant p(\alpha c)$, поскольку $p(\cdot)$ неотрицательна. Следовательно, для всех $x\in L$ выполняется неравенство $\langle l, x\rangle\leqslant p(x)$ и по теореме Хана—Банаха l можно продолжить до линейного функционала Λ такого, что

$$\langle\Lambda, \alpha c\rangle=\langle l, \alpha c\rangle=\alpha p(c),\quad \alpha\in\mathbf{R} \tag{3}$$

и

$$\langle\Lambda, x\rangle\leqslant p(x),\quad x\in X. \tag{4}$$

Поскольку $p(\cdot)$ непрерывна в 0, из (4) следует непрерывность функционала Λ (предложение 2 п. 2.1.3).

Г) Для любых $a\in A$ и $b\in B$ имеем

$$\langle\Lambda, a-b\rangle=\langle\Lambda, a-a_0-b+b_0\rangle+\langle\Lambda, a_0-b_0\rangle\leqslant$$
$$\leqslant p(a-a_0-b+b_0)+\langle l, (-1)(b_0-a_0)\rangle\leqslant 1-p(b_0-a_0),$$

поскольку $a-a_0-b+b_0\in C$, а на C функция $p(x)\leqslant 1$; кроме того, мы использовали (4). Но при $0<t\leqslant 1$ точка $(b_0-a_0)/t=c/t$ не может принадлежать множеству C, ибо C выпукло и содержит 0, а на $[0, c/t]$ лежит точка $c=b_0-a_0\notin C$. Поэтому

$$p(b_0-a_0)=\inf\left\{t>0 \,\middle|\, \frac{b_0-a_0}{t}\in C\right\}\geqslant 1. \tag{5}$$

Следовательно, $\langle\Lambda, a-b\rangle\leqslant 1-p(b_0-a_0)\leqslant 0$ для любых $a\in A$ и $b\in B$. В неравенстве $\langle\Lambda, a\rangle\leqslant\langle\Lambda, b\rangle$ $a\in A$ и $b\in B$ можно выбирать независимо, поэтому

$$\sup_{a\in A}\langle\Lambda, a\rangle\leqslant\inf_{b\in B}\langle\Lambda, b\rangle.$$

Кроме того, согласно (3) и (5) $\langle\Lambda, b_0-a_0\rangle = p(b_0-a_0) \geqslant 1$, так что $\Lambda \neq 0$. Значит, Λ разделяет A и B. ■

Замечание. В п. 1.3.3 мы уже доказали теорему отделимости в случае конечномерного пространства X. Сравнивая ее условия с условиями только что доказанной теоремы, мы видим, что в конечномерном случае можно опустить условие открытости A (то, что в п. 1.3.3 одно из множеств было точкой,—не существенно). Условия первой теоремы отделимости можно ослабить и требовать, чтобы $\operatorname{int} A \neq \varnothing$ и $(\operatorname{int} A) \cap B = \varnothing$. Однако совсем обойтись без существования внутренних точек хотя бы у одного из множеств нельзя.

Упражнения.

1. Пусть A—выпуклое множество в линейном топологическом пространстве X, $\operatorname{int} A \neq \varnothing$ и $x^* \in X^*$. Докажите, что

$$\sup_{x\in A} \langle x^*, x\rangle = \sup_{x\in \operatorname{int} A} \langle x^*, x\rangle \qquad (6)$$

2 Выведите из (6) теорему отделимости для случая выпуклых A и B, таких, что $\operatorname{int} A \neq \varnothing$, $B \neq \varnothing$, $B \cap \operatorname{int} A = \varnothing$.

3. Докажите, что в пространстве $X = l_2$ компактный эллипсоид

$$A = \left\{ x = [x_k] \in l_2 \;\middle|\; \sum_{k=1}^{\infty} k^2 x_k^2 \leqslant 1 \right\}$$

и луч

$$B = \{x = [x_k] \mid x_k = t/k, \quad t > 0\}$$

нельзя разделить никаким непрерывным линейным функционалом.

Вторая теорема отделимости. *Пусть X—локально выпуклое линейное топологическое пространство* [КФ, стр. 169], *A—непустое, замкнутое выпуклое подмножество в X и $\hat{x} \in X$—точка, не принадлежащая A. Тогда найдется ненулевой линейный непрерывный функционал, строго разделяющий A и $\hat{x}$.*

Доказательство. Поскольку $\hat{x} \notin A$ и A замкнуто, существует окрестность $V \ni \hat{x}$ такая, что $A \cap V = \varnothing$. Ввиду локальной выпуклости X существует выпуклая окрестность $B \subset V$ точки $\hat{x}$. Ясно, что $B \cap A = \varnothing$ и по первой теореме отделимости существует ненулевой функционал x^*, разделяющий A и B:

$$\sup_{x\in A} \langle x^*, x\rangle \leqslant \inf_{x\in B} \langle x^*, x\rangle.$$

Остается заметить, что

$$\inf_{x\in B} \langle x^*, x\rangle < \langle x^*, \hat{x}\rangle,$$

ибо нижняя грань ненулевого линейного функционала x^* не может достигаться во внутренней точке $\hat{x}$ множества B. ■

Определение 2. *Аннулятором* $A^\perp$ подмножества X линейного пространства X называется множество тех линейных функционалов l на X, для которых $\langle l, x\rangle = 0$ для всех $x \in A$.

Заметим, что $A^\perp$ всегда содержит $0 \in X^*$.

Лемма о нетривиальности аннулятора. *Пусть L — замкнутое подпространство локально выпуклого линейного топологического пространства X, причем $L \neq X$. Тогда аннулятор $L^\perp$ содержит ненулевой элемент $x^* \in X^*$.*

Доказательство. Возьмем любую точку $\hat{x} \notin L$. По второй теореме отделимости существует ненулевой функционал $x^* \in X^*$, строго разделяющий $\hat{x}$ и L (L — подпространство линейного пространства и, следовательно, выпукло):

$$\sup_{x \in L} \langle x^*, x\rangle < \langle x^*, \hat{x}\rangle. \tag{7}$$

Если бы существовало $x_0 \in L$, для которого $\langle x^*, x_0\rangle \neq 0$ то, поскольку $\alpha x_0 \in L$ для любого $\alpha \in \mathbf{R}$, было бы

$$\sup_{x \in L} \langle x^*, x\rangle \geqslant \sup_{\alpha \in \mathbf{R}} \langle x^*, \alpha x_0\rangle = +\infty,$$

вопреки (7). Следовательно, $\langle x^*, x\rangle \equiv 0$ на L, и потому $x^* \in L^\perp$. ■

2.1.5. Теорема Банаха об обратном операторе и лемма о правом обратном отображении. Вторым принципом линейного анализа является следующая теорема:

Теорема Банаха об обратном операторе. *Пусть X и Y — банаховы пространства, Λ: $X \to Y$ — непрерывный линейный оператор. Если Λ является мономорфизмом, т. е.*

$$\operatorname{Ker} \Lambda = \{x \mid \Lambda x = 0\} = \{0\},$$

и эпиморфизмом, т. е.

$$\operatorname{Im} \Lambda = \{y \mid y = \Lambda x, \quad x \in X\} = Y,$$

то Λ — изоморфизм между X и Y, т. е. существует линейный непрерывный оператор $M = \Lambda^{-1}$: $Y \to X$ такой, что

$$M\Lambda = I_X, \quad \Lambda M = I_Y.$$

Доказательство [КФ, стр. 225]. В нашем курсе будет многократно использовано такое следствие теоремы Банаха:

Лемма о правом обратном отображении. *Пусть X и Y — банаховы пространства, Λ — непрерывный линейный оператор из X в Y, являющийся эпиморфизмом. Тогда существует отображение $M\colon Y \to X$ (вообще говоря, разрывное и нелинейное), удовлетворяющее условиям*

$$\Lambda \circ M = I_Y,$$

$$\|M(y)\| \leqslant C\|y\| \text{ для некоторого } C > 0.$$

Отметим, что в силу второго условия M — непрерывное *в нуле* отображение.

Доказательство. Оператор Λ непрерывен, следовательно, его ядро $\operatorname{Ker}\Lambda$, являющееся прообразом точки нуль, будет замкнутым подпространством. По теореме о фактор-пространстве (п. 2.1.2) пространство $X/\operatorname{Ker}\Lambda$ банахово. Определим оператор $\tilde{\Lambda}\colon X/\operatorname{Ker}\Lambda \to Y$, положив $\tilde{\Lambda}\xi = \Lambda x$ на классе смежности $\xi = \pi x$. Это определение корректно: $x_1, x_2 \in \xi \Rightarrow \pi x_1 = \pi x_2 = \xi \Rightarrow x_1 - x_2 \in \operatorname{Ker}\Lambda \Rightarrow \Lambda x_1 = \Lambda x_2$. Оператор $\tilde{\Lambda}$ линеен: если $\pi x_i = \xi_i$, $i = 1, 2$, то

$$\tilde{\Lambda}(\alpha_1\xi_1 + \alpha_2\xi_2) = \Lambda(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1\Lambda x_1 + \alpha_2\Lambda x_2 = \alpha_1\tilde{\Lambda}\xi_1 + \alpha_2\tilde{\Lambda}\xi_2.$$

Оператор $\tilde{\Lambda}$ непрерывен: для любого $x \in \xi$

$$\|\tilde{\Lambda}\xi\| = \|\Lambda x\| \leqslant \|\Lambda\|\|x\|,$$

откуда

$$\|\tilde{\Lambda}\xi\| \leqslant \|\Lambda\| \inf_{x\in\xi}\|x\| = \|\Lambda\|\|\xi\|_{X/\operatorname{Ker}\Lambda}.$$

Наконец, оператор $\tilde{\Lambda}$ биективен (взаимно однозначен). Действительно, $\tilde{\Lambda}$ — эпиморфизм вместе с Λ: $\forall y \in Y\ \exists x \in X$, $\Lambda x = y \Rightarrow$ для $\xi = \pi(x)$, $\tilde{\Lambda}\xi = y$. С другой стороны, $\operatorname{Ker}\tilde{\Lambda} = 0$: $\tilde{\Lambda}\xi = 0$, $\pi x = \xi \Rightarrow \Lambda x = 0 \Rightarrow x \in \operatorname{Ker}\Lambda \Rightarrow \xi = 0$.

Итак, доказано, что оператор $\tilde{\Lambda}$ есть непрерывный линейный биективный оператор из $X/\operatorname{Ker}\Lambda$ в Y. По теореме Банаха об обратном операторе существует непрерывный обратный к $\tilde{\Lambda}$ оператор $\tilde{M} = \tilde{\Lambda}^{-1}$, $\tilde{M}\colon Y \to X/\operatorname{Ker}\Lambda$. Для элемента $\xi \in \tilde{M}y$ найдем в соответствии с (4) п. 2.1.2 такой элемент $x = x(\xi)$, что $x \in \xi$ и $\|x\| \leqslant 2\|\xi\|_{X/\operatorname{Ker}\Lambda}$.

Положим $M(y)=x(\xi)$. В итоге

$$(\Lambda\circ M)(y)=\Lambda x(\xi)=\tilde{\Lambda}\xi=\tilde{\Lambda}\tilde{M}y=y,$$

$$\|M(y)\|=\|x(\xi)\|\leqslant 2\|\xi\|_{X/\mathrm{Ker}\,\Lambda}=2\|\tilde{M}y\|_{X/\mathrm{Ker}\,\Lambda}\leqslant$$
$$\leqslant 2\|\tilde{M}\|\|y\|. \blacksquare$$

2.1.6. Лемма о замкнутости образа. Пусть X, Y и Z — банаховы пространства, A: $X\to Y$ и B: $X\to Z$ — линейные непрерывные операторы. Равенство

$$Cx=(Ax,\ Bx)$$

определяет линейный непрерывный оператор C: $X\to Y\times Z$.

Упражнение. Докажите, что если норма в $Y\times Z$ определяется равенством (1) п. 2.1.2, то $\|C\|=\max\{\|A\|,\ \|B\|\}$.

Лемма о замкнутости образа. *Если подпространство* $\mathrm{Im}\,A$ *замкнуто в* Y *и подпространство* $B\,\mathrm{Ker}\,A$ [1]) *замкнуто в* Z, *то подпространство* $\mathrm{Im}\,C$ *замкнуто в* $Y\times Z$.

Доказательство. Замкнутое подпространство $\tilde{Y}=\mathrm{Im}\,A$ банахова пространства Y само является банаховым пространством и A: $X\to\tilde{Y}$ — эпиморфизм. По лемме из п. 2.1.5 существует правое обратное к A отображение M: $\tilde{Y}\to X$. Пусть $(y,\ z)$ принадлежит замыканию образа оператора C. Это означает, что найдется последовательность $\{x_n\,|\,n\geqslant 1\}$ такая, что $y=\lim\limits_{n\to\infty}Ax_n$, $z=\lim\limits_{n\to\infty}Bx_n$. Положим

$$\xi_n=M(Ax_n-y)$$

($y\in\tilde{Y}$ как предел элементов $Ax_n\in Y$). Учитывая свойства M, получаем

$$A(x_n-\xi_n)=Ax_n-(A\circ M)(Ax_n-y)=y,$$
$$\|\xi_n\|=\|M(Ax_n-y)\|\leqslant C\|Ax_n-y\|\to 0.$$

Поэтому $B\xi_n\to 0$ и $\lim\limits_{n\to\infty}B(x_n-\xi_n)=\lim\limits_{n\to\infty}Bx_n=z$, т. е. z принадлежит замыканию множества

$$\Sigma=\{\zeta=Bx\,|\,Ax=y\}.$$

Это множество является сдвигом подпространства $B\,\mathrm{Ker}\,A$:

$$\zeta_1,\zeta_2\in\Sigma\Rightarrow\zeta_i=Bx_i,\quad Ax_i=y,\quad i=1,2\Rightarrow$$
$$\Rightarrow\zeta_2-\zeta_1=B(x_2-x_1),\quad A(x_2-x_1)=0\Rightarrow\zeta_2=\zeta_1+(\zeta_2-\zeta_1),$$
$$\zeta_1\in\Sigma,\quad \zeta_2-\zeta_1\in B\,\mathrm{Ker}\,A.$$

[1]) То есть образ ядра оператора A при отображении B.

Но $B\,\mathrm{Ker}\,A$ замкнуто по условию, значит, замкнуто и множество Σ. Следовательно, z, принадлежащий замыканию Σ, принадлежит и самому Σ, что означает существование элемента $\hat{x}$, для которого $A\hat{x}=y$, $B\hat{x}=z$. Итак, $(y, z)\in \mathrm{Im}\,C$. ■

2.1.7. Лемма об аннуляторе ядра регулярного оператора. *Пусть X, Y — банаховы пространства, $A\colon X\to Y$ — линейный непрерывный эпиморфизм. Тогда* $(\mathrm{Ker}\,A)^{\perp}=\mathrm{Im}\,A^*$.

Доказательство. А) Пусть $x^*\in \mathrm{Im}\,A^* \Leftrightarrow x^*=A^*y^*$. Тогда, если $x\in \mathrm{Ker}\,A$, то $\langle x^*, x\rangle=\langle A^*y^*, x\rangle=\langle y^*, Ax\rangle=0$, т. е. $\mathrm{Im}\,A^*\subset(\mathrm{Ker}\,A)^{\perp}$.

Б) Пусть $x^*\in(\mathrm{Ker}\,A)^{\perp}$, т. е. $Ax=0\Rightarrow\langle x^*, x\rangle=0$. Рассмотрим отображение $C\colon X\to \mathbf{R}\times Y$, $Cx=(\langle x^*, x\rangle, Ax)$. Образ ядра A при отображении x^* есть нуль, т. е. замкнутое множество, и по лемме о замкнутости образа (п. 2.1.6) $\mathrm{Im}\,C$ — замкнутое подпространство в $\mathbf{R}\times Y$. Оно не совпадает с $\mathbf{R}\times Y$, поскольку, например, $(1, 0)\notin \mathrm{Im}\,C$ (действительно, если $(\alpha, 0)=(\langle x^*, x\rangle, Ax)$, то $Ax=0\Rightarrow\alpha=$ $=(x^*, x)=0$).

По лемме о нетривиальности аннулятора замкнутого собственного подпространства (п. 2.1.4) существует ненулевой линейный непрерывный функционал $\Lambda\in(\mathrm{Im}\,C)^{\perp}\subset$ $\subset(\mathbf{R}\times Y)^*$ и, вспоминая лемму об общем виде линейного функционала на произведении пространств (п. 2.1.2), находим число $\hat{\lambda}_0\in\mathbf{R}^*=\mathbf{R}$ и элемент $\hat{y}^*\in Y^*$ такие, что

$$\langle\hat{\lambda}_0x^*+A^*\hat{y}^*, x\rangle=\hat{\lambda}_0\langle x^*, x\rangle+\langle\hat{y}^*, Ax\rangle=$$
$$=\langle\Lambda, Cx\rangle\equiv 0, \ \forall x.$$

Случай $\hat{\lambda}_0=0$ невозможен, ибо $\mathrm{Im}\,A=Y$ и, значит,

$$\langle\hat{y}^*, Ax\rangle\equiv 0\Rightarrow\hat{y}^*=0\Rightarrow\Lambda=0.$$

Следовательно, $x^*=A^*(-\hat{y}^*/\hat{\lambda}_0)$, т. е. $x^*\in \mathrm{Im}\,A^*$ и $(\mathrm{Ker}\,A)^{\perp}\subset$ $\subset \mathrm{Im}\,A^*$. ■

2.1.8. Абсолютно непрерывные функции. В теории оптимального управления постоянно приходится иметь дело с дифференциальными уравнениями вида

$$\dot{x}=\varphi(t, x, u(t)), \tag{1}$$

где $u(\cdot)$ — заданная функция, называемая *управлением*. При этом, хотя функция $\varphi(t, x, u)$ достаточно хорошая (непрерывная или даже гладкая), управление $u(\cdot)$ таковым, вообще говоря, не является. Оно может быть кусочно-непрерывным, а иногда даже удобно считать $u(\cdot)$ всего лишь измеримой функцией. Поэтому правая часть уравнения (1) может оказаться разрывной, и следует уточнить, что понимается под его решением.

Замечание. В этом пункте «измеримость», «интегрируемость», «почти всюду» и т. д. понимаются в обычном лебеговом смысле [КФ, гл. V, §§ 4, 5]. То же правило сохраняется и в остальной части книги, за редкими исключениями, оговариваемыми особо: например, в формулах следующего пункта интеграл понимается в смысле Стилтьеса или Лебега—Стилтьеса [КФ, гл. VI, § 6]. Если речь идет о векторнозначных или матричных функциях, то соответствующими свойствами должна обладать каждая их компонента. Например, функция $x(\cdot) = (x_1(\cdot), \ldots, x_n(\cdot))\colon [\alpha, \beta] \to \mathbf{R}^n$ измерима и интегрируема, если каждая из числовых функций $x_i(\cdot)\colon [\alpha, \beta] \to \mathbf{R}$ обладает тем же свойством, и по определению

$$\int_\alpha^\beta x(t)\,dt = \left(\int_\alpha^\beta x_1(t)\,dt, \ \ldots, \ \int_\alpha^\beta x_n(t)\,dt\right). \tag{2}$$

Наиболее естественно считать (мы так и поступим), что дифференциальное уравнение (1) вместе с начальным условием $x(t_0) = x_0$ эквивалентны интегральному уравнению

$$x(t) = x_0 + \int_{t_0}^t \varphi(s, x(s), u(s))\,ds. \tag{3}$$

Для того чтобы это было справедливо, $x(\cdot)$ должна быть неопределенным интегралом своей производной, т. е. необходимо, чтобы имела место формула Ньютона—Лейбница

$$x(t) - x(\tau) = \int_\tau^t \dot{x}(s)\,ds. \tag{4}$$

Функции $x(\cdot)$, для которых выполняется (4), в лебеговской теории интегрирования называются абсолютно непрерывными [КФ, гл. VI, § 4]. Важно, что этим функциям можно дать другое, эквивалентное и эффективно проверяемое определение.

5*

Определение. Пусть X — линейное нормированное пространство, $\Delta = [\alpha, \beta]$ — отрезок числовой прямой. Функция $x(\cdot)\colon \Delta \to X$ называется *абсолютно непрерывной*, если для всякого $\varepsilon > 0$ найдется такое $\delta > 0$, что для любой конечной системы попарно непересекающихся интервалов $(\alpha_k, \beta_k) \subset [\alpha, \beta]$, $k = 1, 2, \ldots, N$, сумма длин которых

$$\sum_{k=1}^{N} (\beta_k - \alpha_k) < \delta, \tag{5}$$

выполняется неравенство

$$\sum_{k=1}^{N} \| x(\beta_k) - x(\alpha_k) \| < \varepsilon. \tag{6}$$

Пример. Функция $x(\cdot)\colon \Delta \to X$, удовлетворяющая условию Липшица

$$\| x(t_1) - x(t_2) \| \leqslant K |t_1 - t_2|, \quad t_1, t_2 \in \Delta,$$

абсолютно непрерывна. Действительно, здесь $\delta = \varepsilon/K$. Непосредственно из определения видно, что абсолютно непрерывная функция равномерно непрерывна на Δ (берем $N = 1$).

Предложение 1. *Пусть X и Y — линейные нормированные пространства, Δ — отрезок числовой прямой.*

Если функции $x_i(\cdot)\colon \Delta \to X$, $i = 1, 2$, абсолютно непрерывны и $c_i \in \mathbf{R}$, то и функция $c_1 x_1(\cdot) + c_2 x_2(\cdot)$ абсолютно непрерывна.

Если функции $x(\cdot)\colon \Delta \to X$ и $A(\cdot)\colon \Delta \to \mathscr{L}(X, Y)$ абсолютно непрерывны, то и функция $A(\cdot)x(\cdot)\colon \Delta \to Y$ абсолютно непрерывна.

Если функция $\Phi(\cdot)\colon G \to Y$ удовлетворяет условию Липшица на множестве $\mathscr{K} \subset G$, а функция $x(\cdot)\colon \Delta \to G$ абсолютно непрерывна и $\operatorname{im} x(\cdot) = \{y \mid y = x(t),\ t \in \Delta\} \subset \mathscr{K}$, то функция $\Phi(x(\cdot))\colon \Delta \to Y$ абсолютно непрерывна.

Доказательство. Первое утверждение непосредственно следует из определения (см. также [КФ, стр. 343]). Для доказательства второго заметим, что $x(\cdot)$ и $A(\cdot)$, будучи непрерывными, ограничены, так что

$$\| x(t) \| \leqslant M_x, \quad \| A(t) \| \leqslant M_A, \quad t \in \Delta.$$

Теперь для любой системы непересекающихся интервалов $(\alpha_k, \beta_k)\subset\Delta$, $k=1, \ldots, N$, имеем

$$\sum_{k=1}^{N}\|A(\beta_k)x(\beta_k)-A(\alpha_k)x(\alpha_k)\|\leqslant$$
$$\leqslant\sum_{k=1}^{N}\|A(\beta_k)x(\beta_k)-A(\beta_k)x(\alpha_k)\|+$$
$$+\sum_{k=1}^{N}\|A(\beta_k)x(\alpha_k)-A(\alpha_k)x(\alpha_k)\|\leqslant$$
$$\leqslant\sum_{k=1}^{N}\|A(\beta_k)\|\|x(\beta_k)-x(\alpha_k)\|+\sum_{k=1}^{N}\|A(\beta_k)-A(\alpha_k)\|\|x(\alpha_k)\|\leqslant$$
$$\leqslant M_A\sum_{k=1}^{N}\|x(\beta_k)-x(\alpha_k)\|+M_x\sum_{k=1}^{N}\|A(\beta_k)-A(\alpha_k)\|.$$

Выберем $\delta>0$ так, чтобы из (5) следовали неравенства

$$\sum_{k=1}^{N}\|x(\beta_k)-x(\alpha_k)\|<\frac{\varepsilon}{2M_A},\quad \sum_{k=1}^{N}\|A(\beta_k)-A(\alpha_k)\|<\frac{\varepsilon}{2M_x}.$$

Тогда

$$\sum_{k=1}^{N}\|A(\beta_k)x(\beta_k)-A(\alpha_k)x(\alpha_k)\|<\frac{\varepsilon}{2M_A}M_A+\frac{\varepsilon}{2M_x}M_x=\varepsilon,$$

так что $A(\cdot)x(\cdot)$ абсолютно непрерывна.

Аналогично, если

$$\|\Phi(x_1)-\Phi(x_2)\|\leqslant K\|x_1-x_2\|,\ \forall x_1,\ x_2\in\mathcal{K},$$

то

$$\sum_{k=1}^{N}\|\Phi(x(\beta_k))-\Phi(x(\alpha_k))\|\leqslant\sum_{k=1}^{N}\|x(\beta_k)-x(\alpha_k)\|<\varepsilon,$$

если $\delta>0$ выбрано так, чтобы из (5) следовало (6) с заменой ε на ε/k. ■

Основную теорему этого пункта мы сформулируем только для конечномерного X, чтобы не определять операцию интегрирования в произвольном нормированном пространстве (для $X=\mathbf{R}^n$ достаточно иметь в виду (2)). Только этот случай нам в дальнейшем и понадобится.

Теорема Лебега [КФ, стр 345] *Если функция* $x(\cdot)\colon\ \Delta\to\mathbf{R}^n$ *абсолютно непрерывна, то она дифферен-*

цируема почти всюду, ее производная $\dot{x}(\cdot)$ *интегрируема на* Δ *и для всех* t, $\tau \in \Delta$ *имеет место равенство* (4).

Если функция $\xi(\cdot)\colon \Delta \to \mathbf{R}^n$ *интегрируема на* Δ *и* $\tau \in \Delta$, *то функция* $x(t) = \int\limits_{\tau}^{t} \xi(s)\,ds$ *абсолютно непрерывна и* $\dot{x}(t) = \xi(t)$ *почти всюду.*

В соответствии с этой теоремой мы будем в дальнейшем функцию $x(\cdot)$ называть решением дифференциального уравнения (1), если она абсолютно непрерывна и удовлетворяет (1) почти всюду. Действительно, по теореме Лебега этого достаточно, чтобы из (1) следовало (3). Обратно, если $x(\cdot)$ удовлетворяет (3) (т. е., в частности, $t \to \varphi(t, x(t), u(t))$ — интегрируемая функция), то по теореме Лебега $x(\cdot)$ абсолютно непрерывна и почти всюду имеет место (1).

В заключение этого пункта приведем еще одно утверждение, числовой аналог которого хорошо известен в лебеговской теории интегрирования, и доказательство которого предоставляется читателю в качестве упражнения.

Предложение 2. *Измеримая функция* $x(\cdot)\colon \Delta \to \mathbf{R}^n$ *интегрируема тогда и только тогда, когда* $|x(\cdot)|$ *интегрируема. При этом*

$$\left| \int\limits_{\Delta} x(t)\,dt \right| \leqslant \int\limits_{\Delta} |x(t)|\,dt. \tag{7}$$

2.1.9. Теорема Рисса об общем виде линейного функционала в пространстве *C*. Формула Дирихле. Определения функции ограниченной вариации и интеграла Стилтьеса мы будем предполагать известными [КФ, гл. VI §§ 2 и 6]. Функцию ограниченной вариации $v(\cdot)\colon [\alpha, \beta] \to \mathbf{R}$ будем называть *канонической*, если она непрерывна справа во всех точках интервала (α, β) и $v(\alpha) = 0$.

Теорема Ф. Рисса [КФ, стр. 369]. *Каждому непрерывному линейному функционалу* $x^* \in C([\alpha, \beta])^*$ *соответствует каноническая функция ограниченной вариации* $v(\cdot)\colon [\alpha, \beta] \to \mathbf{R}$ *такая, что для всех* $x(\cdot) \in C([\alpha, \beta])$

$$\langle x^*, x(\cdot) \rangle = \int\limits_{\alpha}^{\beta} x(t)\,dv(t), \tag{1}$$

и такое представление единственно: если для всех $x(\cdot)\in$ $\in C([\alpha, \beta])$ *и для канонической* $v(\cdot)$

$$\int_\alpha^\beta x(t)\,dv(t)=0,$$

то $v(t)\equiv 0$.

Эта теорема легко обобщается на векторный случай. Пусть $e_1=(1, 0, \ldots, 0)$, $e_2=(0, 1, \ldots, 0)$, ... — единичные векторы стандартного базиса в $\mathbf{R}^n$. Произвольный элемент $x(\cdot)=(x_1(\cdot), \ldots, x_n(\cdot))\in C([\alpha, \beta], \mathbf{R}^n)$ представляется в виде

$$x(\cdot)=\sum_{k=1}^n x_k(\cdot)\,e_k.$$

Если теперь $x^*\in C([\alpha, \beta], \mathbf{R}^n)^*$, то

$$\langle x^*, x(\cdot)\rangle=\sum_{k=1}^n \langle x^*, x_k(\cdot)\,e_k\rangle=\sum_{k=1}^n \langle x_k^*, x_k(\cdot)\rangle,$$

где функционалы $x_k^*\in C([\alpha, \beta])^*$ определяются равенствами $\langle x_k^*, \xi(\cdot)\rangle=\langle x^*, \xi(\cdot)\,e_k\rangle$. Применяя к x_k^* теорему Рисса, получаем представление

$$\langle x^*, x(\cdot)\rangle=\sum_{k=1}^n \int_\alpha^\beta x_k(t)\,dv_k(t). \tag{2}$$

Набор функций ограниченной вариации $v(\cdot)=(v_1(\cdot), \ldots$ $\ldots, v_n(\cdot))$ естественно назвать векторнозначной функцией ограниченной вариации (значения — вектор-строчки) $v(\cdot)\colon [\alpha, \beta]\to\mathbf{R}^{n*}$. Тогда формула (2) может быть переписана в виде

$$\langle x^*, x(\cdot)\rangle=\int_\alpha^\beta dv(t)\,x(t), \tag{3}$$

совпадающем с (1) с точностью до порядка сомножителей. Как и в теореме Рисса, соответствие между x^* и $v(\cdot)$ будет взаимнооднозначным при соблюдении условия «каноничности» функции $v(\cdot)$.

Каждая функция ограниченной вариации $v(\cdot)$: $[\alpha, \beta]\to\mathbf{R}$ определяет обобщенную меру (или «заряд» [КФ, гл. VI § 5]). Интеграл Стилтьеса (1) и есть интеграл по этой мере. Аналогично векторнозначная функция ограниченной вариации $v(\cdot)\colon [\alpha, \beta]\to\mathbf{R}^{n*}$ определяет на

$[\alpha, \beta]$ меру с векторными значениями, интегралом по которой является (3). Меру, соответствующую функции ограниченной вариации $v(t)$, принято обозначать $dv(t)$.

Если $v_1(\cdot)$ и $v_2(\cdot)$—две функции ограниченной вариации на отрезке $[\alpha, \beta]$, то на квадрате $[\alpha, \beta]\times[\alpha, \beta]$ определено произведение мер $dv_1(\cdot)\times dv_2(\cdot)$ и справедлива теорема Фубини [КФ, стр. 317]. В частности, имеет место известная формула для перестановки пределов интегрирования («формула Дирихле»):

$$\int_\alpha^\beta \left\{\int_\alpha^t f(t, s)\, dv_1(s)\right\} dv_2(t) = \int_\alpha^\beta \left\{\int_s^\beta f(t, s)\, dv_2(t)\right\} dv_1(s). \quad (4)$$

При обобщении формулы (4) на векторнозначные функции (а это нам далее понадобится) следует быть внимательным к порядку сомножителей. Например, если

$$f(\cdot, \cdot)\colon [\alpha, \beta]\times[\alpha, \beta] \to \mathscr{L}(\mathbf{R}^n, \mathbf{R}^n),$$
$$v_1(\cdot)\colon [\alpha, \beta] \to \mathbf{R}^n, \quad v_2(\cdot)\colon [\alpha, \beta] \to \mathbf{R}^{n*},$$

то (4) следует писать так:

$$\int_\alpha^\beta dv_2(t) \left\{\int_\alpha^t f(t, s)\, dv_1(s)\right\} = \int_\alpha^\beta \left\{\int_s^\beta dv_2(t) f(t, s)\right\} dv_1(s). \quad (5)$$

§ 2.2. Основы дифференциального исчисления в линейных нормированных пространствах

Дифференциальное исчисление—это тот раздел математического анализа, в котором изучаются локальные свойства гладких отображений. На интуитивном уровне гладкость отображения означает возможность локально аппроксимировать его подходящим аффинным отображением—дифференциалом, или, более общо, полиномиальным отображением, и получить отсюда информацию о свойствах изучаемого отображения. Наряду с выпуклым анализом дифференциальное исчисление является основным аппаратом теории экстремальных задач.

Основы дифференциального исчисления скалярных функций на прямой были заложены Ньютоном и Лейбницем. Задачи вариационного исчисления потребовали распространить понятие производной и дифференциала на отображения многомерных и бесконечномерных пространств. Это было сделано Лагранжем в XVIII в. и в наше время более полно Фреше, Гато, Леви и другими математиками.

2.2.1. Производная по направлению, первая вариация, производные Гато и Фреше, строгая дифференцируемость. Хорошо известно, что для вещественных функций одного вещественного переменного два определения — существование конечного предела

$$\lim_{h \to 0} \frac{F(\hat{x}+h)-F(\hat{x})}{h} \tag{1}$$

и возможность асимптотического разложения при $h \to 0$

$$F(\hat{x}+h) = F(\hat{x}) + F'(\hat{x})h + o(h) \tag{2}$$

— приводят к одному и тому же понятию дифференцируемости. Для функций нескольких переменных, а тем более для функций с бесконечномерной областью определения дело обстоит не так просто. Определение (1) и обобщающее его определение частной производной приводят к понятиям производной по направлению, первой вариации и производной Гато. В то же время обобщение определения (2) приводит к производной Фреше и строгой дифференцируемости.

Пусть X и Y — линейные нормированные пространства, U — окрестность точки $\hat{x}$ в X, F — отображение из U в Y.

Определение 1. Предел

$$\lim_{\lambda \downarrow 0} \frac{F(\hat{x}+\lambda h)-F(\hat{x})}{\lambda} \tag{3}$$

в предположении, что он существует, называется *производной F в точке $\hat{x}$ по направлению* h и обозначается $F'(\hat{x}; h)$ (а также у других авторов $(\nabla_h F)(\hat{x})$, $D_h F(\hat{x})$, $\partial_h F(\hat{x})$). Для вещественных функций $(Y = \mathbf{R})$ мы будем понимать (1) несколько расширенно, допуская в качестве предела $-\infty$ и $+\infty$.

Определение 2. Пусть для любого $h \in X$ существует производная по направлению $F'(\hat{x}; h)$. Отображение $\delta_+ F(\hat{x}; \cdot): X \to Y$, определяемое формулой $\delta_+ F(x; h) = F'(x; h)$, называется *первой вариацией отображения F в точке* $\hat{x}$ (про функцию F мы говорим в этом случае, что она имеет в точке $\hat{x}$ первую вариацию).

Если $\delta_+(\hat{x}, -h) = -\delta_+(\hat{x}, h)$ при всех h, иначе говоря, если для любого $h \in X$ существует предел

$$\delta F(\hat{x}, h) = \lim_{\lambda \to 0} \frac{F(\hat{x} + \lambda h) - F(\hat{x})}{\lambda},$$

то отображение $h \mapsto \delta F(\hat{x}; h)$ называют *первой вариацией по Лагранжу* отображения F в точке $\hat{x}$ (см. п. 1.4.1).

Определение 3. Пусть F имеет в точке $\hat{x}$ первую вариацию и при этом существует линейный непрерывный оператор $\Lambda \in \mathscr{L}(X, Y)$ такой, что $\delta_+ F(x; h) \equiv \Lambda h$.

Тогда оператор Λ называется *производной Гато* отображения F в точке x и обозначается $F'_\Gamma(x)$.

Таким образом, $F'_\Gamma(x)$ — это такой элемент из $\mathscr{L}(X, Y)$, что для любого $h \in X$ при $\lambda \downarrow 0$ выполняется соотношение

$$F(x + \lambda h) = F(x) + \lambda F'_\Gamma h + o(\lambda). \tag{4}$$

Несмотря на внешнее сходство соотношений (2) и (4), между ними есть глубокое различие. Как мы увидим ниже, уже для $X = \mathbf{R}^2$ функция, дифференцируемая по Гато, может не быть непрерывной. Дело в том, что в разложении (4) $o(\lambda)$ не обязана быть равномерной по h.

Определение 4. Пусть в окрестности точки $\hat{x}$ отображение F можно представить в виде

$$F(\hat{x} + h) = F(\hat{x}) + \Lambda(h) + \alpha(h)\|h\|, \tag{5}$$

где $\Lambda \in \mathscr{L}(X, Y)$ и

$$\lim_{\|h\| \to 0} \|\alpha(h)\| = \|\alpha(0)\| = 0. \tag{6}$$

Тогда отображение $F(\cdot)$ называют *дифференцируемым по Фреше в точке* $\hat{x}$ и пишут $F \in D^1(\hat{x})$. Оператор Λ называется *производной Фреше* (или просто производной) отображения F в точке $\hat{x}$ и обозначается $F'(\hat{x})$.

Соотношения (5) и (6) можно записать еще и так:

$$F(\hat{x} + h) = F(\hat{x}) + F'(\hat{x})[h] + o(\|h\|). \tag{7}$$

Отсюда легко следует, что производная Фреше определена однозначно (для производной Гато это очевидно, поскольку однозначно определение производных по направлению), ибо равенство $\|\Lambda_1 h - \Lambda_2 h\| = o(\|h\|)$ для операторов $\Lambda_i \in \mathscr{L}(X, Y)$, $i = 1, 2$, возможно лишь при $\Lambda_1 = \Lambda_2$.

Наконец, на языке $\varepsilon-\delta$ (5) и (6) формулируются так: для любого $\varepsilon>0$ найдется $\delta>0$, при котором для всех h таких, что $\|h\|<\delta$, выполняется неравенство

$$\|F(x+h)-F(x)-\Lambda h\|\leqslant\varepsilon\|h\|. \tag{8}$$

Это делает естественным дальнейшее усиление.

Определение 5. Отображение F называют *строго дифференцируемым в точке* $\hat{x}$ (и пишут $F\in SD^1(\hat{x})$), если существует такой оператор $\Lambda\in\mathscr{L}(X, Y)$, при котором для всякого $\varepsilon>0$ найдется такое $\delta>0$, что для всех x_1 и x_2, удовлетворяющих неравенствам $\|x_1-\hat{x}\|<\delta$, $\|x_2-\hat{x}\|<\delta$, выполняется неравенство

$$\|F(x_1)-F(x_2)-\Lambda(x_1-x_2)\|\leqslant\varepsilon\|x_1-x_2\|. \tag{9}$$

Полагая в (9) $x_2=\hat{x}$ и $x_1=\hat{x}+h$, получаем (8), так что строго дифференцируемая функция дифференцируема также и по Фреше и $\Lambda=F'(\hat{x})$.

Производная $F'(\hat{x})$ (Гато, Фреше или строгая) по определению является линейным отображением из X в Y. Значение этого отображения на векторе $h\in X$ мы будем часто обозначать $F'(\hat{x})[h]$. Для числовых функций одного переменного (когда $X=Y=\mathbf{R}$) пространство $\mathscr{L}(X, Y)=$ $=\mathscr{L}(\mathbf{R}, \mathbf{R})$ линейных непрерывных отображений из $\mathbf{R}$ в $\mathbf{R}$ естественно отождествляется с $\mathbf{R}$ (линейной функции $y=kx$ сопоставляется ее угловой коэффициент k). Именно в этом смысле в элементарном анализе производная в точке — это число (угловой коэффициент касательной), а соответствующее линейное отображение из $\mathbf{R}$ в $\mathbf{R}$ — это дифференциал $dy=F'(\hat{x})\,dx$ (здесь dx и dy — элементы одномерного векторного пространства $\mathbf{R}$).

Хорошо известно также, что для функций одного переменного определение 3 (или совпадающее с ним в этом случае определение первой вариации по Лагранжу) и определение 4 эквивалентны (функция имеет производную в точке тогда и только тогда, когда в этой точке она имеет дифференциал) и приводят к одному и тому же понятию производной, введенному по сути дела Ньютоном и Лейбницем. Определение 2 применимо здесь к функциям, имеющим в точке $\hat{x}$ обе односторонние производные, не обязательно совпадающие. Для функций нескольких переменных ($1<\dim X<\infty$) в элементарном анализе употребляются определение 1 (если h — единичный вектор, то (3) — это обычное определение производной по направлению; если h — базисный вектор оси OX_k, то заменив в (3) $\lambda \downarrow 0$ на $\lambda \to 0$, получим определение частной производной $\partial F/\partial x_k$) и определение 4 (существование полного дифференциала).

Обобщение понятия производной на бесконечномерные пространства было стимулировано задачами вариационного исчисления. Определение первой вариации $\delta F(\hat{x}; h)$, ее обозначение и название дал Лагранж. На языке функционального анализа это же определение было дано Гато, а потому первую вариацию по Лагранжу иногда называют дифференциалом Гато. Требование линейности в определении производной Гато стало общепринятым после работ П. Леви. Наиболее употребительно в анализе (как конечномерном, так и бесконечномерном) определение производной Фреше, однако для многих целей (в частности, и для наших дальнейших) существования производной Фреше в точке не хватает, надо чуть больше. Это и привело к понятию строгой дифференцируемости.

Предложение 1. *Между определениями* 1—5 *и непрерывностью функции имеют место следующие импликации:*

$$\begin{tikzcd} 1 & 2 \arrow[l, Rightarrow] & 3 \arrow[l, Rightarrow] & 4 \arrow[l, Rightarrow] \arrow[d, Rightarrow] & & 5 \arrow[ll, Rightarrow] \arrow[d, Rightarrow] \\ & & & \text{непрерывность в точке } \hat{x} & & \text{непрерывность в окрестности точки } \hat{x} \end{tikzcd}$$

ни одна из которых не может быть обращена.

Доказательство. Положительная часть доказательства (т. е. существование импликаций) немедленно следует из определений, а отрицательная (невозможность обращений) подтверждается набором контрпримеров, подробный разбор которых предоставляется читателю в качестве упражнения.

1) 2 не влечет 3: $f_1\colon \mathbf{R} \to \mathbf{R}$, $f_1(x) = |x|$.

В точке $x = 0$ вариация нелинейна. Тот же пример показывает, что из непрерывности в точке не следует дифференцируемости по Фреше или Гато.

2) 3 не влечет 4: $f_2\colon \mathbf{R}^2 \to \mathbf{R}$,

$$f_2(x_1, x_2) = \begin{cases} 1, & x_1 = x_2^2,\ x_2 > 0, \\ 0 & \text{в остальных точках.} \end{cases}$$

Этот пример показывает, что функция может быть дифференцируемой по Гато (в точке (0, 0)), не будучи непрерывной.

3) 4 не влечет 5: $f_3\colon \mathbf{R} \to \mathbf{R}$,

$$f_3(x) = \begin{cases} x^2, & x \text{ рационально}, \\ 0, & x \text{ иррационально}. \end{cases}$$

В точке $x=0$ эта функция дифференцируема (по Фреше), но не строго дифференцируема. ■

Упражнения.

1. Покажите, что функция $f_4\colon \mathbf{R}^2 \to \mathbf{R}$, определяемая в полярных координатах на $\mathbf{R}^2$ равенством

$$f_4(x) = r\cos 3\varphi \qquad (x = (x_1, x_2) = (r\cos\varphi, r\sin\varphi)),$$

имеет в точке (0, 0) первую вариацию по Лагранжу, но не дифференцируема по Гато.

2. Если функция $F \in SD^1(\hat{x})$, то в некоторой окрестности точки $\hat{x}$ она удовлетворяет условию Липшица.

3. Если числовая ($X = Y = \mathbf{R}$) функция $F \in SD^1(\hat{x})$, то $F'(x)$ существует почти во всех точках некоторой окрестности точки $\hat{x}$.

Предложение 2. *Если отображения $F_i\colon U \to Y$, $i = 1, 2$, и $A\colon U \to \mathscr{L}(Y, Z)$, где Z — нормированное линейное пространство, дифференцируемы в смысле одного из определений 1—5 (одного и того же для всех трех отображений), то: для любых $\alpha_i \in \mathbf{R}$, $i = 1, 2$, отображение $F = \alpha_1 F_1 + \alpha_2 F_2$ в точке $\hat{x}$ дифференцируемо в том же смысле, причем*

$$F'(\hat{x}) = \alpha_1 F_1'(\hat{x}) + \alpha_2 F_2'(\hat{x}),$$

или

$$F'(\hat{x};\ h) = \alpha_1 F_1'(\hat{x};\ h) + \alpha_2 F_2'(\hat{x};\ h);$$

отображение

$$\Phi = AF_i$$

в точке $\hat{x}$ дифференцируемо в том же смысле, причем

$$\Phi'(\hat{x};\ h) = A'(\hat{x};\ h)\,F_i(\hat{x}) + A(\hat{x})\,F_i'(\hat{x};\ h).$$

Доказательство непосредственно получается из определений. ■

В частном случае, когда $X=\mathbf{R}$ и $A\colon U \longrightarrow \mathscr{L}(Y, \mathbf{R})=Y^*$, $\Phi(x)=\langle A(x),\ F_\iota(x)\rangle$ — числовая функция, и последнюю формулу можно переписать так:

$$\langle A(x),\ F_\iota(x)\rangle'\big|_{x=\hat{x}}=\langle A'(\hat{x}),\ F_\iota(\hat{x})\rangle+\langle A(\hat{x}),\ F'_\iota(\hat{x})\rangle,$$

что вполне соответствует обычной формуле дифференцирования.

Приведем два простейших примера вычисления производных.

1) Аффинное отображение. Отображение $A\colon X \longrightarrow Y$ одного линейного пространства в другое называется *аффинным*, если существует линейное отображение $\Lambda\colon X \longrightarrow Y$ и константа $a \in Y$ такие, что

$$A(x)=\Lambda x+a.$$

Если X и Y — нормированные пространства, а $\Lambda \in \mathscr{L}(X, Y)$, то отображение A строго дифференцируемо в любой точке x и при этом

$$A'(x)=\Lambda.$$

Это утверждение проверяется непосредственно. В частности, если A линейно ($a=0$), то $A'(x)[h]=A(h)$, а производная постоянного отображения ($\Lambda=0$) равна нулю.

Упражнение 1. Докажите обратное. если производная (достаточно в смысле Гато) отображения $A\colon X \longrightarrow Y$ существует в каждой точке $x \in X$ и для всех x одна и та же, то A — аффинное отображение.

2) Полилинейное отображение. Пусть X и Y — линейные топологические пространства, а $\mathscr{L}^n(X, Y)$ — линейное пространство непрерывных полилинейных отображений декартова произведения $X^n=\underbrace{X\times\ldots\times X}_{n \text{ раз}}$ в Y. Напомним, что отображение $\Pi(x_1, \ldots, x_n)$ называется *полилинейным*, если при каждом i отображение

$$x_\iota \longmapsto \Pi(\hat{x}_1, \ldots, \hat{x}_{\iota-1},\ x_\iota,\ \hat{x}_{\iota+1}, \ldots, \hat{x}_n)$$

линейно. Если X и Y нормированы, то непрерывность Π равносильна ограниченности, т. е. конечности числа

$$\|\Pi\|=\sup_{\|x_1\|\leqslant 1, \ldots, \|x_n\|\leqslant 1}\|\Pi(x_1, \ldots, x_n)\|. \tag{10}$$

В этом случае $\mathscr{L}^n(X, Y)$ является нормированным линейным пространством с нормой (10) и выполняется неравенство

$$\|\Pi(x_1, \ldots, x_n)\| \leqslant \|\Pi\| \|x_1\| \ldots \|x_n\|. \qquad (11)$$

Функция $Q_n(x) = \Pi(x, \ldots, x)$ называется *формой степени* n (при $n=2$ квадратичной, при $n=3$ — тернарной), отвечающей полилинейному отображению Π. Из определений вытекает, что

$$Q_n(\hat{x}+h) = Q_n(\hat{x}) + \sum_{i=1}^{n} \Pi(\hat{x}, \ldots, h, \ldots, \hat{x}) + o(\|h\|^2), \qquad (12)$$

а потому $Q_n(\cdot)$ дифференцируема по Фреше в каждой точке $\hat{x}$ и

$$Q_n'(\hat{x})[h] = \sum_{i=1}^{n} \Pi(\hat{x}, \ldots, h, \ldots, \hat{x}) \qquad (13)$$

(в обеих формулах (12) и (13) в i-м члене суммы аргумент h стоит на i-м месте, а остальные аргументы равны $\hat{x}$). Если отображение Π симметрично, т. е. $\Pi(x_1, \ldots, x_n)$ не меняется при любой перестановке аргументов x_i, то (13) превращается в

$$Q_n'(\hat{x})[h] = n\Pi(h, \hat{x}, \ldots, \hat{x}).$$

В частности, если X — вещественное гильбертово пространство со скалярным произведением $(\cdot|\cdot)$, $A \in \mathscr{L}(X, X)$ и $\Pi(x_1, x_2) = (Ax_1 | x_2)$, то $Q_2(x) = (Ax|x)$ — квадратичная форма оператора A — дифференцируема и

$$Q_2'(\hat{x})[h] = (A\hat{x}|h) + (Ah|\hat{x}) = (A\hat{x} + A^*\hat{x}|h),$$

так что $Q_2'(\hat{x})$ естественно отождествляется с вектором $A\hat{x} + A^*\hat{x}$. Например, для $Q_2(x) = {}^1/_2(x|x) = {}^1/_2\|x\|^2$ производная $Q'(x) = x$.

Упражнение 2 Докажите, что $Q_n(x) = \Pi(x, \ldots, x)$, где $\Pi \in \mathscr{L}^n(X, Y)$ строго дифференцируема при всех x

Более сложный пример доставляет

Предложение 3. *Пусть $U \subset \mathscr{L}(X, Y)$ — множество непрерывных линейных операторов $A\colon X \to Y$, для которых существует обратный оператор $A^{-1} \in \mathscr{L}(Y, X)$. Если хотя бы одно из пространств X или Y банахово, то U открыто и функция $\Phi(A) = A^{-1}$ дифференцируема по Фреше в любой точке $A \in U$, причем $\Phi'(\hat{A})[H] = -\hat{A}^{-1}H\hat{A}^{-1}$.*

Доказательство. Если пространство X банахово, то ряд $\sum_{k=0}^{\infty}(-\hat{A}^{-1}H)^k$ сходится в $\mathscr{L}(X, X)$ при $\|H\|<\|\hat{A}^{-1}\|^{-1}$ и, как нетрудно проверить непосредственно,

$$(\hat{A}+H)\left[\sum_{k=0}^{\infty}(-\hat{A}^{-1}H)^k\hat{A}^{-1}\right]=I_Y,$$

$$\left[\sum_{k=0}^{\infty}(-\hat{A}^{-1}H)^k\hat{A}^{-1}\right](\hat{A}+H)=I_X.$$

Следовательно,

$$(\hat{A}+H)^{-1}=\sum_{k=0}^{\infty}(-\hat{A}^{-1}H)^k\hat{A}^{-1} \tag{14}$$

при $\|H\|<\|\hat{A}^{-1}\|^{-1}$, так что $\mathring{B}(\hat{A}, \|\hat{A}^{-1}\|^{-1})\subset U$ и U открыто.

Если пространство Y банахово, то (14) следует заменить на

$$(\hat{A}+H)^{-1}=\hat{A}^{-1}\sum_{k=0}^{\infty}(-H\hat{A}^{-1})^k.$$

Далее, из (14) имеем

$$\|(\hat{A}+H)^{-1}-\hat{A}^{-1}+\hat{A}^{-1}H\hat{A}^{-1}\|=\left\|\sum_{k=2}^{\infty}(-\hat{A}^{-1}H)^k\hat{A}^{-1}\right\|\leqslant$$

$$\leqslant\sum_{k=2}^{\infty}(\|\hat{A}^{-1}\|\|H\|)^k\|\hat{A}^{-1}\|=\frac{\|\hat{A}^{-1}\|^3\|H\|^2}{1-\|\hat{A}^{-1}\|\|H\|}=o(\|H\|)$$

при $H\to 0$, а следовательно, функция $\Phi(A)=\hat{A}^{-1}$ дифференцируема и ее производная Фреше $\Phi'(\hat{A})[\cdot]=-\hat{A}^{-1}(\cdot)\hat{A}^{-1}$. ■

Упражнение 3. Докажите, что $\Phi(A)$ строго дифференцируема в любой точке $\hat{A}\in U$.

2.2.2. Теорема о суперпозиции дифференцируемых отображений.

Теорема о суперпозиции. *Пусть X, Y, Z — нормированные пространства; U — окрестность точки $\hat{x}$ в X, V — окрестность точки $\hat{y}$ в Y; φ: $U\to V$, $\varphi(\hat{x})=\hat{y}$; ψ: $V\to Z$; $f=\psi\circ\varphi$: $U\to Z$ — суперпозиция отображений φ и ψ.*

Если ψ дифференцируема по Фреше в точке $\hat{y}$, а φ в точке $\hat{x}$ дифференцируема по Фреше (дифференцируема по Гато, имеет первую вариацию или имеет производную по направлению h), то f обладает в точке $\hat{x}$ тем же свойством, что и φ, и при этом

$$f'(\hat{x}) = \psi'(\hat{y}) \circ \varphi'(\hat{x}) \tag{1}$$

или

$$f'(\hat{x};\ h) = \psi'(\hat{y})[\varphi'(\hat{x};\ h)]. \tag{2}$$

Если ψ строго дифференцируема в $\hat{y}$, а φ строго дифференцируема в $\hat{x}$, то f строго дифференцируема в $\hat{x}$.

Доказательство. Рассмотрим подробно два крайних случая — производную по направлению и строгую дифференцируемость.

А) По определению производной Фреше

$$\psi(y) = \psi(\hat{y}) + \psi'(\hat{y})[y - \hat{y}] + \alpha(y)\|y - \hat{y}\|,$$

где

$$\lim_{y \to \hat{y}} \alpha(y) = \alpha(\hat{y}) = 0.$$

Если существует

$$\varphi'(\hat{x};\ h) = \lim_{\lambda \downarrow 0} \frac{\varphi(\hat{x} + \lambda h) - \varphi(\hat{x})}{\lambda} = \lim_{\lambda \downarrow 0} \frac{\varphi(\hat{x} + \lambda h) - \hat{y}}{\lambda},$$

то

$$\lim_{\lambda \downarrow 0} \frac{f(\hat{x} + \lambda h) - f(\hat{x})}{\lambda} = \lim_{\lambda \downarrow 0} \frac{\psi(\varphi(\hat{x} + \lambda h)) - \psi(\hat{y})}{\lambda} =$$

$$= \lim_{\lambda \downarrow 0} \frac{\psi'(\hat{y})[\varphi(\hat{x} + \lambda h) - \hat{y}] + \alpha(\varphi(\hat{x} + \lambda h))\|\varphi(\hat{x} + \lambda h) - \hat{y}\|}{\lambda} =$$

$$= \psi'(\hat{y})\left[\lim_{\lambda \downarrow 0} \frac{\varphi(\hat{x} + \lambda h) - \hat{y}}{\lambda}\right] +$$

$$+ \lim_{\lambda \downarrow 0} \alpha(\varphi(\hat{x} + \lambda h)) \lim_{\lambda \downarrow 0} \left\| \frac{\varphi(\hat{x} + \lambda h) - \hat{y}}{\lambda} \right\| =$$

$$= \psi'(\hat{y})[\varphi'(\hat{x};\ h)] + \alpha(\hat{y})\|\varphi'(\hat{x};\ h)\| =$$

$$= \psi'(\hat{y})[\varphi'(\hat{x};\ h)],$$

что и доказывает (2).

Б) Обозначим для краткости $A_1 = \varphi'(\hat{x})$, $A_2 = \psi'(\hat{y})$. По определению строгой дифференцируемости для любых

$\varepsilon_1 > 0$, $\varepsilon_2 > 0$ найдутся такие $\delta_1 > 0$, $\delta_2 > 0$, что

$$\|x_1-\hat{x}\| < \delta_1, \|x_2-\hat{x}\| < \delta_1 \Rightarrow \|\varphi(x_1)-\varphi(x_2)-A_1(x_1-x_2)\| \leqslant \leqslant \varepsilon_1\|x_1-x_2\|, \quad (3)$$

$$\|y_1-\hat{y}\| < \delta_2, \|y_2-\hat{y}\| < \delta_2 \Rightarrow \|\psi(y_1)-\psi(y_2)-A_2(y_1-y_2)\| \leqslant \leqslant \varepsilon_2\|y_1-y_2\|. \quad (4)$$

Для любого $\varepsilon > 0$ подберем $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$ так, чтобы выполнялось неравенство

$$\varepsilon_1\|A_2\|+\varepsilon_2\|A_1\|+\varepsilon_1\varepsilon_2 < \varepsilon;$$

по этим ε_1, ε_2 найдем $\delta_1 > 0$, $\delta_2 > 0$ так, чтобы имели место соотношения (3) и (4), и наконец положим

$$\delta = \min\left(\delta_1, \frac{\delta_2}{\varepsilon_1+\|A_1\|}\right).$$

Если теперь $\|x_1-\hat{x}\| < \delta$ и $\|x_2-\hat{x}\| < \delta$, то в силу (3)

$$\|\varphi(x_1)-\varphi(x_2)\| \leqslant \leqslant \|\varphi(x_1)-\varphi(x_2)-A_1(x_1-x_2)\|+\|A_1(x_1-x_2)\| \leqslant \leqslant \varepsilon_1\|x_1-x_2\|+\|A_1\|\|x_1-x_2\| = (\|A_1\|+\varepsilon_1)\|x_1-x_2\|. \quad (5)$$

Полагая в этом неравенстве поочередно $x_1=\hat{x}$ и $x_2=\hat{x}$, получаем

$$\|\varphi(x_i)-\varphi(\hat{x})\| = \|\varphi(x_i)-\hat{y}\| < (\|A_1\|+\varepsilon_1)\delta \leqslant \delta_2,$$

так что для $y_i=\varphi(x_i)$ справедливо (4). Используя теперь (4), (3) и (5), получаем

$$\|f(x_1)-f(x_2)-A_2A_1(x_1-x_2)\| \leqslant \leqslant \|\psi(\varphi(x_1))-\psi(\varphi(x_2))-A_2(\varphi(x_1)-\varphi(x_2))\| + +\|A_2(\varphi(x_1)-\varphi(x_2))-A_2A_1(x_1-x_2)\| \leqslant \leqslant \varepsilon_2\|\varphi(x_1)-\varphi(x_2)\|+\|A_2\|\|\varphi(x_1)-\varphi(x_2)-A(x_1-x_2)\| \leqslant \leqslant \varepsilon_2(\|A_1\|+\varepsilon_1)\|x_1-x_2\|+\|A_2\|\varepsilon_1\|x_1-x_2\| = =(\varepsilon_2\|A_1\|+\varepsilon_1\varepsilon_2+\|A_2\|\varepsilon_1)\|x_1-x_2\| \leqslant \varepsilon\|x_1-x_2\|,$$

что и означает строгую дифференцируемость f в точке $\hat{x}$.

Полагая в этих рассуждениях $x_2=\hat{x}$, $x_1=\hat{x}+h$, мы получим доказательство теоремы для случая дифференцируемости φ по Фреше. Остальные утверждения получаются анализом уже доказанного равенства (2). ■

Следующий контрпример показывает, что теорема о суперпозиции не имеет, вообще говоря, места, если ψ дифференцируема лишь по Гато.

Пример. Пусть $X=Y=\mathbf{R}^2$, $Z=\mathbf{R}$;

$$\varphi(x)=(\varphi_1(x_1, x_2), \varphi_2(x_1, x_2))=(x_1^2, x_2);$$

$$\psi(y)=\psi(y_1, y_2)=\begin{cases} 1 & \text{при } y_1=y_2^2,\ y_2>0, \\ 0 & \text{в остальных случаях} \end{cases}$$

(ср. с контрпримером в доказательстве предложения 1 п. 2.1.1). Здесь φ дифференцируема по Фреше в точке (0, 0) и даже строго (проверьте!), ψ дифференцируема по Гато в (0, 0), однако функция

$$f(x)=(\psi\circ\varphi)(x)=\psi(x_1^2, x_2)=\begin{cases} 1 & \text{при } x_2=|x_1|>0, \\ 0 & \text{в остальных случаях} \end{cases}$$

не дифференцируема по Гато в (0, 0) (и даже не имеет в этой точке производных по направлениям $h=(1,1)$ и $h=(-1,1)$).

Следствие. *Пусть X и Y — нормированные пространства. U — окрестность точки $\hat{y}$ в Y, $f\colon U\to\mathbf{R}$ дифференцируема по Фреше в точке $\hat{y}$ и $\Lambda\in\mathscr{L}(X, Y)$. Тогда $f\circ\Lambda\colon \Lambda^{-1}(U)\to\mathbf{R}$ дифференцируема по Фреше в точке $\hat{x}=\Lambda^{-1}(\hat{y})$ и*

$$(f\circ\Lambda)'(\hat{x})=\Lambda^*(f'\circ\Lambda)(\hat{x}). \tag{6}$$

Доказательство. В соответствии с определениями

$$(f'\circ\Lambda)(\hat{x})=f'(\Lambda\hat{x})=f'(\hat{y})\in\mathscr{L}(Y, \mathbf{R})=Y^*$$

и

$$(f\circ\Lambda)'(\hat{x})=f'(\Lambda\hat{x})\circ\Lambda=f'(\hat{y})\circ\Lambda\in\mathscr{L}(X, \mathbf{R})=X^*.$$

Теперь имеем для любого $h\in X$

$$\langle\Lambda^*(f'\circ\Lambda)(\hat{x}), h\rangle=\langle(f'\circ\Lambda)(\hat{x}), \Lambda h\rangle=f'(\hat{y})[\Lambda h]=$$
$$=(f'(\hat{y})\circ\Lambda)h=\langle f'(\hat{y})\circ\Lambda, h\rangle=\langle(f\circ\Lambda)'(\hat{x}), h\rangle,$$

чем доказано (6).

2.2.3. Теорема о среднем и ее следствия. Хорошо известно, что для числовых функций одного переменного справедлива теорема Лагранжа, называемая также теоремой о среднем значении или формулой конечных приращений: *если функция $f\colon[a, b]\to\mathbf{R}$ непрерывна на отрезке $[a, b]$ и дифференцируема в интервале (a, b), то существует точка $c\in(a, b)$ такая, что*

$$f(b)-f(a)=f'(c)(b-a). \tag{1}$$

Нетрудно убедиться также в том, что формула (1) остается справедливой и для числовых функций $f(x)$, аргумент которых принадлежит произвольному линейному топологическому пространству.

В этом случае

$$[a, b]=\{x \mid x=a+t(b-a),\ 0\leqslant t\leqslant 1\},$$

аналогично определяется интервал (a, b), а дифференцируемость можно понимать в смысле Гато. Полагая

$$\Phi(t)=f(a+t(b-a)),$$

мы сводим доказательство к случаю одного вещественного переменного.

Совсем не так обстоит дело для векторнозначных функций.

Пример. Пусть отображение f: $\mathbf{R}\to\mathbf{R}^2$ определяется равенством $f(t)=(\sin t, -\cos t)$. Тогда для каждого t существует (строгая) производная Фреше $f'(t)$ (докажите!):

$$f'(t)[\alpha]=(\cos t,\ \sin t)\alpha=(\alpha\cos t,\ \alpha\sin t).$$

В то же время для любого c

$$f(2\pi)-f(0)=0\neq f'(c)[2\pi-0]=(2\pi\cos c,\ 2\pi\sin c),$$

так что формула (1) не имеет места.

Можно заметить, однако, что сама формула (1) используется в анализе много реже, чем вытекающая из нее оценка $|f(b)-f(a)|\leqslant M|b-a|$, где $M=\sup|f'(x)|$. Мы покажем сейчас, что в этом более слабом виде утверждение распространяется уже на случай произвольных нормированных пространств. По традиции оно сохраняет название «теорема о среднем», хотя, конечно, должно было бы именоваться «теоремой об оценке конечного приращения».

Теорема о среднем. *Пусть X и Y — нормированные линейные пространства и открытое множество $U\subseteq X$ содержит отрезок $[a, b]$.*

Если функция f: $U\to Y$ дифференцируема по Гато в каждой точке $x\in[a, b]$, то

$$\|f(b)-f(a)\|\leqslant\sup_{c\in[a,b]}\|f'(c)\|\,\|b-a\|. \tag{2}$$

Доказательство. Возьмем произвольно $y^*\in Y^*$ и рассмотрим функцию

$$\Phi(t)=\langle y^*,\ f(a+t(b-a))\rangle.$$

В каждой точке отрезка $[0, 1]$ эта функция имеет лево- и правостороннюю производные:

$$\Phi'_-(t)=\lim_{\alpha\downarrow 0}\frac{\Phi(t-\alpha)-\Phi(t)}{-\alpha}=$$
$$=\lim_{\alpha\downarrow 0}\left\langle y^*,\ \frac{f(a+(t-\alpha)(b-a))-f(a+t(b-a))}{-\alpha}\right\rangle=$$
$$=-\left\langle y^*,\ \lim_{\alpha\downarrow 0}\frac{f(a+t(b-a)-\alpha(b-a))-f(a+t(b-a))}{\alpha}\right\rangle=$$
$$=-\langle y^*,\ f'(a+t(b-a))[-(b-a)]\rangle=$$
$$=\langle y^*,\ f'(a+t(b-a))[b-a]\rangle,$$

и аналогично

$$\Phi'_+(t)=\lim_{\alpha\downarrow 0}\frac{\Phi(t+\alpha)-\Phi(t)}{\alpha}=\langle y^*,\ f'(a+t(b-a))[b-a]\rangle.$$

Так как эти производные совпадают, то $\Phi(t)$ дифференцируема (в обычном смысле) на $[0, 1]$, а потому и непрерывна на том же отрезке. По формуле Лагранжа существует такое $\theta\in(0, 1)$, что

$$\langle y^*,\ f(b)-f(a)\rangle=\Phi(1)-\Phi(0)=\Phi'(\theta)=$$
$$=\langle y^*,\ f'(a+\theta(b-a))[b-a]\rangle.$$

Воспользуемся теперь следствием 1 из теоремы Хана — Банаха (п. 2.1.3), согласно которому для любого элемента $y\in Y$ найдется линейный функционал $y^*\in Y^*$ такой, что $\|y^*\|=1$ и $\langle y^*,\ y\rangle=\|y\|$. Выбрав именно так функционал y^* для элемента $y=f(b)-f(a)$, получаем

$$\|f(b)-f(a)\|=\langle y^*,\ f(b)-f(a)\rangle=$$
$$=\langle y^*,\ f'(a+\theta(b-a))[b-a]\rangle\leqslant$$
$$\leqslant\|y^*\|\|f'(a+\theta(b-a))\|\|b-a\|\leqslant\sup_{c\in[a,b]}\|f'(c)\|\|b-a\|,$$

что и требовалось доказать. ■

Приведем несколько следствий из теоремы о среднем.

Следствие 1. *Пусть выполнены все условия теоремы о среднем и* $\Lambda\in\mathscr{L}(X, Y)$. *Тогда*

$$\|f(b)-f(a)-\Lambda(b-a)\|\leqslant\sup_{c\in[a,b]}\|f'(c)-\Lambda\|\|b-a\|. \tag{3}$$

Доказательство. Применим теорему о среднем к отображению $g(x)=f(x)-\Lambda x$. ■

Следствие 2. *Пусть X и Y — нормированные пространства, U — окрестность точки $\hat{x}$ в X и отображение*

f: $U \longrightarrow Y$ дифференцируемо по Гато в каждой точке $x \in U$. Если отображение $x \mapsto f'_\Gamma(x)$ непрерывно (в равномерной операторной топологии пространства $\mathscr{L}(X, Y)$) в точке $\hat{x}$, то отображение f строго дифференцируемо в $\hat{x}$ (а следовательно, и дифференцируемо по Фреше в той же точке).

Доказательство. По заданному $\varepsilon > 0$ найдем $\delta > 0$ так, чтобы выполнялось соотношение

$$\|x - \hat{x}\| < \delta \Rightarrow \|f'_\Gamma(x) - f'_\Gamma(\hat{x})\| < \varepsilon. \tag{4}$$

Если $\|x_1 - \hat{x}\| < \delta$ и $\|x_2 - \hat{x}\| < \delta$, то для любого $x = x_1 + t(x_2 - x_1) \in [x_1, x_2]$, $0 \leqslant t \leqslant 1$, и

$$\|x - \hat{x}\| = \|x_1 + t(x_2 - x_1) - \hat{x}\| = \\ = \|t(x_2 - \hat{x}) + (1 - t)(x_1 - \hat{x})\| \leqslant \\ \leqslant t\|x_2 - \hat{x}\| + (1 - t)\|x_1 - x\| < t\delta + (1 - t)\delta = \delta,$$

так что в силу (4) $\|f'_\Gamma(x) - f'(\hat{x})\| < \varepsilon$.

Применяя следствие 1 для $\Lambda = f'_\Gamma(\hat{x})$, получаем

$$\|f(x_1) - f(x_2) - f'_\Gamma(\hat{x})(x_1 - x_2)\| \leqslant \\ \leqslant \sup_{x \in [x_1, x_2]} \|f'_\Gamma(x) - f'_\Gamma(\hat{x})\| \|x_1 - x_2\| \leqslant \varepsilon \|x_1 - x_2\|,$$

что и означает строгую дифференцируемость f в $\hat{x}$. ■

Определение. Пусть X и Y — нормированные пространства. Отображение F: $U \longrightarrow Y$, определенное на некотором открытом подмножестве $U \subset X$, принадлежит классу $C^1(U)$, если в каждой точке $x \in U$ оно имеет производную и отображение $x \mapsto F'(x)$ непрерывно (в равномерной операторной топологии).

Следствие 2 показывает нам, что здесь можно не оговаривать, какая производная имеется в виду. Этим замечанием постоянно пользуются при проверке дифференцируемости конкретных функционалов: доказывается существование производной Гато и проверяется ее непрерывность, а это уже гарантирует строгую дифференцируемость (и, значит, существование производной Фреше).

Упражнения. Пусть в условиях теоремы о среднем $Y = \mathbf{R}^n$ и отображение $x \longrightarrow f'_\Gamma(x)$ непрерывно в $U \subset X$. Тогда

1) $$f(b) - f(a) = \int_0^1 f'_\Gamma(a + t(b - a))\, dt\, [b - a]. \tag{5}$$

2) Существует оператор $\Lambda \in \mathscr{L}(X, Y)$, принадлежащий замкнутой выпуклой оболочке (см. п. 2.6.2) множества $\{f'_\Gamma(x),\ x \in [a, b]\}$ и такой, что

$$f(b) - f(a) = \Lambda(b - a).$$

3) Обобщите 1) и 2) на случай произвольного банахова пространства Y, определив подходящим образом интеграл в (5).

2.2.4. Дифференцирование в произведении пространств. Частные производные. Теорема о полном дифференциале. В этом пункте X, Y, Z — нормированные пространства. Рассмотрим сначала случай отображения, значения которого лежат в произведении $Y \times Z$, $F\colon U \to Y \times Z$, $U \subset X$. Поскольку точкой $Y \times Z$ является пара (y, z), отображение F также состоит из двух компонент: $F(x) = (G(x), H(x))$, где $G\colon U \to Y$, $H\colon U \to Z$. Непосредственно из определений выводится следующее

Предложение 1. *Пусть X, Y, Z — нормированные пространства, U — окрестность точки $\hat{x}$ в X, $G\colon U \to Y$, $H\colon U \to Z$.*

Для того чтобы отображение $F = (G, H)\colon U \to Y \times Z$ было дифференцируемо в точке $\hat{x}$ в смысле одного из определений 1—5 п. 2.2.1, *необходимо и достаточно, чтобы этим же свойством обладали G и H. При этом*

$$F'(\hat{x}) = (G'(\hat{x}),\ H'(\hat{x})),$$

или

$$F'(\hat{x}, h) = (G'(\hat{x}, h),\ H'(\hat{x}, h)).$$

Перейдем теперь к случаю, когда область определения отображения $F\colon U \to Z$ лежит в произведении пространств $U \subset X \times Y$.

Определение. Пусть X, Y, Z — нормированные пространства, U — окрестность точки $(\hat{x}, \hat{y})$ в $X \times Y$, $F\colon U \to Z$.

Если отображение $x \mapsto F(x, \hat{y})$ дифференцируемо в точке $\hat{x}$ (по Гато, по Фреше или строго), то его производная называется *частной производной по x* отображения F в точке $(\hat{x}, \hat{y})$ и обозначается $F_x(\hat{x}, \hat{y})$ или $\frac{\partial F}{\partial x}(\hat{x}, \hat{y})$.

Аналогично определяется частная производная по y

$$F_y(\hat{x}, \hat{y}) = \frac{\partial F}{\partial y}(\hat{x}, \hat{y}).$$

Теорема о полном дифференциале. *Пусть X, Y и Z — нормированные пространства, U — окрест-*

ность в $X\times Y$, $F\colon U\to Z$ — *отображение, имеющее в каждой точке* $(x, y)\in W$ *частные производные* $F_x(x, y)$ *и* $F_z(x, y)$ *в смысле Гато.*

Если отображения $(x, y)\mapsto F_x(x, y)$ *и* $(x, y)\mapsto F_y(x, y)$ *непрерывны в точке* $(\hat{x}, \hat{y})\in U$ *в равномерной операторной топологии, то* F *строго дифференцируема в той же точке и при этом*

$$F'(\hat{x}, \hat{y})[\xi, \eta]=F_x(\hat{x}, \hat{y})\xi+F_y(\hat{x}, \hat{y})\eta.$$

Доказательство. Задавшись произвольным $\varepsilon>0$, выберем $\delta>0$ столь малым, чтобы «прямоугольная» окрестность

$$V=\mathring{B}(\hat{x}, \delta)\times\mathring{B}(\hat{y}, \delta)=\{(x, y)\,|\,\|x-\hat{x}\|<\delta, \|y-\hat{y}\|<\delta\}$$

точки $(\hat{x}, \hat{y})$ содержалась в U, и в ней выполнялись неравенства

$$\|F_x(x, y)-F_x(\hat{x}, \hat{y})\|<\varepsilon,\quad \|F_y(x, y)-F_y(\hat{x}, \hat{y})\|<\varepsilon. \tag{1}$$

Теперь имеем

$$\begin{aligned}\Delta&=\\&=F(x_1, y_1)-F(x_2, y_2)-F_x(\hat{x}, \hat{y})(x_1-x_2)-F_y(\hat{x}, \hat{y})(y_1-y_2)=\\&=[F(x_1, y_1)-F(x_2, y_1)-F_x(\hat{x}, \hat{y})(x_1-x_2)]+\\&\qquad+[F(x_2, y_1)-F(x_2, y_2)-F_y(\hat{x}, \hat{y})(y_1-y_2)].\end{aligned}$$

Легко видеть, что если точки (x_1, y_1), (x_2, y_2) лежат в V, то и $(x_2, y_1)\in V$ и, более того, оба отрезка $[(x_1, y_1), (x_2, y_1)]$ и $[(x_2, y_1), (x_2, y_2)]$ содержатся в $V\subset U$. Поэтому функции $x\mapsto F(x, y_1)$ и $y\mapsto F(x_2, y)$ дифференцируемы по Гато: первая имеет производную F_x на $[x_1, x_2]$, вторая F_y на $[y_1, y_2]$. Применяя теорему о среднем к этим функциям (в форме неравенства (3) п. 2.2.3 с соответствующими Λ), получаем в силу (1)

$$\begin{aligned}&(x_1, y_1)\in V,\ (x_2, y_2)\in V\Rightarrow\\&\Rightarrow\|\Delta\|\leqslant\sup_{\xi\in[x_1, x_2]}\{\|F_x(\xi, y_1)-F_x(\hat{x}, \hat{y})\|\,\|x_1-x_2\|\}+\\&\qquad+\sup_{\eta\in[y_1, y_2]}\{\|F_y(x_2, \eta)-F_y(\hat{x}, \hat{y})\|\,\|y_1-y_2\|\}\leqslant\\&\qquad\qquad\leqslant\varepsilon\|x_1-x_2\|+\varepsilon\|y_1-y_2\|.\ \blacksquare\end{aligned}$$

Следствие. *Для того чтобы* $F\in C^1(U)$, *необходимо и достаточно, чтобы в* U *частные производные* F_x *и* F_y *были непрерывны.*

Как предложение 1, так и теорема о полном дифференциале без труда обобщаются на случай произведения любого конечного числа пространств.

Остановимся еще на конечномерном случае. Пусть $F\colon U \longrightarrow \mathbf{R}^m$ определена на открытом множестве $U \subset \mathbf{R}^n$. Поскольку естественным образом

$$\mathbf{R}^m = \underbrace{\mathbf{R} \times \ldots \times \mathbf{R}}_{m \text{ раз}}, \quad \mathbf{R}^n = \underbrace{\mathbf{R} \times \ldots \times \mathbf{R}}_{n \text{ раз}},$$

можно воспользоваться как предложением 1, так и теоремой о полном дифференциале. Если

$$F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \\ \cdot \\ \cdot \\ F_m(x) \end{pmatrix} = \begin{pmatrix} F_1(x_1, \ldots, x_n) \\ F_2(x_1, \ldots, x_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ F_m(x_1, \ldots, x_n) \end{pmatrix} \quad \text{и} \quad h = \begin{pmatrix} h_1 \\ h_2 \\ \\ \\ h_n \end{pmatrix},$$

то

$$F'(\hat{x})\,[h] = \begin{pmatrix} F_1'(\hat{x})[h] \\ F_2(\hat{x})[h] \\ \\ F_m'(\hat{x})[h] \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} \frac{\partial F_1}{\partial x_j}(\hat{x})\, h_j \\ \sum_{j=1}^{n} \frac{\partial F_2}{\partial x_j}(\hat{x})\, h_j \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \sum_{j=1}^{n} \frac{\partial F_m}{\partial x_j}(\hat{x})\, h_j \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\hat{x}) & \ldots & \frac{\partial F_1}{\partial x_n}(\hat{x}) \\ \cdot & & \\ \cdot & & \\ \frac{\partial \Gamma_m}{\partial x_1}(\hat{x}) & \ldots & \frac{\partial F_m}{\partial x_n}(\hat{x}) \end{pmatrix} \begin{pmatrix} h_1 \\ \\ \cdot \\ h_n \end{pmatrix} = \frac{\partial F}{\partial x}(\hat{x})\; h. \qquad (2)$$

Матрица порядка $m \times n$

$$\frac{\partial F}{\partial x}(\hat{x}) = \left(\frac{\partial F_i}{\partial x_j}(\hat{x})\right)$$

называется *матрицей Якоби* отображения F в точке $\hat{x}$. Легко видеть, что это есть не что иное, как матрица линейного оператора $F'(\hat{x})\colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$, отвечающая стандартным базисам в $\mathbf{R}^n$ и $\mathbf{R}^m$.

В классическом анализе обычно обозначают

$$h = dx = \begin{pmatrix} dx_1 \\ \cdot \\ \cdot \\ \cdot \\ dx_n \end{pmatrix}$$

и формула (2) имеет вид

$$dF(\hat{x}) = \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(\hat{x})\, dx_j.$$

Доказанное в этом пункте утверждение является обобщением хорошо известной теоремы о том, что существование непрерывных частных производных является достаточным условием дифференцируемости функции нескольких переменных.

2.2.5. Производные высших порядков. Формула Тейлора. В этом пункте X и Y — нормированные пространства, U — открытое подмножество в X. Дифференцируемость всюду понимается в смысле Фреше.

Если отображение $f\colon U \to Y$ дифференцируемо в каждой точке $x \in U$, то определено отображение $f'(x)\colon U \to \mathscr{L}(X, Y)$. Поскольку $\mathscr{L}(X, Y)$ также является нормированным пространством, можно ставить вопрос о существовании второй производной

$$f''(x) = (f')'(x) \in \mathscr{L}(X, \mathscr{L}(X, Y)).$$

По индукции определяются производные высших порядков: если в U уже определена $f^{(n-1)}(x)$, то

$$f^{(n)}(x) = (f^{(n-1)})'(x) \in \underbrace{\mathscr{L}(X, \ldots, \mathscr{L}(X, Y)}_{n \text{ раз}} \ldots).$$

Определение 1. Пусть $f\colon U \to Y$. Будем говорить, что $f^{(n)}$ существует в точке $\hat{x} \in U$, если в некоторой окрестности этой точки существуют $f'(x)$, $f''(x)$, ..., $f^{(n-1)}(x)$ и существует $f^{(n)}(\hat{x})$.

Если $f^{(n)}(x)$ существует в каждой точке $x \in U$ и отображение $x \mapsto f^{(n)}(x)$ непрерывно в равномерной (порожденной нормой) топологии пространства $\mathscr{L}(X, \ldots, \mathscr{L}(X, Y) \ldots)$, то f называется отображением класса $C^n(U)$.

В дальнейшем нам понадобятся некоторые свойства полилинейных непрерывных отображений (см. пример 2 в п. 2.2.1).

П р е д л о ж е н и е 1. *Нормированные пространства* $\mathscr{L}^n(X, \mathscr{L}^m(X, Y))$ *и* $\mathscr{L}^{n+m}(X, Y)$ *изометрично изоморфны.*

Д о к а з а т е л ь с т в о. Если $\pi \in \mathscr{L}^n(X, \mathscr{L}^m(X, Y))$, то $\pi(x_1, \ldots, x_n) \in \mathscr{L}^m(X, Y)$ и равенство

$$\Pi(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m}) = \\ = \pi(x_1, \ldots, x_n)[x_{n+1}, \ldots, x_{n+m}] \qquad (1)$$

определяет полилинейное отображение Π пространства $X^{n+m} = \underbrace{X \times \ldots \times X}_{n+m \text{ раз}}$ в Y. Обратно, всякое такое отображение Π определяет при помощи равенства (1) полилинейное отображение π пространства $X^n = \underbrace{X \times \ldots \times X}_{n \text{ раз}}$ в пространство полилинейных отображений X^m в Y Остается заметить, что

$$\|\Pi\| = \sup_{\substack{\|x_1\| \leqslant 1 \\ \cdots \\ \|x_{n+m}\| \leqslant 1}} \|\Pi(x_1, \ldots, x_{n+m})\| = \\ = \sup_{\substack{\|x_1\| \leqslant 1 \\ \cdots \\ \|x_n\| \leqslant 1}} \sup_{\substack{\|x_{n+1}\| \leqslant 1 \\ \cdots \\ \|x_{n+m}\| \leqslant 1}} \pi(x_1, \ldots, x_n)[x_{n+1}, \ldots, x_{n+m}] = \\ = \sup_{\substack{\|x_1\| \leqslant 1 \\ \cdots \\ \|x_n\| \leqslant 1}} \|\pi(x_1, \ldots, x_n)\| = \|\pi\|. \blacksquare$$

С л е д с т в и е 1. $\underbrace{\mathscr{L}(X, \mathscr{L}(X, \ldots, \mathscr{L}(X, Y) \ldots))}_{n \text{ раз}}$ *изометрично изоморфно* $\mathscr{L}^n(X, Y)$.

Таким образом, можно считать, что $f^{(n)}(x) \in \mathscr{L}^n(X, Y)$. Значение этого полилинейного отображения на векторах $(\xi_1, \ldots, \xi_n)$ будем обозначать $f^{(n)}(x)[\xi_1, \ldots, \xi_n]$. В соответствии с индуктивным определением

$$f^{(n)}(x_0)[\xi_1, \ldots, \xi_n] = \\ = \lim_{\alpha \downarrow 0} \frac{f^{(n-1)}(x_0 + \alpha\xi_1)[\xi_2, \ldots, \xi_n] - f^{(n-1)}(x_0)[\xi_2, \ldots, \xi_n]}{\alpha}. \qquad (2)$$

Для произвольного $\Pi \in \mathscr{L}^n(X, Y)$ и произвольного набора различных индексов $(i_1, \ldots, i_k)$, каждый из которых принимает одно из значений $1, 2, \ldots, n$, обозначим

$$\Pi_{i_1 \ldots i_k}(x; h_1, \ldots, h_k) = \Pi(x_1, \ldots, x_n)\Big|_{\substack{x_{i_l} = h_l,\ l = 1, 2, \ldots, k, \\ x_i = x,\ i \neq i_1, \ldots, i_k.}}$$

Предложение 2. *Если* $\Pi \in \mathscr{L}^n(X, Y)$ *и* $Q(x) = \Pi(x, \ldots, x)$, *то*

$$Q'(x)[h_1] = \sum_{i=1}^{n} \Pi_i(x; h_1),$$
$$Q''(x)[h_1, h_2] = \sum_{i=1}^{n} \sum_{\substack{j=1\\ j\neq i}}^{n} \Pi_{ij}(x; h_1, h_2), \tag{3}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$Q^{(n)}(x)[h_1, \ldots, h_n] = \sum_{(i_1, \ldots, i_n)} \Pi(h_{i_1}, \ldots, h_{i_n}),$$
$$Q^{(k)}(x) \equiv 0 \quad \text{при} \quad k > n$$

(в формуле для $Q^{(n)}$ суммирование производится по всем перестановкам индексов).

Доказательство этих формул получается непосредственной выкладкой (ср. также пример в п. 2.2.1). Заметим, что из (3) видно, что каждая производная $Q^{(l)}$ является симметрической функцией от $(h_1, \ldots, h_l)$. Ниже мы увидим, что это не случайно.

Предложение 3. *Если отображение* $\Pi \in \mathscr{L}^n(X, Y)$ *симметрично и* $Q(x) = \Pi(x, \ldots, x) \equiv 0$, *то* $\Pi \equiv 0$.

Доказательство. Положим

$$\varphi(t_1, \ldots, t_n) = Q(t_1x_1 + \ldots + t_nx_n) =$$
$$= \sum_{i_1, \ldots, i_n} t_{i_1} \ldots t_{i_n} \Pi(x_{i_1} \ldots x_{i_n}) =$$
$$= \sum_{\substack{0 \leqslant \alpha_i \leqslant n\\ \Sigma\alpha_i = n}} t_1^{\alpha_1} \ldots t_n^{\alpha_n} \Sigma'_\alpha \Pi(x_{i_1}, \ldots, x_{i_n}) =$$
$$= \sum_{\substack{0 \leqslant \alpha_i \leqslant n\\ \Sigma\alpha_i = n}} t_1^{\alpha_1} \ldots t_n^{\alpha_n} N_{\alpha_1 \ldots \alpha_n} \Pi(\underbrace{x_1, \ldots, x_1}_{\alpha_1 \text{ раз}}, \ldots, \underbrace{x_n, \ldots, x_n}_{\alpha_n \text{ раз}}),$$

где в Σ'_α собраны члены, у которых среди индексов $i_1, \ldots, i_n$ равны единице α_1 штук, ..., равны $n - \alpha_n$ штук, а $N_{\alpha_1 \ldots \alpha_n}$ — число членов в этой сумме (в последнем переходе мы воспользовались симметрией Π).

По условию $\varphi(t_1, \ldots, t_n) \equiv 0$, поэтому равны нулю все коэффициенты этого многочлена. В частности, равен нулю и коэффициент $n!\Pi(x_1, \ldots, x_n)$ при $t_1t_2 \ldots t_n$. ■

Теорема о смешанных производных. *Если для функции* $f\colon U \to Y$ *существует вторая производная* $f''(\hat{x})$, *то для всех* $\xi, \eta \in X$

$$f''(\hat{x})[\xi, \eta] = f''(\hat{x})[\eta, \xi].$$

Доказательство. По определению второй производной

$$f'(x)-f'(\hat{x})=(f')'(\hat{x})[x-\hat{x}]+\alpha(x)\|x-\hat{x}\|,$$

где $\alpha(x)\in\mathscr{L}(X, Y)$ и

$$\lim_{x\to\hat{x}}\alpha(\hat{x})=\alpha(\hat{x})=0. \qquad (4)$$

При достаточно малых η и x, близких к $\hat{x}$, функция

$$\varphi(x)=f(x+\eta)-f(x)$$

определена и

$$\begin{aligned}\varphi'(x)&=f'(x+\eta)-f'(x)=f'(x+\eta)-f'(\hat{x})-(f'(x)-f'(\hat{x}))=\\&=(f')'(\hat{x})[x+\eta-\hat{x}]+\alpha(x+\eta)\|x+\eta-\hat{x}\|-\\&\quad-(f')'(\hat{x})[x-\hat{x}]-\alpha(x)\|x-\hat{x}\|=\\&=(f')'(\hat{x})[\eta]+\alpha(x+\eta)\|x-\hat{x}+\eta\|-\alpha(x)\|x-\hat{x}\|.\end{aligned}$$

В частности,

$$\varphi'(\hat{x})=(f')'(\hat{x})[\eta]+\alpha(\hat{x}+\eta)\|\eta\|, \qquad (5)$$

и потому

$$\begin{aligned}&\varphi'(x)-\varphi'(\hat{x})=\\&\quad=\alpha(x+\eta)\|x-\hat{x}+\eta\|-\alpha(x)\|x-\hat{x}\|-\alpha(\hat{x}+\eta)\|\eta\|. \qquad (6)\end{aligned}$$

В силу (4) для любого $\varepsilon>0$ найдется такое $\delta>0$, что

$$\|x-\hat{x}\|<\delta\Rightarrow\|\alpha(x)\|<\varepsilon, \qquad (7)$$

откуда ввиду (6)

$$\begin{aligned}&\|x-\hat{x}\|<\delta/2,\\&\quad\|\eta\|<\delta/2\Rightarrow\|\varphi'(x)-\varphi'(\hat{x})\|<2\varepsilon(\|x-\hat{x}\|+\|\eta\|). \qquad (8)\end{aligned}$$

При достаточно малых ξ и η вторая разность

$$\Delta(\eta, \xi)=f(\hat{x}+\xi+\eta)-f(\hat{x}+\xi)-f(\hat{x}+\eta)+f(\hat{x}) \qquad (9)$$

имеет смысл и, используя (5), (7) и (8), мы получаем при $\|\xi\|<\delta/2$, $\|\eta\|<\delta/2$

$$\begin{aligned}&\|\Delta(\eta, \xi)-f''(\hat{x})[\eta, \xi]\|=\\&\quad=\|\varphi(\hat{x}+\xi)-\varphi(\hat{x})-((f')'(\hat{x})[\eta])[\xi]\|=\\&\quad=\|\varphi(\hat{x}+\xi)-\varphi(\hat{x})-\varphi'(\hat{x})[\xi]+\alpha(\hat{x}+\eta)[\xi]\|\eta\|\|\leqslant\\&\quad\leqslant\sup_{x\in[\hat{x},\,\hat{x}+\xi]}\|\varphi'(x)-\varphi'(\hat{x})\|\|\xi\|+\|\alpha(\hat{x}+\eta)\|\|\xi\|\|\eta\|\leqslant\\&\quad\leqslant 2\varepsilon(\|\xi\|+\|\eta\|)\|\xi\|+\varepsilon\|\xi\|\|\eta\|\leqslant 3\varepsilon(\|\xi\|+\|\eta\|)\|\xi\|.\end{aligned}$$

Заменяя ξ и η на $t\xi$, $t\eta$, где теперь ξ и η могут быть любыми фиксированными, а $t \in \mathbb{R}$ должно быть мало, мы получаем неравенство

$$\|\Delta(t\eta,\ t\xi) - t^2 f''(\hat{x})[\eta,\ \xi]\| \leqslant 3\varepsilon t^2 (\|\xi\| + \|\eta\|)\|\xi\|. \quad (10)$$

Меняя местами в проведенных выше рассуждениях ξ и η, мы получаем наряду с (10) неравенство

$$\|\Delta(t\xi,\ t\eta) - t^2 f''(\hat{x})[\xi,\ \eta]\| \leqslant 3\varepsilon t^2 (\|\xi\| + \|\eta\|)\|\eta\|. \quad (11)$$

Но $\Delta(t\eta,\ t\xi) = \Delta(t\xi,\ t\eta)$, поскольку вторая разность (9) симметрична относительно ξ и η и из (10) и (11) вытекает по сокращении на t^2 неравенство

$$\|f''(\hat{x})[\eta,\ \xi] - f''(\hat{x})[\xi,\ \eta]\| \leqslant 3\varepsilon (\|\xi\| + \|\eta\|)^2.$$

Ввиду произвольности ε это приводит к (3). ■

Название, которое мы дали этой теореме, связано с тем, что в классическом анализе ей соответствует теорема о независимости смешанной производной от порядка дифференцирования: $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ (почему?).

Следствие 2. *Производная $f^{(n)}(\hat{x})$ (если существует) является симметрической полилинейной функцией.*

Доказательство. При $n = 2$ это утверждение только что доказанной теоремы. Далее рассуждаем по индукции, используя равенства

$$((f^{(n-1)})'(\hat{x})[\xi])[\eta_1,\ \ldots,\ \eta_{n-1}] = f^{(n)}(\hat{x})[\xi_1,\ \eta_1,\ \ldots,\ \eta_{n-1}] =$$
$$= ((f^{(n-2)})''(\hat{x})[\xi,\ \eta_1])[\eta_2,\ \ldots,\ \eta_{n-1}].$$

Воспользовавшись индукционной гипотезой, получаем из левого равенства инвариантность $f^{(n)}(\hat{x})[\xi,\ \eta_1,\ \ldots,\ \eta_{n-1}]$ относительно перестановки аргументов η_i, а из теоремы 1 и правого равенства следует инвариантность относительно транспозиции ξ и η_1. Для доказательства симметричности $f^{(n)}(\hat{x})$ этого достаточно. ■

Переходя от симметрической полилинейной функции к соответствующей ей форме (см. пример в 2.2.1), мы получаем из производной дифференциал.

Определение 2.

$$d^n f(x_0;\ h) = f^n(x_0)[h,\ \ldots,\ h].$$

Пример. Если $Q(x) = \Pi(x, \ldots, x)$, где $\Pi \in \mathscr{L}^n(X, Y)$, то, согласно предложению 2,

$$d^k Q(0;\ h) = 0 \quad \text{при} \quad k \neq n$$

и
$$d^n Q(0;\ h) = k!\,Q(h).$$

Лемма. *Если $f^{(n)}(\hat{x})$ существует и $f(\hat{x}) = 0$, $f'(\hat{x}) = 0, \ldots, f^{(n)}(\hat{x}) = 0$, то*
$$\lim_{h\to 0} \frac{\| f(\hat{x}+h)\|}{\|h\|^n} = 0.$$

Доказательство. При $n=1$ утверждение верно по определению производной. Далее рассуждаем по индукции. Пусть $n>1$. Положим $g(x)=f'(x)$. Тогда $g(\hat{x}) = 0, \ldots, g^{(n-1)}(\hat{x}) = 0$, и по индукционной гипотезе для всякого $\varepsilon>0$ найдется такое $\delta>0$, что
$$\|g(\hat{x}+h_1)\| < \varepsilon \|h_1\|^{n-1} \quad \text{при} \quad \|h_1\| < \delta.$$
Теперь по теореме о среднем при $\|h\|<\delta$ имеем
$$\|f(\hat{x}+h)\| = \|f(\hat{x}+h) - f(\hat{x}) - f'(\hat{x})h\| \leqslant$$
$$\leqslant \sup_{h_1\in[0,\,h]} \|f'(\hat{x}+h_1) - f'(\hat{x})\|\,\|h\| =$$
$$= \sup_{h_1\in[0,\,h]} \|g(\hat{x}+h_1)\|\,\|h\| \leqslant \sup_{h_1\in[0,\,h]} \varepsilon \|h_1\|^{n-1}\|h\| = \varepsilon\|h\|^n. \blacksquare$$

Теорема о формуле Тейлора. *Если $f^n(\hat{x})$ существует, то*
$$f(\hat{x}+h) = f(\hat{x}) + df(\hat{x};\ h) + \ldots + \frac{1}{n!} d^n f(\hat{x};\ h) + \alpha_n(h)\|h\|^n,$$
где
$$\lim_{h\to 0} \alpha(h) = \alpha_n(0) = 0.$$

Доказательство. Рассмотрим функцию
$$g(\xi) = f(\hat{x}+\xi) - f(\hat{x}) - df(\hat{x};\ \xi) - \ldots - \frac{1}{n!} d^n f(\hat{x};\ \xi).$$
Воспользовавшись приведенным выше примером и равенством $d^k f(\hat{x};\ \xi) = f^{(k)}(\hat{x})[\xi, \ldots, \xi]$, $f^{(k)}(\hat{x}) \in \mathscr{L}^k(X, Y)$, получаем $d^k g(0;\ h) = d^k f(\hat{x},\ h) - \frac{1}{k!}(k!\,d^k f(\hat{x};\ h)) = 0$.

Применяя следствие 1, согласно которому $g^{(k)}(0)$ — симметрическая функция из $\mathscr{L}^k(X, Y)$, и предложение 3, согласно которому симметрическая функция обращается в нуль, коль скоро обращается в нуль соответствующая ей форма, мы заключаем, что $g(0)=0, \ldots, g^{(n)}(0)=0$.

По лемме $\|g(\xi)\| = o(\|\xi\|^n)$, и утверждение теоремы доказано, причем $\alpha_n(h) = \frac{g(h)}{\|h\|^n}$ при $h \neq 0$. ■

Упражнение. Если $f^{(n)}(x)$ существует в звездной окрестности точки $\hat{x}$, то

1) $$\left\| f(\hat{x}+h) - f(\hat{x}) - df(\hat{x}, h) - \ldots - \frac{1}{(n-1)!} d^{n-1}f(\hat{x}, h) \right\| \leqslant \leqslant \sup_{c \in [\hat{x}, \hat{x}+h]} \|f^{(n)}(c)\| \frac{\|h\|^n}{n!};$$

2) $$\left\| f(\hat{x}+h) - f(\hat{x}) - df(\hat{x}, h) - \ldots - \frac{1}{n!} d^n f(\hat{x}, h) \right\| \leqslant \leqslant \sup_{c \in [\hat{x}, \hat{x}+h]} \|f^{(n)}(c) - f^{(n)}(\hat{x})\| \frac{\|h\|^n}{n!}.$$

Указание. Один из возможных путей решения: при $n > 1$ производная $f'(x)$ непрерывна в окрестности точки $\hat{x}$, и можно воспользоваться равенством

$$f(\hat{x}+h) - f(\hat{x}) = \int_0^1 f'(x_0 + th)\,dt\,[h]$$

(см. упражнение в п. 2.2.3).

Неравенства этой задачи являются обобщениями классической формулы Тейлора с остаточным членом в форме Лагранжа, в то время как теорема дает нам ту же формулу с остаточным членом в форме Пеано.

Доказательство существования старших производных в смысле Фреше у конкретных отображений зачастую сопряжено с громоздкими выкладками. В то же время вычисление производных, а точнее, дифференциалов, т. е. соответствующих однородных форм, обычно можно произвести сравнительно легко. Чаще всего это делается с помощью понятия вариации по Лагранжу.

Определение 3. Функция $f = U \to Y$ в некоторой точке $\hat{x} \in U$ имеет *n-ю вариацию по Лагранжу*, если для любого $h \in X$ функция $\varphi(\alpha) = f(\hat{x} + \alpha h)$ дифференцируема в точке $\alpha = 0$ до порядка n включительно. При этом n-й вариацией f в точке называется отображение $h \mapsto \delta^n f(\hat{x}; h)$, определяемое равенством

$$\delta^n f(\hat{x}; h) = \varphi^{(n)}(0) = \frac{d^n}{d\alpha^n} f(\hat{x} + \alpha h)\Big|_{\alpha=0}. \qquad (12)$$

Следствие 3. *Если существует $f^{(n)}(\hat{x})$, то f имеет n-ю вариацию по Лагранжу в точке $\hat{x}$, причем*

$$\delta^n f(\hat{x};\ h) = d^n f(\hat{x};\ h) = f^{(n)}(\hat{x})[h,\ \ldots,\ h]. \qquad (13)$$

Доказательство легко получается по индукции из формулы Тейлора, так как, согласно последней,

$$f(\hat{x}+\alpha h) = f(\hat{x}) + \frac{\alpha}{1!} df(\hat{x};\ h) + \ldots + \frac{\alpha^n}{n!} d^n f(\hat{x};\ h) + o(\alpha^n).$$

Таким образом, если мы заранее уверены в существовании производной $f^{(n)}(\hat{x})$, то вычислить $d^n f(\hat{x}; h)$ можно, пользуясь формулами (12) и (13) на основе дифференциального исчисления для функций скалярного аргумента, что, конечно, много проще. Согласно предложению 3 дифференциалом $d^n f(\hat{x};\ h)$ производная $f^{(n)}(\hat{x})$ (т. е. полилинейная функция) однозначно определяется.

Обратное, конечно, неверно: функция может иметь n-ю вариацию, не имея производной Фреше $f^{(n)}$.

Читателю предоставляется самостоятельно выяснить, какими дополнительными свойствами должна обладать n-я вариация, чтобы был верен аналог достаточного условия дифференцируемости (следствие 2 п. 2.2.3).

§ 2.3. Теорема о неявной функции

Доказываемое в этом параграфе обобщение классической теоремы о неявной функции может быть с успехом применено в различных разделах анализа. В дальнейшем мы будем иметь возможность в этом убедиться.

2.3.1. Формулировка теоремы о существовании неявной функции. *Пусть X — топологическое пространство, Y и Z — банаховы пространства, W — окрестность точки $(x_0,\ y_0)$ в $X \times Y$, Ψ — отображение из W в Z, $\Psi(x_0, y_0) = z_0$. Если:*

1) *отображение $x \mapsto \Psi(x,\ y_0)$ непрерывно в точке x_0;*

2) *существует линейный непрерывный оператор $\Lambda\colon Y \to Z$ такой, что для любого $\varepsilon > 0$ найдутся число $\delta > 0$ и окрестность Ξ точки x_0, обладающие тем свойством, что из условия $x \in \Xi$ и неравенств $\|y' - y_0\| < \delta$, $\|y'' - y_0\| < \delta$ следует неравенство*

$$\|\Psi(x,\ y') - \Psi(x,\ y'') - \Lambda(y' - y'')\| < \varepsilon \|y' - y''\|;$$

3)
$$\Lambda Y = Z,$$

то найдутся число $K>0$, окрестность $\mathcal{U}$ точки (x_0, z_0) в $X\times Z$ и отображение $\varphi.\ \mathcal{U}\to Y$ такие, что:

а) $\Psi(x, \varphi(x, z))=z$;

б) $\|\varphi(x, z)-y_0\|\leqslant K\|\Psi(x, y_0)-z\|$.

2.3.2. Модифицированный принцип сжимающих отображений.

Лемма. *Пусть T — топологическое пространство, Y — банахово пространство, V — окрестность точки (t_0, y_0) в $T\times Y$, Φ — отображение из V в Y. Тогда, если существует окрестность U точки t_0 в T, число $\beta>0$ и число θ, $0<\theta<1$, такие, что из $t\in U$, $\|y-y_0\|<\beta$ следует:*

а) $$(t, \Phi(t, y))\in V;$$

б) $$\|\Phi(t, \Phi(t, y))-\Phi(t, y)\|\leqslant\theta\|\Phi(t, y)-y\|;$$

в) $$\|\Phi(t, y_0)-y_0\|<\beta(1-\theta),$$

то последовательность $\{y_n(t)\mid n\geqslant 0\}$, рекуррентно определяемая равенствами

$$y_0(t)=y_0,\quad y_n(t)=\Phi(t, y_{n-1}(t)),$$

при любом $t\in U$ содержится в шаре $\mathring{B}(y_0, \beta)=\{y\mid\|y-y_0\|<\beta\}$ и равномерно по $t\in U$ сходится к отображению $t\mapsto\varphi(t)$, причем

$$\|\varphi(t)-y_0\|\leqslant\frac{\|\Phi(t, y_0)-y_0\|}{1-\theta}.$$

Доказательство. Доказываем лемму по индукции. Пусть t — некоторый элемент из U. Обозначим через Γ_n утверждение: «элемент $y_k(t)$ определен для $0\leqslant k\leqslant n$ и принадлежит $\mathring{B}(y_0, \beta)$». Вследствие того, что $y_0(t)=y_0\in\mathring{B}(y_0, \beta)$, утверждение Γ_0 верно. Пусть Γ_n верно. Тогда для $1\leqslant k\leqslant n$ имеем

$$\begin{aligned}\|y_{k+1}(t)-y_k(t)\|&\overset{\text{def}}{=}\|\Phi(t, y_k(t))-\Phi(t, y_{k-1}(t))\|\overset{\text{def}}{=}\\&=\|\Phi(t, \Phi(t, y_{k-1}(t))-\Phi(t, y_{k-1}(t))\|\leqslant\\&\leqslant\theta\|\Phi(t, y_{k-1}(t))-y_{k-1}(t)\|\overset{\text{def}}{=}\theta\|y_k(t)-y_{k-1}(t)\|.\end{aligned}$$

Отсюда, продолжая проведенное рассуждение, получим

$$\begin{aligned}\|y_{k+1}(t)-y_k(t)\|&\leqslant\theta\|y_k(t)-y_{k-1}(t)\|\leqslant\\&\leqslant\theta^2\|y_{k-1}(t)-y_{k-2}(t)\|\leqslant\ldots\leqslant\theta^k\|y_1(t)-y_0\|=\\&=\theta^k\|\Phi(t, y_0)-y_0\|.\end{aligned}$$

Пусть теперь $k \geqslant 1$, $k+l \leqslant n+1$; тогда вследствие неравенства треугольника

$$\|y_{k+l}(t)-y_k(t)\| =$$
$$= \|y_{k+l}(t)-y_{k+l-1}(t)+y_{k+l-1}(t)-\ldots+y_{k+1}(t)-y_k(t)\| \leqslant$$
$$\leqslant (\theta^{k+l-1}+\ldots+\theta^k)\|\Phi(t, y_0)-y_0\| <$$
$$< \left(\sum_{s=k}^{\infty}\theta^s\right)\|\Phi(t, y_0)-y_0\| = \frac{\theta^k}{1-\theta}\|\Phi(t, y_0)-y_0\| < \theta^k\beta \quad (1)$$

В частности, полагая в (1) $k+l=n+1$, $k=1$, и используя условие в), получаем

$$\|y_{n+1}(t)-y_0\| = \|y_{n+1}(t)-y_1(t)+y_1(t)-y_0\| \leqslant$$
$$\leqslant \|y_{n+1}(t)-y_1(t)\|+\|y_1(t)-y_0\| < \theta\beta+\beta(1-\theta)=\beta.$$

Значит, Γ_{n+1} верно, откуда по индукции утверждение верно для всех n. Но тогда из (1) получаем

$$\|y_{k+l}(t)-y_k(t)\| \leqslant \frac{\theta^k}{1-\theta}\|\Phi(t, y_0)-y_0\|, \quad \forall l, \quad k \geqslant 1. \quad (2)$$

Из (2) следует, что последовательность $\{y_k(t) \mid k \geqslant 0\}$ фундаментальна и вследствие полноты Y сходится к некоторому элементу, который мы обозначим $\varphi(t)$. Переходя в (2) к пределу при $l \to \infty$, получаем, что

$$\|\varphi(t)-y_k(t)\| \leqslant \frac{\theta^k}{1-\theta}\|\Phi(t, y_0)-y_0\| < \theta^k\beta. \quad (3)$$

Из (3) следует, что отображение $t \mapsto y_k(t)$, $t \in U$, равномерно сходится к отображению $t \mapsto \varphi(t)$. Наконец, положив в (3) $k=1$, получим

$$\|\varphi(t)-y_0\| = \|\varphi(t)-y_1(t)+y_1(t)-y_0\| \leqslant$$
$$\leqslant \frac{\theta}{1-\theta}\|\Phi(t, y_0)-y_0\|+\|\Phi(t, y_0)-y_0\| = \frac{\|\Phi(t, y_0)-y_0\|}{1-\theta}.$$

2.3.3. Доказательство теоремы. Доказательство сформулированной в п. 2 3 1 теоремы разбиваем на несколько пунктов.

А) По условию теоремы Λ есть эпиморфизм из Y в Z. Значит, по лемме о правом обратном отображении (п. 2.1.5) существует отображение $\mathrm{M}: Z \to Y$ такое, что

$$(\Lambda \circ \mathrm{M})(z) = z, \quad (1)$$
$$\|\mathrm{M}(z)\| \leqslant C\|z\| \quad (2)$$

для некоторого $C > \theta$. Положим $T = X \times Z$ и проверим применимость леммы п. 2.3.2 для отображения

$$\Phi(t, y) = \Phi(x, z, y) = y + M(z - \Psi(x, y)).$$

Б) Задав θ, $0 < \theta < 1$ и $\varepsilon = \theta/C$, где C — константа из (2), найдем окрестность Ξ_1 точки x_0 и число $\beta_0 > 0$ так, чтобы из неравенств $\|y' - y_0\| < \beta_0$, $\|y'' - y_0\| < \beta_0$ следовало (в соответствии с условием 2) теоремы) неравенство (при $x \in \Xi_1$)

$$\|\Psi(x, y') - \Psi(x, y'') - \Lambda(y' - y'')\| < \varepsilon \|y' - y''\|. \quad (3)$$

Обозначим $V = \Xi_1 \times Z \times \mathring{B}(y_0, \beta_0)$.

В) Пусть $\beta = \beta_0/(C\|\Lambda\| + 2)$. Используя непрерывность функции $(x, z) \mapsto \Psi(x, y_0) - z$ в точке (x_0, z_0) (см. условие 1) теоремы), выберем окрестность $\Xi_2 \subset \Xi_1$ точки x_0 и $\gamma > 0$ так, чтобы из $(x, z) \in \mathcal{U} = \Xi_2 \times \mathring{B}(z_0, \gamma)$ следовало неравенство

$$\|\Psi(x, y_0) - z\| < \beta(1 - \theta)/C. \quad (4)$$

Г) Если $(x, z) \in \mathcal{U}$ и $y \in \mathring{B}(y_0, \beta)$, то

$$\|\Phi(x, z, y) - y_0\| \overset{\text{def}}{=} \|y - y_0 + M(z - \Psi(x, y))\| \overset{(2)}{\leqslant}$$
$$\leqslant \|y - y_0\| + C\|z - \Psi(x, y)\| \leqslant \|y - y_0\| +$$
$$+ C\{\|z - \Psi(x, y_0) - \Psi(x, y) + \Psi(x, y_0) + \Lambda(y - y_0) -$$
$$- \Lambda(y - y_0)\|\} \leqslant \|y - y_0\| + C\{\|z - \Psi(x, y_0)\| +$$
$$+ \|\Psi(x, y) - \Psi(x, y_0) - \Lambda(y - y_0)\| + \|\Lambda(y - y_0)\|\} \overset{(3)}{\leqslant}$$
$$\leqslant \|y - y_0\| + C\{\|z - \Psi(x, y_0)\| + \varepsilon\|y - y_0\| + \|\Lambda\|\|y - y_0\|\} \leqslant$$
$$\leqslant (1 + \theta + C\|\Lambda\|)\|y - y_0\| + C\|z - \Psi(x, y_0)\| \quad (5)$$

(ибо $\varepsilon C = \theta$!).

Из (5) в силу (4) получаем, во-первых, что

$$\|\Phi(x, z, y) - y_0\| <$$
$$< (1 + \theta + C\|\Lambda\|)\beta + C\frac{\beta(1-\theta)}{C} = (2 + C\|\Lambda\|)\beta = \beta_0. \quad (5')$$

Значит, выполнено условие а) леммы п 2.3.2: $(x, z, \Phi(x, z, y)) \in V$ Во-вторых, при $y = y_0$ получаем

$$\|\Phi(x, z_0, y_0) - y_0\| \leqslant C\|z - \Psi(x, y_0)\| \overset{(4)}{<} C\frac{\beta(1-\theta)}{C} = \beta(1-\theta). \quad (5'')$$

Значит, выполнено условие в) леммы п. 2.3.2.

Д) Если снова $(x, z) \in \mathcal{U}$, $y \in \mathring{B}(y_0, \beta)$, то

$$\Phi(x, z, \Phi(x, z, y)) - \Phi(x, z, y) \overset{\text{def}}{=}$$
$$= \Phi(x, z, y) + M(z - \Psi(x, \Phi(x, z, y))) - \Phi(x, z, y) =$$
$$= M(z - \Psi(x, \Phi(x, z, y))). \quad (6)$$

При этом вследствие (1)
$$\Lambda\Phi(x, z, y) = \Lambda y + \Lambda M(z - \Psi(x, y)) = \Lambda y + z - \Psi(x, y),$$
откуда
$$z = \Psi(x, y) + \Lambda(\Phi(x, z, y) - y). \quad (7)$$

Подставляя (7) в (6) и учитывая, что в силу (5′) $\Phi(x, z, y) \in \mathring{B}(y_0, \beta_0)$, получаем

$$\|\Phi(x, z, \Phi(x, z, y)) - \Phi(x, z, y)\| \overset{(6)}{=}$$
$$= \|M(z - \Psi(x, \Phi(x, z, y)))\| \overset{(2)}{\leqslant} C\|z - \Psi(x, \Phi(x, z, y))\| \overset{(7)}{=}$$
$$= C\|\Psi(x, \Phi(x, z, y)) - \Psi(x, y) - \Lambda(\Phi(x, z, y) - y)\| \overset{(3)}{\leqslant}$$
$$\leqslant C\varepsilon\|y - \Phi(x, z, y)\| = \theta\|y - \Phi(x, z, y)\|. \quad (8)$$

Таким образом, выполнено условие б) леммы п. 2.3.2.

Е) Еще раз сравним лемму п. 2.3.2 и наши построения. Топологическому пространству T в лемме соответствует у нас произведение $X \times Z$, точке t_0 — точка (x_0, z_0), окрестности U точки t_0 — окрестность $\mathcal{U}$ точки (x_0, z_0), окрестности V в лемме соответствует окрестность V, построенная в Б), отображению $(t, y) \mapsto \Phi(t, y)$ соответствует у нас отображение $(x, z, y) \mapsto \Phi(x, z, y)$, числу β в лемме соответствует число β, построенное в В).

Тогда соотношение (5′) означает, что выполнено условие а) леммы, соотношение (5″) — что выполнено условие в), наконец, соотношение (8) означает, что выполнено условие б). Значит, можно применять лемму.

Ж) По лемме, последовательность $\{y_n(x, z) \mid n \geqslant 0\}$, $(x, z) \in \mathcal{U}$, определяемая рекуррентно равенствами
$$y_0(x, z) = y_0, \quad y_{n+1}(x, z) = \Phi(x, z, y_n(x, z)),$$
содержится при всех $n \geqslant 0$ в шаре $\mathring{B}(y_0, \beta)$ и равномерно сходится к функции $\varphi(x, z)$, причем

$$\|\varphi(x, z) - y_0\| \leqslant \|\Phi(x, z, y_0) - y_0\| / (1 - \theta) \overset{\text{def}}{=}$$
$$= \|M(z - \Psi(x, y_0))\| / (1 - \theta) \overset{(2)}{\leqslant} \frac{C}{1 - \theta}\|z - \Psi(x, y_0)\|,$$

так что выполняется утверждение б) теоремы с $K = C/(1-\theta)$.

3) $\|\Psi(x, \varphi(x, z)) - z\| \leqslant$

$$\leqslant \|\Psi(x, \varphi(x, z)) - \Psi(x, y_n(x, z))\| + \|\Psi(x, y_n(x, z)) - z\| \overset{(1)}{\leqslant}$$
$$\leqslant \|\Psi(x, \varphi(x, z)) - \Psi(x, y_n(x, z)) - \Lambda(\varphi(x, z) - y_n(x, z))\| +$$
$$+ \|\Lambda(\varphi(x, z) - y_n(x, z))\| + \|\Lambda M(z - \Psi(x, y_n(x, z)))\| \overset{(3),(2)}{<}$$
$$< \varepsilon\|\varphi(x, z) - y_n(x, z)\| + \|\Lambda\|\|\varphi(x, z) - y_n(x, z)\| +$$
$$+ \|\Lambda\|\|y_{n+1}(x, z) - y_n(x, z)\| \to 0$$

при $n \to \infty$. Поэтому $\Psi(x, \varphi(x, z)) = z$. ■

2.3.4. Классические теоремы о неявной функции и об обратном отображении. Из доказанной выше общей теоремы легко выводятся две классические теоремы, постоянно используемые в различных прикладных для анализа областях, например в дифференциальной геометрии. По сравнению с п. 2.3.1 мы усилим в двух направлениях требования, предъявляемые к функции Ψ. Во-первых, теперь она будет предполагаться непрерывно дифференцируемой, а во-вторых, — и это наиболее существенно — оператор Λ, входящий в формулировку теоремы п. 2.3.1, будет предполагаться обратимым. Все это обеспечивает гладкость и единственность неявной функции.

А) Классическая теорема о неявной функции. *Пусть X, Y, Z — банаховы пространства, W — окрестность в $X \times Y$, $\Psi: W \to Z$ — отображение из класса $C^1(W)$. Если:*

1) $\Psi(x_0, y_0) = 0$; (1)

2) *существует обратный оператор $[\Psi_y(x_0, y_0)]^{-1} \in \mathscr{L}(Z, Y)$, то существуют такие $\varepsilon > 0$ и $\delta > 0$ и такое отображение φ: $\mathring{B}(x_0, \delta) \to Y$ класса $C^1(\mathring{B}(x_0, \delta))$, что:*

а) $$\varphi(x_0) = y_0; \tag{2}$$

б) $\|x - x_0\| < \delta \Rightarrow \|\varphi(x) - y_0\| < \varepsilon$ *и* $\Psi(x, \varphi(x)) \equiv 0$;

в) *в «прямоугольнике» $\mathring{B}(x_0, \delta) \times \mathring{B}(y_0, \varepsilon)$ равенство $\Psi(x, y) = 0$ возможно только при $y = \varphi(x)$;*

г) $$\varphi'(x) = -[\Psi_y(x, \varphi(x))]^{-1}\Psi_x(x, \varphi(x)). \tag{3}$$

Доказательство. Из непрерывной дифференцируемости Ψ в W следует выполнение условий 1) и 2) теоремы существования неявной функции из п. 2.3.1 (для проверки 2) надо положить $\Lambda = \Psi_y(x_0, y_0)$ и воспользоваться теоремой о среднем). Условие 3) выполнено, поскольку существует $\Psi_y^{-1}(x_0, y_0)$. Значит, по этой теореме найдутся число $K > 0$, окрестность $U \ni x_0$ и отображение $\varphi\colon U \to Y$ такие, что

$$\Psi(x, \varphi(x)) \equiv 0, \tag{4}$$

$$\|\varphi(x) - y_0\| \leqslant K \|\Psi(x, y_0)\|. \tag{5}$$

Полагая в (5) $x = x_0$ и используя (1), получаем (2), а (4) дает вторую половину утверждения б).

Б) Отображение $\Psi \in C^1(W)$ и потому строго дифференцируемо в точке (x_0, y_0) (следствие 2 п. 2.2.3). Следовательно, для всякого $\varkappa > 0$ найдется такое $\varepsilon(\varkappa) > 0$, что

$$\|x_i - x_0\| < \varepsilon(\varkappa), \ \|y_i - y_0\| < \varepsilon(\varkappa), \quad i = 1, 2, \Rightarrow$$
$$\Rightarrow \|\Psi(x_1, y_1) - \Psi(x_2, y_2) - \Psi_x(x_0, y_0)(x_1 - x_2) -$$
$$- \Psi_y(x_0, y_0)(y_1 - y_2)\| \leqslant \varkappa \max\{\|x_1 - x_2\|, \|y_1 - y_2\|\}. \tag{6}$$

Сначала положим $\varkappa_0 = {}^1/_2 \|\Psi_y^{-1}(x_0, y_0)\|^{-1}$ и найдем соответствующее $\varepsilon = \varepsilon(\varkappa_0)$. Поскольку $\Psi(x, y)$ непрерывна и $\Psi(x_0, y_0) = 0$, найдется такое δ, $0 < \delta < \varepsilon$, что

$$\mathring{B}(x_0, \delta) = \{x \mid \|x - x_0\| < \delta\} \subset U$$

и

$$\|x - x_0\| < \delta \Rightarrow \|\Psi(x, y_0)\| < \frac{\varepsilon}{K} \overset{(5)}{\Rightarrow} \|\varphi(x) - y_0\| < \varepsilon, \tag{7}$$

а следовательно, выполняется первая половина утверждения б).

Предположим теперь, что $\Psi(\hat{x}, \hat{y}) = 0$ в точке $(\hat{x}, \hat{y}) \in \in \mathring{B}(x_0, \delta) \times \mathring{B}(y_0, \varepsilon)$. Применяя (6) с указанным выше $\varkappa_0$ $(x_1 = x_2 = \hat{x},\ y_1 = \hat{y},\ y_2 = \varphi(\hat{x}))$, имеем

$$\|\hat{y} - \varphi(\hat{x})\| = \|\Psi_y^{-1}(x_0, y_0)\Psi_y(x_0, y_0)(\hat{y} - \varphi(\hat{x}))\| \leqslant$$
$$\leqslant \|\Psi_y^{-1}(x_0, y_0)\| \|\Psi(\hat{x}, \hat{y}) - \Psi(\hat{x}, \varphi(\hat{x})) - \Psi_y(x_0, y_0)(\hat{y} - \varphi(\hat{x}))\| \overset{(6)}{\leqslant}$$
$$\leqslant \|\Psi_y(x_0, y_0)^{-1}\| \varkappa_0 \|\hat{y} - \varphi(\hat{x})\| = \frac{1}{2}\|\hat{y} - \varphi(\hat{x})\|,$$

что возможно лишь при $\hat{y} = \varphi(\hat{x})$, т. е. верно утверждение в).

В) При тех же $\varkappa_0$, ε, δ положим в (6) $x_1=x$, $x_2=x_0$, $y_1=y_2=y_0$. Тогда

$$\|x-x_0\|<\delta\overset{(6)}{\Longrightarrow}\|\Psi(x, y_0)-\Psi(x_0, y_0)-\Psi_x(x_0, y_0)(x-x_0)\|\leqslant$$
$$\leqslant\varkappa_0\|x-x_0\|\Rightarrow\|\Psi(x, y_0)\|\leqslant(\|\Psi_x(x_0, y_0)\|+\varkappa_0)\|x-x_0\|\overset{(5)}{\Longrightarrow}$$
$$\Rightarrow\|\varphi(x)-y_0\|\leqslant K(\|\Psi_x(x_0, y_0)\|+\varkappa_0)\|x-x_0\|. \qquad (8)$$

Теперь зададим произвольно $\varkappa_1>0$ и применим (6) к

$$x_1=x,\ x_2=x_0,\quad y_1=\varphi(x),\quad y_2=y_0.$$

Для

$$\delta_1=\min\left\{\varepsilon(\varkappa_1),\quad \frac{\varepsilon(\varkappa_1)}{K(\|\Psi_x(x_0, y_0)\|+\varkappa_0)}\right\}$$

имеем

$$\|x-x_0\|<\delta_1\Rightarrow\|\Psi_y^{-1}(x_0, y_0)\Psi_x(x_0, y_0)(x-x_0)+\varphi(x)-y_0\|=$$
$$=\|\Psi_y^{-1}(x_0, y_0)\{\Psi_x(x_0, y_0)(x-x_0)+\Psi_y(x_0, y_0)(\varphi(x)-y_0)\}\|\leqslant$$
$$\leqslant\|\Psi_y^{-1}(x_0, y_0)\|\|\Psi(x, \varphi(x))-\Psi(x_0, y_0)-$$
$$-\Psi_x(x_0, y_0)(x-x_0)-\Psi_y(x_0, y_0)(\varphi(x)-y_0)\|\overset{(6)}{\leqslant}$$
$$\leqslant\|\Psi_y^{-1}(x_0, y_0)\|\varkappa_1\max\{\|x-x_0\|, \|\varphi(x)-y_0\|\}\overset{(8)}{\leqslant}C\varkappa_1\|x-x_0\|,$$

где

$$C=\|\Psi_y^{-1}(x_0, y_0)\|\max\{1, K(\|\Psi_x(x_0, y_0)\|+\varkappa_0)\}.$$

Поскольку $\varkappa_1$ можно выбирать произвольно, доказанное неравенство означает, что при $x\to x_0$

$$\Psi_y^{-1}(x_0, y_0)\Psi_x(x_0, y_0)(x-x_0)+\varphi(x)-y_0=o(x-x_0)$$

или

$$\varphi(x)=y_0+[-\Psi_y^{-1}(x_0, y_0)\Psi_x(x_0, y_0)](x-x_0)+o(x-x_0),$$

а следовательно, существует производная Фреше $\varphi'(x_0)$, и равенство (3) имеет место при $x=x_0$.

Чтобы доказать дифференцируемость φ при остальных x, напомним, что множество операторов Λ, для которых существует обратный, открыто в $\mathscr{L}(X, Y)$ (предложение 3 в п. 2.2.1). Поэтому, уменьшив, если нужно, δ, мы можем считать, что $\Psi_y^{-1}(x, \varphi(x))$ существует при $\|x-x_0\|<\delta$.

Теперь можно повторить все доказательство, заменив точку (x_0, y_0) точкой $(\hat{x}, \varphi(\hat{x}))$, где $\|\hat{x}-x_0\|<\delta$. В силу доказанной выше единственности мы получим ту же самую функцию $\varphi(x)$ (по крайней мере, если $\|x-\hat{x}\|$ достаточно мало) и для нее формула (3) будет верна при $x=\hat{x}$. Так как $\hat{x}$ было произвольно, то утверждение г) имеет место при всех $x \in \mathring{B}(x_0, \delta)$.

Наконец, производная $\varphi'(x)$ непрерывна, так как, согласно (3), она представляется в виде суперпозиции непрерывных отображений $x \mapsto (x, \varphi(x))$, $(x, y) \mapsto \Psi_x(x, y)$, $(x, y) \mapsto \Psi_y(x, y)$ и $A \mapsto A^{-1}$ (последнее непрерывно, поскольку оно дифференцируемо по Фреше согласно предложению 3 п. 2.2.1). ■

Замечание. Если в дополнение к условиям теоремы

$$\Psi \in C^r(W), \quad r \geqslant 2, \text{ то } \quad \varphi \in C^r(\mathring{B}(x_0, \delta)).$$

Действительно, из формулы (3) и теоремы о суперпозиции можно вывести, что

$$\varphi''(x) = -\{\Psi_y(x, \varphi(x))^{-1} \circ \Psi_x(x, \varphi(x))\}'$$

существует и непрерывна. Далее рассуждаем по индукции.

Теорема об обратном отображении. *Пусть Y и Z — банаховы пространства, W — окрестность точки y_0 в Y, Ψ: $W \to Z$ — отображение класса $C^1(W)$, $z_0 = \Psi(y_0)$. Если существует обратный оператор $\Psi'(y_0)^{-1} \in \mathscr{L}(Z, Y)$, то:*

а) *существует такое $\varepsilon > 0$, что $\mathring{B}(y_0, \varepsilon) \subset W$ и $V = \{z \mid z = \Psi(y), \quad \|y-y_0\| < \varepsilon\} = \Psi(\mathring{B}(y_0, \quad \varepsilon))$ — окрестность z_0 в Z;*

б) *существует отображение Φ: $V \to Y$ класса $C^1(V)$, обратное к $\Psi \mid \mathring{B}(y_0, \varepsilon)$, т. е. такое, что*

$$y = \Phi(z) \Leftrightarrow z = \Psi(y), \quad (y, z) \in \mathring{B}(y_0, \varepsilon) \times V; \tag{9}$$

в)
$$\Phi'(z) = [\Psi'(\Phi(z))]^{-1}. \tag{10}$$

Доказательство. А). Отображение $\Psi \in C^1(W)$ строго дифференцируемо в каждой точке W, в частности и в y_0. Поэтому для всякого $\varkappa > 0$ найдется $\varepsilon(\varkappa) > 0$ такое, что

$$\|y_i - y_0\| < \varepsilon(\varkappa), \quad i = 1, 2, \Rightarrow$$
$$\Rightarrow \|\Psi(y_1) - \Psi(y_2) - \Psi'(y_0)(y_1 - y_2)\| \leqslant \varkappa \|y_1 - y_2\|. \tag{11}$$

Выберем $\varepsilon > 0$ так, чтобы для всех $y \in \mathring{B}(y_0, \varepsilon)$ существовал обратный оператор $\Psi'(y)^{-1}$ (см. предложение 3 п. 2 2 1) и чтобы было $\varepsilon < \varepsilon(\varkappa_0)$, где $\varkappa_0 = {}^1/_2 \|\Psi'(y_0)^{-1}\|^{-1}$.

Последнее обеспечивает инъективность отображения Ψ на $\mathring{B}(y_0, \varepsilon)$:

$$y_1,\ y_2 \in \mathring{B}(y_0, \varepsilon),\ \Psi(y_1) = \Psi(y_2) \Rightarrow$$
$$\Rightarrow \|y_1 - y_2\| = \|\Psi'(y_0)^{-1}\Psi'(y_0)(y_1 - y_2)\| \leqslant$$
$$\leqslant \|\Psi'(y_0)^{-1}\| \|\Psi(y_1) - \Psi(y_2) - \Psi'(y_0)(y_1 - y_2)\| \overset{(11)}{\leqslant}$$
$$\leqslant \|\Psi'(y_0)^{-1}\| \frac{1}{2\|\Psi'(y_0)^{-1}\|} \|y_1 - y_2\| = \frac{1}{2}\|y_1 - y_2\| \Rightarrow y_1 = y_2.$$

Далее мы докажем, что шар $\mathring{B}(y_0, \varepsilon)$ — тот самый, о котором говорится в п. а) теоремы, а пока положим $V = \Psi(\mathring{B}(y_0, \varepsilon))$. Ввиду инъективности отображения $\Psi \mid \mathring{B}(y_0, \varepsilon)$ оно обратимо, так что существует $\Phi\colon V \to \mathring{B}(y_0, \varepsilon)$, обладающее свойством (9).

Б) Для любой точки $\hat{y} \in \mathring{B}(y_0, \varepsilon)$ применим теорему п. 2.3.1 к отображению $\Psi(y)$ (не зависящему от x, так что пространство X может быть любым) в окрестности точки $\hat{y}$. Условие 1) этой теоремы здесь тривиально; условие 2) превращается в условие строгой дифференцируемости в точке $\hat{y}$ и, как было отмечено, выполнено, причем $\Lambda = \Psi'(\hat{y})$; условие 3) выполнено, поскольку $\hat{y} \in \mathring{B}(y_0, \varepsilon)$ и в соответствии с выбором ε существует $\Psi'(\hat{y})^{-1}$.

По теореме п. 2.3.1 существует окрестность U точки $\hat{z} = \Psi(\hat{y})$ и отображение $\varphi\colon U \to Y$ такие, что

$$\Psi(\varphi(z)) \equiv z, \tag{12}$$
$$\|\varphi(z) - \hat{y}\| \leqslant K\|z - \hat{z}\| \tag{13}$$

для некоторого $K > 0$. Возьмем $\delta > 0$ столь малым, чтобы $\mathring{B}(\hat{z}, \delta) \subset U$ и чтобы было $\delta < (\varepsilon - \|y_0 - \hat{y}\|)/K$. Тогда

$$\|z - \hat{z}\| < \delta \overset{(13)}{\Rightarrow} \|\varphi(z) - \hat{y}\| < K\left(\frac{\varepsilon - \|y_0 - \hat{y}\|}{K}\right) \Rightarrow$$
$$\Rightarrow \|\varphi(z) - y_0\| < \varepsilon \Rightarrow z = \Psi(\varphi(z)) \in \Psi(\mathring{B}(y_0, \varepsilon)) =$$
$$= V \Rightarrow \mathring{B}(\hat{z}, \delta) \subset V.$$

Таким образом, каждая точка $\hat{z} \in V$ имеет окрестность $\mathring{B}(\hat{z}, \delta) \subset V$, так что V открыто. Кроме того, для $z \in \mathring{B}(z_0, \delta)$

$$\Psi(\varphi(z)) \overset{(12)}{=} z = \Psi(\Phi(z)) \Rightarrow \varphi(z) = \Phi(z) \tag{14}$$

в силу инъективности Ψ.

В) Докажем теперь дифференцируемость отображения Φ и формулу (10). В силу строгой дифференцируемости Ψ в точке $\hat{y}=\Phi(\hat{z})$ имеет место аналог (11): для всякого $\varkappa>0$ найдется $\varepsilon(\varkappa)>0$ такое, что

$$\|y_i-\hat{y}\|<\varepsilon(\varkappa), \quad i=1,\ 2, \Rightarrow$$
$$\Rightarrow \|\Psi(y_1)-\Psi(y_2)-\Psi'(\hat{y})(y_1-y_2)\| \leqslant \varkappa\|y_1-y_2\|. \tag{15}$$

Полагая здесь $y_1=\Phi(z)$, $y_2=\hat{y}=\Phi(\hat{z})$, имеем

$$\|z-\hat{z}\|<\frac{\varepsilon(\varkappa)}{K} \overset{(13)\,(14)}{\Longrightarrow} \|\Phi(z)-\hat{y}\|<\varepsilon(\varkappa) \overset{(15)}{\Rightarrow}$$
$$\Rightarrow \|\Psi(\Phi(z))-\Psi(\Phi(\hat{z}))-\Psi'(\hat{y})(\Phi(z)-\Phi(\hat{z})\| \leqslant$$
$$\leqslant \varkappa\|\Phi(z)-\hat{y}\| \overset{(13)}{\leqslant} \varkappa K\|z-\hat{z}\| \overset{(12)}{\Rightarrow}$$
$$\Rightarrow \|z-\hat{z}-\Psi'(\hat{y})(\Phi(z)-\Phi(\hat{z}))\| \leqslant \varkappa K\|z-\hat{z}\| \Rightarrow$$
$$\Rightarrow \|\Phi(z)-\Phi(\hat{z})-\Psi'(\hat{y})^{-1}(z-\hat{z})\|=$$
$$=\|\Psi'(\hat{y})^{-1}(z-\hat{z}-\Psi'(\hat{y})(\Phi(z)-\Phi(\hat{z})))\| \leqslant$$
$$\leqslant \|\Psi'(\hat{y})^{-1}\|\varkappa K\|z-\hat{z}\|,$$

и потому

$$\Phi(z)-\Phi(\hat{z})-\Psi'(\hat{y})^{-1}(z-\hat{z})=o(z-\hat{z})$$

при $z\to\hat{z}$, так как $\varkappa$ произвольно. Поэтому существует производная Фреше $\Phi'(\hat{z})=\Psi'(\hat{y})^{-1}$, так что (10) доказано. Производная $\Phi'(z)$ непрерывна на V, так как, согласно (10), она представляется в виде суперпозиции непрерывных отображений $z\mapsto\Phi(z)$, $y\mapsto\Psi'(y)$ и $A\mapsto A^{-1}$ (последнее непрерывно в силу предложения 3 п. 2.2.1).

2.3.5. Касательное пространство и теорема Люстерника. В этом пункте X — нормированное пространство, M — некоторое его подмножество.

Определение. Элемент $h\in X$ называется *касательным вектором к M в точке $x_0\in M$*, если существуют такие $\varepsilon>0$ и отображение r: $[-\varepsilon,\ \varepsilon]\to X$, что:

а) $x_0+\alpha h+r(\alpha)\in M$,

б) $r(\alpha)=o(\alpha)$ при $\alpha\to 0$.

Элемент $h\in X$ называется *односторонним касательным вектором к M в точке $x_0\in M$*, если существуют $\varepsilon>0$ и отображение r: $(0,\ \varepsilon)\to X$ такие, что выполнено а), и б) имеет место при $\alpha\downarrow 0$.

Множество всех касательных векторов к M в точке x_0 обозначается $T_{x_0}M$, множество односторонних касательных векторов — $T^+_{x_0}M$. Очевидно, что $T_{x_0}M \subset T^+_{x_0}M$ и что $T_{x_0}M = T^+_{x_0}M \cap (-T^+_{x_0}M)$. Если множество $T_{x_0}M$ является подпространством в X, то оно называется *касательным пространством* к M в точке x_0.

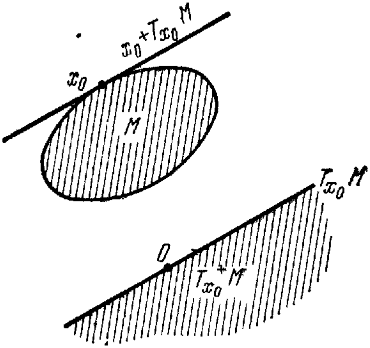


Рис. 33

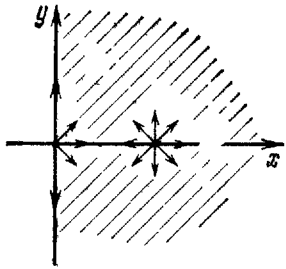


Рис. 34.

Замечание. В геометрии обычно касательной прямой, плоскостью и т д. называют не $T_{x_0}M$, а аффинное многообразие $x_0 + T_{x_0}M$ (рис. 33).

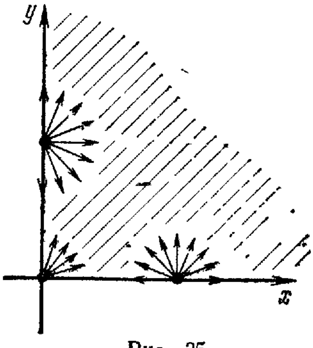


Рис. 35.

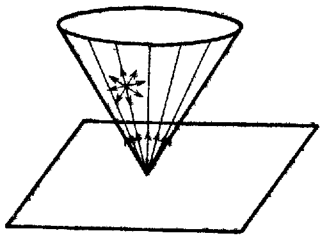


Рис. 36

Примеры (подробные доказательства предоставляются читателю).

1) $X = \mathbf{R}^2$, $M = \{(x, y) \mid x \geqslant 0\}$ (рис. 34); $T^+_{(0,0)}M = M$, $T_{(0,0)}M = \{(0, b) \mid b \in \mathbf{R}\}$, $T_{(1,0)}M = T^+_{(1,0)}M = \mathbf{R}^2$.

2) $X = \mathbf{R}^2$, $M = \{(x, y) \mid x \geqslant 0,\ y \geqslant 0\} = \mathbf{R}^2_+$ (рис. 35); $T_{(1,0)}M = \{(a, 0) \mid a \in \mathbf{R}\}$, $T_{(0,0)}M = \{0\}$, $T_{(0,1)}M = \{(0, b) \mid b \in \mathbf{R}\}$, $T^+_{(1,0)}M = \{(a, b) \mid a \in \mathbf{R},\ b \geqslant 0\}$, $T^+_{(0,0)}M = M$, $T^+_{(0,1)}M = \{(a, b) \mid a \geqslant 0,\ b \in \mathbf{R}\}$.

3) $X = Y \times \mathbf{R}$, где Y — нормированное пространство, $M = \{(y, z) \mid z = \|y\|\}$ — конус (рис. 36).

$$T_{(0,0)}M = \{0\}, \quad T^+_{(0,0)}M = M.$$

Упражнения.

1 Предполагая, что Y — гильбертово пространство (например, $Y = \mathbf{R}^2$), найдите $T_{(y, \|y\|)}M$ и $T^+_{(y, \|y\|)}M$ в примере 3) при $y \neq 0$ (рис. 36).

2. Найдите $T_a M$ и $T^+_a M$, где:

а) $M = \{2^{-n} \mid n = 1, 2, \ldots\} \cup \{0\} \subset \mathbf{R}$, $a = 0$;

б) $M = \left\{\frac{1}{n} \middle| n = 1, 2, \ldots\right\} \cup \{0\} \subset \mathbf{R}$, $a = 0$;

в) $M = \{(x, y) \mid x^2 = y^3\} \subset \mathbf{R}^2$, $a = (0, 0)$ и $a = (1, 1)$;

г) $M = \{(x, y) \mid x^2 \leqslant y^3\} \subset \mathbf{R}^2$, $a = (0, 0)$ и $a = (1, 1)$;

д) $M = \{(x, y) \mid x^2 \geqslant y^3\} \subset \mathbf{R}^2$, $a = (0, 0)$ и $a = (1, 1)$;

е) $M = \{(x, y, z) \mid z = 1$, если x и y рациональны, $z = 0$ в остальных случаях$\}$, $a = (0, 0, 1)$

Во многих случаях, в том числе и представляющих интерес для теории экстремальных задач, множество касательных векторов может быть найдено при помощи следующей общей теоремы.

Теорема Люстерника. *Пусть X, Z — банаховы пространства; U — окрестность точки x_0 в X, F: $U \to Z$, $F(x_0) = 0$.*

Если F строго дифференцируемо в точке x_0 и $F'(x_0)$ является эпиморфизмом, то множество $M = \{x \mid F(x) = 0\}$ имеет в точке x_0 касательное пространство

$$T_{x_0}M = T^+_{x_0}M = \operatorname{Ker} F'(x_0).$$

Доказательство. А) Пусть $h \in T^+_{x_0}M$. Тогда, если $r(\cdot)$ — отображение из определения одностороннего касательного вектора, то в силу строгой дифференцируемости

$$0 = F(x_0 + \alpha h + r(\alpha)) = F(x_0) + \alpha F'(x_0)h + o(\alpha) = \\ = \alpha F'(x_0)h + o(\alpha).$$

Значит, $F'(x_0)h = 0$ и $T^+_{x_0}M \subset \operatorname{Ker} F'(x_0)$.

Б) Применим теорему п. 2.3.1 к отображению $\Psi(x, y) = F(x + y)$. Вследствие строгой дифференцируемости F отображение $x \mapsto \Psi(x, 0)$ непрерывно в точке x_0 и

$$\|\Psi(x, y') - \Psi(x, y'') - F'(x_0)(y' - y'')\| \leqslant \varepsilon \|y' - y''\|,$$

если $\|x - x_0\| < \delta$, $\|y'\| < \delta$, $\|y''\| < \delta$. Условие 3) теоремы п. 2.3.1 также выполняется, ибо $F'(x_0)$ отображает X

на Z. В соответствии с этой теоремой существует отображение $\varphi\colon U \to X$ окрестности $U \subset X$ точки x_0 такое, что

$$\Psi(x, \varphi(x)) \equiv 0 \Leftrightarrow F(x+\varphi(x)) \equiv 0,$$
$$\|\varphi(x)\| \leqslant K\|\Psi(x, 0)\| = K\|F(x)\|.$$

Остается лишь положить $r(\alpha) = \varphi(x_0+\alpha h)$. Тогда,

$$F(x_0+\alpha h+r(\alpha)) = F(x_0+\alpha h+\varphi(x_0+\alpha h)) = 0,$$
$$\|r(\alpha)\| = \|\varphi(x_0+\alpha h)\| \leqslant K\|F(x_0+\alpha h)\| =$$
$$= K\|F(x_0+\alpha h)-F(x_0)\| = K\|F'(x_0)\alpha h+o(\|\alpha h\|)\| =$$
$$= K\|\alpha F'(x_0)h+o(\alpha)\|,$$

и если $h \in \operatorname{Ker} F'(x_0)$, то $\|r(\alpha)\| = o(\alpha)$. Следовательно,

$$\operatorname{Ker} F'(x_0) \subset T_{x_0}M \subset T^+_{x_0}M. \blacksquare$$

На уровне геометрической интуиции теорема Люстерника выглядит весьма прозрачно. Пренебрегая строгостью, можно определить касательное пространство L к множеству M в точке x_0 тем, что «с точностью до бесконечно малых высшего порядка» M аппроксимируется аффинным многообразием x_0+L. Но «в бесконечно малой окрестности» точки x_0 функция F аппроксимируется своей линейной частью:

$$F(x) \approx F(x_0)+F'(x_0)(x-x_0) = F'(x_0)(x-x_0) \qquad (1)$$

(точка $x_0 \in M$, и потому $F(x_0)=0$), а множество $M = \{x \mid F(x)=0\}$ аппроксимируется множеством

$$\tilde{M} = \{x \mid F'(x_0)(x-x_0)=0\} = \operatorname{Ker} F'(x_0)+x_0.$$

Таким образом, «с точностью до бесконечно малых высшего порядка» M совпадает с аффинным многообразием $\tilde{M} = x_0+\operatorname{Ker} F'(x_0)$, и, следовательно, $\operatorname{Ker} F'(x_0)$ — касательное пространство.

§ 2.4. Дифференцируемость некоторых конкретных отображений

2.4.1. Оператор Немыцкого и оператор дифференциальной связи. Пусть U — открытое множество в $\mathbb{R}\times\mathbb{R}^n$, и пусть функция $f(t, x)\colon U \to \mathbb{R}^m$ и ее частная производная $f_x(t, x)$ непрерывны в U. Зададим равенством

$$\mathcal{N}(x(\cdot))(t) = f(t, x(t)), \quad t_0 \leqslant t \leqslant t_1 \qquad (1)$$

отображение $\mathcal{N}$: $\mathcal{U} \to C([t_0, t_1], \mathbf{R}^m)$, определенное на множестве

$$\mathcal{U} = \{x(\cdot) \in C([t_0, t_1], \mathbf{R}^n) \mid (t, x(t)) \in U, \ t_0 \leqslant t \leqslant t_1\}. \qquad (2)$$

Это отображение мы будем называть *оператором Немыцкого.*

Упражнение 1. Проверьте, что множество (2) открыто в $C([t_0, t_1], \mathbf{R}^n)$.

Предложение 1. *Оператор Немыцкого, заданный соотношением* (1), *непрерывно дифференцируем на множестве* (2) *и при этом*

$$\mathcal{N}'(\hat{x}(\cdot))[h(\cdot)](t) = \hat{f}_x(t)\, h(t), \qquad (3)$$

где

$$\hat{f}_x(t) \Leftrightarrow \left(\frac{\partial f_i}{\partial x_j}(t, \hat{x}(t)) \quad \begin{matrix} i = 1, \ldots, m, \\ j = 1, \ldots, n \end{matrix} \right) \qquad (4)$$

— матрица Якоби.

Доказательство. Вычислим сначала первую вариацию по Лагранжу отображения $\mathcal{N}$. По определению

$$\delta \mathcal{N}(\hat{x}(\cdot); h(\cdot)) = \lim_{\alpha \to 0} \frac{\mathcal{N}(\hat{x}(\cdot) + \alpha h(\cdot)) - \mathcal{N}(\hat{x}(\cdot))}{\alpha}.$$

Слева и справа здесь стоят элементы пространства $C([t_0, t_1], \mathbf{R}^m)$, т. е. некоторые функции от $t \in [t_0, t_1]$. При фиксированном t имеем

$$\lim_{\alpha \to 0} \frac{\mathcal{N}(\hat{x}(\cdot) + \alpha h(\cdot))(t) - \mathcal{N}(\hat{x}(\cdot))(t)}{\alpha} =$$

$$= \lim_{\alpha \to 0} \frac{f(t, \hat{x}(t) + \alpha h(t)) - f(t, \hat{x}(t))}{\alpha} =$$

$$= f_x(t, \hat{x}(t))\, h(t) = \hat{f}_x(t)\, h(t). \qquad (5)$$

Чтобы убедиться в том, что сходимость здесь имеет место также в смысле пространства $C([t_0, t_1], \mathbf{R}^m)$, т. е. что она равномерна, заметим, что для некоторого $a > 0$ компакт

$$\mathcal{K} = \{(t, x) \mid |x - \hat{x}(t)| \leqslant a, \ t_0 \leqslant t \leqslant t_1\} \subset U$$

и на этом компакте f_x равномерно непрерывна. Это значит, что

$$\forall \varepsilon > 0, \exists \delta > 0, \quad |t' - t''| < \delta, \quad |x' - x''| < \delta \Rightarrow$$
$$\Rightarrow \|f_x(t', x') - f_x(t'', x'')\| < \varepsilon.$$

Если $|\alpha| < \delta/\|h(\cdot)\|_C$, то по теореме о среднем (п. 2.2.3)

$$\left\|\frac{\mathcal{N}(\hat{x}(\cdot)+\alpha h(\cdot))-\mathcal{N}(\hat{x}(\cdot))}{\alpha}-\{f_x(t)h(t)\}\right\|=$$
$$=\max_{t\in[t_0,t_1]}\left\|\frac{f(t,\hat{x}(t)+\alpha h(t))-f(t,\hat{x}(t))-f_x(t,\hat{x}(t))h(t)}{\alpha}\right\|\leqslant$$
$$\leqslant\max_{t\in[t_0,t_1]}\max_{\theta\in[0,1]}\|f_x(t,\hat{x}(t)+\theta\alpha h(t))-f_x(t,\hat{x}(t))\|\,|h(t)|\leqslant$$
$$\leqslant\varepsilon\|h\|_C. \quad (6)$$

Поэтому сходимость (5) действительно равномерна и доказано существование первой вариации по Лагранжу

$$\delta\mathcal{N}(\hat{x}(\cdot),\ h(\cdot))(t)\equiv\hat{f}_x(t)h(t).$$

Поскольку первая вариация задается линейным оператором

$$\mathcal{N}'_\Gamma(\hat{x}(\cdot))\colon\ C([t_0,t_1],\ \mathbf{R}^n)\to C([t_0,t_1],\ \mathbf{R}^m),$$

а именно $\mathcal{N}'_\Gamma(\hat{x}(\cdot))$ — это оператор умножения на матричную функцию $\hat{f}_x(t)$ и этот оператор ограничен:

$$\|\mathcal{N}'_\Gamma(\hat{x}(\cdot))\|\leqslant\max_{t\in[t_0,t_1]}\|f_x(t,\hat{x}(t))\|, \quad (7)$$

то $\mathcal{N}(x(\cdot))$ дифференцируем по Гато. Чтобы доказать непрерывность производной Гато, оценим норму разности

$$\|\mathcal{N}'_\Gamma(x(\cdot))-\mathcal{N}'_\Gamma(\hat{x}(\cdot))\|=$$
$$=\sup_{\|h(\cdot)\|_C\leqslant1}\|\mathcal{N}'_\Gamma(x(\cdot))[h(\cdot)]-\mathcal{N}'_\Gamma(\hat{x}(\cdot))[h(\cdot)]\|=$$
$$=\sup_{\|h(\cdot)\|_C\leqslant1}\max_{t\in[t_0,t_1]}\|f_x(t,x(t))h(t)-f_x(t,\hat{x}(t))h(t)\|\leqslant$$
$$\leqslant\max_{t\in[t_0,t_1]}\|f_x(t,x(t))-f_x(t,\hat{x}(t))\|.$$

Снова с помощью компакта $\mathcal{K}$ и равномерной непрерывности f_x, убеждаемся в непрерывности $\mathcal{N}'_\Gamma(x(\cdot))$ относительно $x(\cdot)\in\mathcal{U}$. Применяя теперь следствие 2 из п. 2.2.3, мы убеждаемся, что $\mathcal{N}(x(\cdot))$ имеет в каждой точке множества $\mathcal{U}$ производную Фреше и строго дифференцируемо. Ввиду равенства $\mathcal{N}'_\Gamma=\mathcal{N}'$ производная Фреше непрерывна в $\mathcal{U}$. ■

Упражнение 2. Докажите, что в (7) имеет место равенство.

Пусть теперь U — открытое множество в $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^r$, и пусть функция $f(t,x,u)\colon U\to\mathbf{R}^m$ и ее частные производные f_x, f_u непрерывны в U. Отображение $\mathcal{N}\colon \mathcal{U} \to C([t_0, t_1], \mathbf{R}^m)$, определенное на множестве

$$\mathcal{U} = \{(x(\cdot), u(\cdot)) \mid (t, x(t), u(t)) \in U,\ t_0 \leqslant t \leqslant t_1\} \subset \\ \subset C^1([t_0, t_1], \mathbf{R}^n) \times C([t_0, t_1], \mathbf{R}^r) \quad (8)$$

равенством

$$\mathcal{N}(x(\cdot), u(\cdot))(t) \equiv f(t, x(t), u(t)), \quad (9)$$

мы также будем называть *оператором Немыцкого.*

Предложение 2. *Оператор Немыцкого, заданный соотношением* (8) *непрерывно дифференцируем на множестве* (9) *и при этом*

$$\mathcal{N}'(\hat{x}(\cdot), \hat{u}(\cdot))[h(\cdot), v(\cdot)](t) = \hat{f}_x(t) h(t) + \hat{f}_u(t) v(t), \quad (10)$$

где

$$\hat{f}_x(t) \Leftrightarrow \left(\frac{\partial f_i(t, \hat{x}(t), \hat{u}(t))}{\partial x_j},\ i = 1, \ldots, m,\ j = 1, \ldots, n\right), \quad (11)$$

$$\hat{f}_u(t) \Leftrightarrow \left(\frac{\partial f_i(t, \hat{x}(t), \hat{u}(t))}{\partial u_k},\ i = 1, \ldots, m,\ k = 1, \ldots, r\right). \quad (12)$$

Доказательство. Как и в предыдущем случае, у отображения $\mathcal{N}$ существуют частные производные

$$\mathcal{N}_x(\hat{x}(\cdot), \hat{u}(\cdot))[h(\cdot)](t) = \hat{f}_x(t) h(t),$$
$$\mathcal{N}_u(\hat{x}(\cdot), \hat{u}(\cdot))[v(\cdot)](t) = \hat{f}_u(t) v(t),$$

и отображения

$$\mathcal{N}_x\colon \mathcal{U} \to \mathscr{L}(C^1([t_0, t_1], \mathbf{R}^n), C([t_0, t_1], \mathbf{R}^m)),$$
$$\mathcal{N}_u\colon \mathcal{U} \to \mathscr{L}(C^1([t_0, t_1], \mathbf{R}^r), C([t_0, t_1], \mathbf{R}^m))$$

непрерывны в $\mathcal{U}$. Остается сослаться на теорему о полном дифференциале из п. 2.2.4. ■

Пусть далее U и $\mathcal{U}$ — те же, что и в предыдущем примере, функция $\varphi(t, x, u)\colon U \to \mathbf{R}^n$ и ее частные производные φ_x, φ_u непрерывны в U. Отображение $\Phi\colon \mathcal{U} \to C([t_0, t_1], \mathbf{R}^n)$, определенное равенством

$$\Phi(x(\cdot), u(\cdot))(t) = \dot{x}(t) - \varphi(t, x(t), u(t)), \quad (13)$$

будем называть *оператором дифференциальной связи.*

Предложение 3. *Оператор дифференциальной связи, заданный соотношением* (13), *непрерывно дифференцируем* *на множестве* (8) *и при этом*

$$\Phi'(\hat{x}(\cdot),\ \hat{u}(\cdot))[h(\cdot),\ v(\cdot)](t)= \\ =\dot{h}(t)-\hat{\varphi}_x(t)h(t)-\hat{\varphi}_u(t)v(t), \quad (14)$$

где

$$\hat{\varphi}_x(t)\Leftrightarrow\left(\frac{\partial\varphi_i(t,\ \hat{x}(t),\ \hat{u}(t))}{\partial x_j},\ i,\ j=1,\ \ldots,\ n\right), \quad (15)$$

$$\hat{\varphi}_u(t)\Leftrightarrow\left(\frac{\partial\varphi_i(t,\ \hat{x}(t),\ \hat{u}(t))}{\partial u_k},\ i=1,\ldots,n,\ k=1,\ \ldots,\ r\right). \quad (16)$$

Доказательство. Отображение Φ есть разность линейного непрерывного оператора $(x(\cdot),\ u(\cdot))\mapsto\dot{x}(\cdot)$ и оператора Немыцкого $\mathcal{N}(x(\cdot),\ u(\cdot))(t)=\varphi(t,\ x(t),\ u(t))$. Поэтому доказываемое утверждение следует из общих свойств производных (п. 2.2.1) и предложения 2. ■

2.4.2. Интегральный функционал. Пусть U — открытое множество в $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^n$, и пусть функция $f(t,x,\dot{x})\colon U\to\mathbf{R}^m$ и ее частные производные f_x, $f_{\dot{x}}$ непрерывны в U. Отображение $\mathcal{I}\colon\ \mathcal{W}\to\mathbf{R}^m$ зададим на множестве

$$\mathcal{W}=\{x(\cdot)\in C^1([t_0,\ t_1],\mathbf{R}^n)\,|\,(t,x(t),\dot{x}(t))\in U,\ t_0\leqslant t\leqslant t_1\} \quad (1)$$

равенством

$$\mathcal{I}(x(\cdot))=\int_{t_0}^{t_1} f(t,\ x(t),\ \dot{x}(t))\,dt. \quad (2)$$

Предложение 1. *Интегральное отображение* (2) *непрерывно дифференцируемо на множестве* (1) *и при этом*

$$\mathcal{I}'(\hat{x}(\cdot))[h(\cdot)]=\int_{t_0}^{t_1}\{\hat{f}_x(t)h(t)+\hat{f}_{\dot{x}}(t)\dot{h}(t)\}\,dt, \quad (3)$$

где

$$\hat{f}_x(t)\Leftrightarrow\left(\frac{\partial f_i}{\partial x_j}(t,\hat{x}(t),\dot{\hat{x}}(t)),\ i=1,\ldots,m,\ j=1,\ldots,n\right), \quad (4)$$

$$\hat{f}_{\dot{x}}(t)\Leftrightarrow\left(\frac{\partial f_i}{\partial\dot{x}_j}(t,\hat{x}(t),\dot{\hat{x}}(t)),\ i=1,\ldots,m,\ j=1,\ldots,n\right). \quad (5)$$

Доказательство. Представим отображение (2) в виде суперпозиции

$$\mathcal{I}=I\circ\mathcal{N}\circ D,$$

где

$$D\colon\ C([t_0,\ t_1],\mathbf{R}^m)\to\mathbf{R}^m$$

— линейный непрерывный оператор, определяемый равенством

$$D(x(\cdot)) = (x(\cdot),\ \dot{x}(\cdot));$$
$$\mathcal{N}\colon \mathcal{U} \to C([t_0,\ t_1],\ \mathbf{R}^m)$$

— оператор Немыцкого, определяемый равенством (9) п. 2.4.1 (при $r = n$), и

$$I\colon\ C([t_0,\ t_1],\ \mathbf{R}^m) \to \mathbf{R}^m$$

— оператор интегрирования

$$I(x(\cdot)) = \int_{t_0}^{t_1} x(t)\,dt$$

также линейный и непрерывный. Все эти операторы дифференцируемы (см. п. 2.2.1 и предложение 2 п. 2.4.1), причем производная линейного оператора совпадает с ним самим, а производная оператора Немыцкого дается формулой (10) п. 2.4.1. По теореме о суперпозиции

$$\mathcal{J}'(\hat{x}(\cdot)) = I \circ \mathcal{N}'(D\hat{x}(\cdot)) \circ D, \tag{6}$$

т. е.

$$\mathcal{J}'(\hat{x}(\cdot))[h(\cdot)] = I\{\mathcal{N}'(D\hat{x}(\cdot))[Dh(\cdot)]\} =$$
$$= I\{\mathcal{N}'(\hat{x}(\cdot),\ \dot{\hat{x}}(\cdot))[h(\cdot),\ \dot{h}(\cdot)] =$$
$$= I\{\hat{f}_x(t)h(t) + \hat{f}_{\dot{x}}(t)\dot{h}(t)\} =$$
$$= \int_{t_0}^{t_1} \{\hat{f}_x(t)h(t) + \hat{f}_{\dot{x}}(t)\dot{h}(t)\}\,dt.$$

Этим доказана формула (3). Непрерывность $\mathcal{J}'(x(\cdot))$ следует из непрерывности производной оператора Немыцкого и равенства (6). ■

Упражнение. Найдите нормы операторов D и I.

В задачах классического вариационного исчисления и оптимального управления рассматриваются также интегральные функционалы вида (2) с переменными пределами интегрирования t_0 и t_1. Чтобы включить их в общую схему, поступим следующим образом.

Пусть предположения относительно $f(t,\ x,\ \dot{x})$ те же, что и раньше, $\Delta \subseteq \mathbf{R}$ — некоторый отрезок,

$$\mathcal{U} = \{(x(\cdot),\ t_0,\ t_1)\,|\,x(\cdot) \in C^1(\Delta,\ \mathbf{R}^n),\ (t,\ x(t),\ \dot{x}(t)) \in U,\quad t \in \Delta,\ t_0,\ t_1 \in \operatorname{int}\Delta\}. \tag{7}$$

Определим отображение $\mathcal{I}\colon \mathcal{U} \to \mathbb{R}^m$ равенством

$$\mathcal{I}(x(\cdot),\ t_0,\ t_1) = \int_{t_0}^{t_1} f(t,\ x(t),\ \dot{x}(t))\,dt. \tag{8}$$

Предложение 2. *Интегральное отображение* (8) *непрерывно дифференцируемо на множестве* (7) *и при этом*

$$\mathcal{I}'(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)[h(\cdot),\ \tau_0,\ \tau_1] = \int_{t_0}^{t_1} (\hat{f}_x(t)\,h(t) + \hat{f}_{\dot{x}}(t)\,\dot{h}(t))\,dt + \hat{f}(\hat{t}_i)\Big|_{i=0}^{i=1}, \tag{9}$$

где $\hat{f}_x(t)$ *и* $\hat{f}_{\dot{x}}(t)$ *даются формулами* (4) *и* (5), *а*

$$\hat{f}(t) = f(t,\ \hat{x}(t),\ \dot{\hat{x}}(t)). \tag{10}$$

Доказательство. Воспользуемся теоремой о полном дифференциале из п. 2.2.4. Частные производные существуют в соответствии с предложением 1 и классической теоремой о производной интеграла по верхнему и нижнему пределам интегрирования:

$$\mathcal{I}_{x(\cdot)}(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)[h(\cdot)] = \int_{\hat{t}_0}^{\hat{t}_1} \{\hat{f}_x(t)\,h(t) + \hat{f}_{\dot{x}}(t)\,\dot{h}(t)\}\,dt,$$

$$\mathcal{I}_{t_0}(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)[\tau_0] = -\hat{f}(\hat{t}_0)\,\tau_0,$$

$$\mathcal{I}_{t_1}(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)[\tau_1] = \hat{f}(\hat{t}_1)\,\tau_1.$$

Приступаем к проверке непрерывности частных производных.

А) $\|\mathcal{I}_{t_0}(x(\cdot),\ t_0,\ t_1) - \mathcal{I}_{t_0}(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)\| =$

$$= \sup_{\|\tau_0\| \leqslant 1} \|f(t_0, x(t_0), \dot{x}(t_0))\,\tau_0 - f(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \dot{\hat{x}}(t_0))\,\tau_0\| \leqslant$$

$$\leqslant |f(t_0,\ x(t_0),\ \dot{x}(t_0)) - f(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \dot{\hat{x}}(\hat{t}_0))|. \tag{11}$$

Кроме того,

$$|x(t_0) - \hat{x}(\hat{t}_0)| \leqslant |x(t_0) - \hat{x}(t_0)| + |\hat{x}(t_0) - \hat{x}(\hat{t}_0)| \leqslant$$

$$\leqslant \|\hat{x}(\cdot) - x(\cdot)\|_{C^1} + |\hat{x}(t_0) - \hat{x}(\hat{t}_0)|, \tag{12}$$

$$|\dot{x}(t_0) - \dot{\hat{x}}(\hat{t}_0)| \leqslant |\dot{x}(t_0) - \dot{\hat{x}}(t_0)| + |\dot{\hat{x}}(t_0) - \dot{\hat{x}}(\hat{t}_0)| \leqslant$$

$$\leqslant \|\hat{x}(\cdot) - x(\cdot)\|_{C^1} + |\dot{\hat{x}}(t_0) - \dot{\hat{x}}(\hat{t}_0)|. \tag{13}$$

Поэтому, если $t_0 \to \hat{t}_0$ и $x(\cdot) \to \hat{x}(\cdot)$ в $C^1(\Delta, \mathbf{R}^n)$, то $x(t_0) \to \hat{x}(\hat{t}_0)$, $\dot{x}(t_0) \to \dot{\hat{x}}(\hat{t}_0)$ и, значит, правая часть неравенства (11) стремится к нулю. Следовательно, $\mathcal{I}_{t_0}(x(\cdot), t_0, t_1)$ непрерывно зависит от $(x(\cdot), t_0, t_1)$ (от t_1 эта производная не зависит вовсе). Непрерывность $\mathcal{I}_{t_1}(x(\cdot), t_0, t_1)$ проверяется аналогично.

Б) Выберем число $a > 0$ так, чтобы компакт

$$\mathcal{K} = \{(t, x, u) \mid |x - \hat{x}(t)| \leqslant a, \ |u - \dot{\hat{x}}(t)| \leqslant a, \ t \in \Delta\} \subset U.$$

На этом компакте производные $f_x(t, x, \dot{x})$ и $f_{\dot{x}}(t, x, \dot{x})$ равномерно непрерывны и ограничены. Теперь при $\|x(\cdot) - x(\cdot)\|_{C^1} \leqslant a$ имеем:

$$\|\mathcal{I}_{x(\cdot)}(x(\cdot), t_0, t_1) - \mathcal{I}_{x(\cdot)}(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)\| =$$

$$= \sup_{\|h(\cdot)\|_{C^1} \leqslant 1} \left| \int_{t_0}^{t_1} (f_x h + f_{\dot{x}} \dot{h})\, dt - \int_{\hat{t}_0}^{\hat{t}_1} (\hat{f}_x h + \hat{f}_{\dot{x}} \dot{h})\, dt \right| \leqslant$$

$$\leqslant \sup_{\|h(\cdot)\|_{C^1} \leqslant 1} \left\{ \left| \int_{\hat{t}_1}^{t_1} \{f_x h + f_{\dot{x}} \dot{h}\}\, dt \right| + \left| \int_{t_0}^{\hat{t}_0} \{f_x h + f_{\dot{x}} \dot{h}\}\, dt \right| + \right.$$

$$\left. + \left| \int_{\hat{t}_1}^{\hat{t}_1} \{(f_x - \hat{f}_x) h + (f_{\dot{x}} - \hat{f}_{\dot{x}}) \dot{h}\}\, dt \right| \right\} \leqslant$$

$$\leqslant \max_{\mathcal{K}} \{\|f_x(t, x, u)\| + \|f_{\dot{x}}(t, x, u)\|\} (|t_1 - \hat{t}_1| + |t_0 - \hat{t}_0|) +$$

$$+ \max_{[\hat{t}_0, \hat{t}_1]} \{\|f_x(t, x(t), \dot{x}(t)) - f_x(t, \hat{x}(t), \dot{\hat{x}}(t))\| +$$

$$+ \|f_{\dot{x}}(t, x(t), \dot{x}(t)) - f_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))\|\}.$$

Первый член стремится к нулю при $t_0 \to \hat{t}_0$ и $t_1 \to \hat{t}_1$. Второй оценивается с использованием равномерной непрерывности так же, как в доказательстве предложения 1 п. 2.4.1, и тем самым стремится к нулю, когда $x(\cdot) \to \hat{x}(\cdot)$ в пространстве $C^1(\Delta, \mathbf{R}^n)$.

Применяя теорему о полном дифференциале, получаем (9). ■

2.4.3. Оператор краевых условий. Пусть функция $\psi(t_0, x_0, t_1, x_1)\colon W \to \mathbf{R}^s$ непрерывно дифференцируема на открытом множестве $W \subset \mathbf{R} \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n$, и пусть

$$\mathscr{W} = \{(x(\cdot),\ t_0,\ t_1) \mid x(\cdot) \in C^1(\Delta,\ \mathbf{R}^n),$$
$$t_0,\ t_1 \in \operatorname{int} \Delta,\ (t_0,\ x(t_0),\ t_1,\ x(t_1)) \in W\}. \quad (1)$$

Отображение $\Psi\colon \mathscr{W} \to \mathbf{R}^s$, определяемое равенством

$$\Psi(x(\cdot),\ t_0,\ t_1) = \psi(t_0,\ x(t_0),\ t_1,\ x(t_1)), \quad (2)$$

называется *оператором краевых условий.*

Предложение. *Оператор краевых условий* (2) *непрерывно дифференцируем на множестве* (1) *и при этом*

$$\Psi'(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)[h(\cdot),\ \tau_0,\ \tau_1] =$$
$$= \hat{\psi}_{t_0}\tau_0 + \hat{\psi}_{x_0}[h(\hat{t}_0) + \dot{\hat{x}}(\hat{t}_0)\tau_0] + \hat{\psi}_{t_1}\tau_1 + \hat{\psi}_{x_1}[h(\hat{t}_1) + \dot{\hat{x}}(t_1)\hat{\tau}_1], \quad (3)$$

где

$$\hat{\psi}_{t_i} = \psi_{t_i}(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \hat{t}_1,\ \hat{x}(\hat{t}_1)), \quad i = 0,\ 1, \quad (4)$$
$$\hat{\psi}_{x_i} = \psi_{x_i}(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ t_1,\ \hat{x}(\hat{t}_1)), \quad i = 0,\ 1. \quad (5)$$

Доказательство. А) Рассмотрим сначала простейшее отображение $\mathrm{ev}\colon C^1(\Delta,\ \mathbf{R}^n) \times \operatorname{int} \Delta \to \mathbf{R}^n$, определяемое равенством

$$\mathrm{ev}(x(\cdot),\ t_0) = x(t_0)$$

(отображение значений). По $x(\cdot)$ это отображение линейно, и потому

$$\mathrm{ev}_{x(\cdot)}(\hat{x}(\cdot),\ \hat{t}_0)[h(\cdot)] = h(\hat{t}_0).$$

Частная производная по t_0 — это обычная производная

$$\mathrm{ev}_{t_0}(\hat{x}(\cdot),\ \hat{t}_0)[\tau_0] = \dot{\hat{x}}(\hat{t}_0)\tau_0.$$

Проверяем непрерывность:

$$\|\mathrm{ev}_{x(\cdot)}(x(\cdot),\ t_0) - \mathrm{ev}_{x(\cdot)}(\hat{x}(\cdot),\ \hat{t}_0)\| =$$
$$= \sup_{\|h(\cdot)\|_{C^1} \leqslant 1} \|h(t_0) - h(\hat{t}_0)\| \leqslant$$
$$\leqslant \sup_{\|h(\cdot)\|_{C^1} \leqslant 1}\ \sup_{\theta \in [t_0,\ \hat{t}_0]} \|\dot{h}(\theta)\|\,|t_0 - \hat{t}_0| \leqslant |t_0 - \hat{t}_0| \to 0$$

при $t_0 \to \hat{t}_0$. Далее

$$\|\mathrm{ev}_{t_0}(x(\cdot),\ t_0) - \mathrm{ev}_{t_0}(\hat{x}(\cdot),\ t_0)\| =$$
$$= \sup_{|\tau_0| \leqslant 1} \|\dot{x}(t_0)\tau_0 - \dot{\hat{x}}(\hat{t}_0)\tau_0\| = \|\dot{\hat{x}}(t_0) - \dot{\hat{x}}(\hat{t}_0)\| \to 0,$$

когда $t_0 \to \hat{t}_0$ и $x(\cdot) \to \hat{x}(\cdot)$ в пространстве $C^1(\Delta, \mathbf{R}^n)$, как это было показано при доказательстве предложения 2 п. 2.4.2 (неравенства (12) и (13)). В силу теоремы о полном дифференциале (п. 2.2.4)

$$\operatorname{ev}(\hat{x}(\cdot), \hat{t}_0)'[h(\cdot), \tau_0] = h(\hat{t}_0) + \dot{\hat{x}}(\hat{t})\tau_0.$$

Аналогично убеждаемся в непрерывной дифференцируемости отображения $(x(\cdot), t_1) \to x(t_1)$.

Б) Воспользовавшись теоремой о суперпозиции, убеждаемся в дифференцируемости отображения (2) и справедливости равенства (3). ■

Упражнение 1. Пусть отображение $\operatorname{ev}: C^1([0, 1]) \times (0, 1) \to \mathbf{R}$ определено формулой $\operatorname{ev}(x(\cdot), t_0) = x(t_0)$ (ср. п. 2.4.3). Докажите, что:

а) для существования второй вариации $\delta^2 \operatorname{ev}(x(\cdot), t_0)$ необходимо и достаточно, чтобы существовала $\ddot{x}(t_0)$ и при этом

$$\delta^2 \operatorname{ev}(x(\cdot), t_0)[h(\cdot), \tau] = 2\dot{h}(t_0)\tau + \ddot{x}(t_0)\tau^2;$$

б) отображение ev не имеет второй производной Фреше, хотя его первая производная Фреше дифференцируема по Гато.

Указание. $\dot{h}(t_0+\tau) - \dot{h}(t_0) \neq o(\|h\|^2 + \tau^2)$.

§ 2.5. Необходимые сведения из теории обыкновенных дифференциальных уравнений

Как уже отмечалось в п. 2.1.8, дифференциальные уравнения $\dot{x} = \varphi(t, x, u(t))$, рассматриваемые в задачах оптимального управления, имеют свою специфику. Поскольку мы допускаем разрывные управления $u(\cdot)$, их правая часть не обязана удовлетворять условиям стандартных теорем из курса дифференциальных уравнений. К тому же нужные сведения о решениях не всегда излагаются в этих курсах в удобной для нас форме. Поэтому для полноты изложения, а также чтобы проиллюстрировать возможность применения здесь общих теорем §§ 2.3 и 2.4, мы приводим в этом параграфе доказательства основных теорем: существования, единственности и дифференцируемости решений, а также некоторые специальные утверждения, относящиеся к линейным системам. Часть формулировок дана в несколько более общей форме, чем это обычно необходимо. Чаще всего мы имеем дело только с кусочно-непрерывными управлениями $u(\cdot)$, но и в этом случае удобно говорить об «измеримости», «интегрируемости» и т. п.

Функция $x(\cdot)$ называется *решением дифференциального уравнения*

$$\dot{x} = F(t, x),$$

если она абсолютно непрерывна (см. п. 2.1.8) и удовлетворяет уравнению почти всюду.

В эквивалентной форме $x(\cdot)$ должно быть решением интегрального уравнения

$$x(t) = x(t_0) + \int_{t_0}^{t} F(s, x(s))\, ds.$$

2.5.1. Основные предположения. Здесь и далее будем предполагать, что G — открытое множество в $\mathbf{R}\times\mathbf{R}^n$ и что функция $F\colon G \to \mathbf{R}^n$ удовлетворяет следующим трем условиям:

А) Для любого x функция $t \mapsto F(t, x)$, определенная на сечении $G_x = \{t \mid (t, x) \in G\}$, измерима и интегрируема на любом конечном отрезке, содержащемся в G_x.

Б) Для любого t функция $x \mapsto F(t, x)$, определенная на сечении $G_t = \{x \mid (t, x) \in G\}$, дифференцируема (хотя бы в смысле Гато).

В) Для любого компакта $\mathcal{K} \subset G$ существует такая локально интегрируемая функция $k(\cdot)$, что

$$\|F_x(t, x)\| \leqslant k(t), \quad \forall (t, x) \in \mathcal{K}. \tag{1}$$

(Функция называется *локально интегрируемой*, если она интегрируема на любом конечном отрезке.)

Типичный пример доставляет функция $F(t, x) = \varphi(t, x, u(t))$, где $\varphi(t, x, u)$ и $\varphi_x(t, x, u)$ непрерывны на $G\times\mathcal{U}$, а $u\colon \mathbf{R} \to \mathcal{U}$ кусочно-непрерывна. Условия А) и Б) здесь очевидны, а $k(t)$ можно положить равной максимуму $\|F_x(t, x)\|$ на $\mathcal{K}$. Другой пример — $F(t, x) = A(t)x$, где $A(\cdot)$ измеримая локально интегрируемая матричная функция,

$$A\colon \mathbf{R} \to \mathscr{L}(\mathbf{R}^n, \mathbf{R}^n).$$

Замечание. Условия А)—В) более ограничительны, чем известные условия Каратеодори [6], но они лучше приспособлены, с одной стороны, к нашим потребностям (рассмотрение уравнений $\dot{x} = \varphi(t, x, u(t))$ с разрывными управлениями $u(\cdot)$, но с существованием производной φ_x), и к нашим возможностям (аппарат дифференциального исчисления § 2.2), с другой.

Лемма 1. *Для любого компакта $\mathscr{K}\subset G$ существует локально интегрируемая функция $\varkappa(\cdot)$ такая, что*

$$|F(t,\ x)|\leqslant\varkappa(t),\quad \forall(t,\ x)\in\mathscr{K}. \tag{2}$$

Доказательство. Поскольку $\mathscr{K}$ — компакт, существуют такие $\delta>0$ и $\varepsilon>0$, что для любой точки $(t_0, x_0)\in\mathscr{K}$ «цилиндр»

$$C_{t_0x_0}=\{(t,\ x)\,|\,|t-t_0|\leqslant\delta,\ |x-x_0|\leqslant\varepsilon\}\subset G.$$

Для этого цилиндра, согласно В), найдем локально интегрируемую функцию $k_{t_0x_0}(t)$.

Снова используя компактность, покроем $\mathscr{K}$ конечным числом цилиндров: $\mathscr{K}\subset\bigcup\limits_{i=1}^{N}C_{t_ix_i}$.

Функция

$$\varkappa(t)=\sum_{i=1}^{N}\left[|F(t,\ x_i)|+k_{t_ix_i}(t)\cdot\varepsilon\right]\chi_{[t_i-\delta,\ t_i+\delta]}(t)$$

— искомая ($\chi_{[\alpha,\ \beta]}(\cdot)$ — характеристическая функция отрезка $[\alpha,\ \beta]$; $|F(t,\ x_i)|$ интегрируема на $[t_i-\delta,\ t_i+\delta]$ в силу А)). Действительно,

$$(t,\ x)\in\mathscr{K}\Rightarrow\exists i,\ (t,\ x)\in C_{t_ix_i}\Rightarrow$$
$$\Rightarrow|F(t,\ x)|\leqslant|F(t,\ x_i)|+|F(t,\ x)-F(t,\ x_i)|\leqslant$$
$$\leqslant|F(t,\ x)|+\sup_{c\in[x,\ x_i]}\|F_x(t,\ c)\|\varepsilon\leqslant\varkappa(t)$$

(отрезок $[(t,\ x),\ (t,\ x_i)]\subset C_{t_ix_i}$ и применима теорема о среднем п. 2.2.3)).

Лемма 2. *Если функция $x\colon\Delta\to\mathbf{R}^n$ непрерывна на отрезке Δ и ее график $\{(t,\ x(t))\,|\,t\in\Delta\}$ лежит в открытом множестве $G\subset\mathbf{R}\times\mathbf{R}^n$, а функция $F\colon G\to\mathbf{R}^n$ удовлетворяет условиям* А)—В), *то функция $t\mapsto f(t)=F(t,\ x(t))$ измерима и интегрируема на Δ.*

Доказательство. При достаточно малом $\varepsilon>0$

$$\mathscr{K}=\{(t,\ x)\,|\,|x-x(t)|\leqslant\varepsilon,\ t\in\Delta\}\subset G,$$

и пусть $k(\cdot)$ — функция, отвечающая этому компакту в силу условия В). Непрерывная функция $x(\cdot)$ является равномерным пределом кусочно-постоянных функций $x_n(\cdot)$ и без ограничения общности графики $x_n(\cdot)$ лежат в $\mathscr{K}$.

На каждом интервале постоянства $x_n(\cdot)$ функция $t\mapsto f_n(t)=F(t,\ x_n(t))$ измерима и интегрируема по условию А), а потому $f_n(\cdot)$ измерима и интегрируема на Δ.

Используя теорему о среднем (п. 2.2.3) и (1), имеем

$$|f_n(t)-f(t)|=|F(t,\ x_n(t))-F(t,\ x(t))|\leqslant$$
$$\leqslant k(t)\,|x_n(t)-x(t)|,$$

откуда $f_n(t)\longrightarrow f(t)$ при всех $t\in\Delta$, а значит, $f(\cdot)$ измерима как предел измеримых функций [КФ, стр. 284]. Наконец,

$$|f(t)|\leqslant|f_n(t)|+k(t)\,|x_n(t)-x(t)|$$

и, следовательно, $f(t)$ интегрируема.

2.5.2. Локальная теорема существования. *Пусть на открытом множестве* $G\subset\mathbf{R}\times\mathbf{R}^n$ *функция* F: $G\longrightarrow\mathbf{R}^n$ *удовлетворяет условиям* A)—B), *и пусть компакт* $\mathcal{K}\subset G$.

Тогда существуют такие $\delta>0$ *и* $\varepsilon>0$, *что для любой точки* $(\hat{t},\ \hat{x})\in\mathcal{K}$ *и для* $(t_0,\ x_0)$, *удовлетворяющих неравенствам*

$$|t_0-\hat{t}|<\delta,\quad |x_0-\hat{x}|<\varepsilon, \tag{1}$$

решение $X(t,\ t_0,\ x_0)$ *задачи Коши*

$$\begin{cases} \dot{x}=F(t,\ x), \\ x(t_0)=x_0, \end{cases} \tag{2}$$

определено на отрезке $[\hat{t}-\delta,\ \hat{t}+\delta]$ *и является непрерывной функцией по совокупности аргументов.*

Доказательство. Применим к рассматриваемой ситуации принцип сжимающих отображений в формулировке п. 2.3.2. Задача Коши (2) эквивалентна интегральному уравнению

$$x(t)=x_0+\int_{t_0}^{t} F(s,\ x(s))\,ds, \tag{3}$$

и мы применим лемму п. 2.3.2 к отображению

$$(t_0,\ x_0,\ x(\cdot))\longmapsto\Phi(t_0,\ x_0,\ x(\cdot))=x_0+\int_{t_0}^{(\cdot)} F(s,\ x(s))\,ds. \tag{4}$$

Выберем γ и β так, чтобы

$$\mathcal{K}_1=\{(t,\ x)\,|\,|t-\hat{t}|\leqslant\gamma,\ |x-\hat{x}|\leqslant\beta,\ (\hat{t},\ \hat{x})\in\mathcal{K}\}\subset G,$$

и пусть $k(\cdot)$ и $\varkappa(\cdot)$ — интегрируемые функции, отвечающие в силу B) и леммы 1 п. 2.5.1 этому компакту. Займемся проверкой условий леммы п. 2.3.2. Топологическим пространством T здесь будет множество (1), $U=T$, $Y=C([\hat{t}-\delta,\ \hat{t}+\delta],\ \mathbf{R}^n)$, где δ, $0<\delta\leqslant\gamma$, мы подберем позже; $y_0(s)\equiv\hat{x}$,

$$V=\{(t_0,\ x_0,\ x(\cdot))\mid |t_0-\hat{t}|<\delta,\ |x_0-\hat{x}|<\varepsilon,\ \|x(\cdot)-y_0(\cdot)\|_C<\beta\}. \tag{5}$$

Константы ε и θ таковы, что $0<\theta<1$ и $0<\varepsilon<\beta(1-\theta)$.

Условие в) п. 2.3.2 выполняется, если для $t\in[\hat{t}-\delta,\ \hat{t}+\delta]$

$$\left|x_0+\int_{t_0}^{t}F(s,\ \hat{x})\,ds-\hat{x}\right|<\beta(1-\theta). \tag{6}$$

По лемме 1 п. 2.5.1 $|F(s,\ \hat{x})|\leqslant\varkappa(s)$, и, выбрав δ достаточно малым, будем иметь

$$\int_{\hat{t}-\delta}^{\hat{t}+\delta}|F(s,\ \hat{x})|\,ds\leqslant\int_{\hat{t}-\delta}^{\hat{t}+\delta}\varkappa(s)\,ds<\beta(1-\theta)-\varepsilon, \tag{7}$$

откуда следует (6), поскольку

$$\left|x_0+\int_{t_0}^{t}F(s,\ \hat{x})\,ds-\hat{x}\right|\leqslant|x_0-\hat{x}|+\left|\int_{t_0}^{t}F(s,\ \hat{x})\,ds\right|\leqslant$$
$$\leqslant\varepsilon+\int_{\hat{t}-\delta}^{\hat{t}+\delta}|F(s,\ \hat{x})|\,ds<\beta(1-\theta).$$

Условие а) п. 2.3.2 означает, что отображение $(t_0,\ x_0,\ x(\cdot))\mapsto(t_0,\ x_0,\ \Phi(t_0,\ x_0,\ x(\cdot)))$ переводит V в себя. Оно выполняется, если

$$\left|x_0+\int_{t_0}^{t}F(s,\ x(s))\,ds-\hat{x}\right|<\beta$$

для $(t_0,\ x_0,\ x(\cdot))\in V$ и $t\in[\hat{t}-\delta,\ \hat{t}+\delta]$. Оценим здесь левую часть, используя (6), теорему о среднем и (1) п. 2.5.1:

$$\left|x_0+\int_{t_0}^{t} F(s,\ x(s))\,ds-\hat{x}\right| \leqslant$$

$$\leqslant\left|x_0+\int_{t_0}^{t} F(s,\ \hat{x})\,ds-\hat{x}\right|+\left|\int_{t_0}^{t}\{F(s,\ x(s))-F(s,\ \hat{x})\}\,ds\right|<$$

$$<\beta(1-\theta)+\int_{\hat{t}-\delta}^{\hat{t}+\delta} k(s)\,|x(s)-\hat{x}|\,ds\leqslant$$

$$\leqslant\beta\Bigl(1-\theta+\int_{\hat{t}-\delta}^{\hat{t}+\delta} k(s)\,ds\Bigr)\leqslant\beta,$$

если выполняется неравенство

$$\int_{\hat{t}-\delta}^{\hat{t}+\delta} k(s)\,ds\leqslant\theta, \tag{8}$$

чего мы можем добиться, уменьшив, если нужно, δ.

Наконец, поскольку уже проверено условие а), условие б) п. 2.3.2 выполняется, если

$$\|\Phi(t_0,\ x_0,\ x(\cdot))-\Phi(t_0,\ x_0,\ y(\cdot))\|\leqslant\theta\|x(\cdot)-y(\cdot)\|$$

для любых $(t_0,\ x_0,\ x(\cdot))\in V$ и $(t_0,\ x_0,\ y(\cdot))\in V$. Но

$$\|\Phi(t_0,\ x_0,\ x(\cdot))-\Phi(t_0,\ x_0,\ y(\cdot))\|=$$

$$=\max_t\left|x_0+\int_{t_0}^{t} F(s,\ x(s))\,ds-x_0-\int_{t_0}^{t} F(s,\ y(s))\,ds\right|=$$

$$=\max_t\left|\int_{t_0}^{t}\{F(s,\ x(s))-F(s,\ y(s))\}\,ds\right|\leqslant$$

$$\leqslant\int_{\hat{t}-\delta}^{\hat{t}+\delta} k(s)\,ds\,\|x(\cdot)-y(\cdot)\|\leqslant\theta\|x(\cdot)-y(\cdot)\|$$

в силу (8).

Таким образом, лемма п. 2.3.2 применима, и потому последовательность $X_n(\cdot\ ,\ t_0,\ x_0)$, определяемая равенствами $X_0(\cdot\ ,\ t_0,\ x_0)\equiv\hat{x}$ и

$$X_{n+1}(\cdot\ ,\ t_0,\ x_0)=x_0+\int_{t_0}^{(\ )} F(s,\ X_n(s,\ t_0,\ x_0))\,ds, \tag{9}$$

равномерно по (t_0, x_0) из (1) сходится в пространстве $C[(\hat{t}-\delta, \hat{t}+\delta], \mathbf{R}^n)$, т. е. равномерно по $t\in[\hat{t}-\delta, \hat{t}+\delta]$. Переходя к пределу в (9), получаем, что $X(t, t_0, x_0)=\lim\limits_{n\to\infty} X_n(t, t_0, x_0)$ удовлетворяет интегральному уравнению (3) и, следовательно, является решением задачи Коши (2). По индукции легко проверяется, что $X(t, t_0, x_0)$ непрерывны, а поскольку сходимость равномерна, то и $X(t, t_0, x_0)$ непрерывна по совокупности аргументов. ■

2.5.3. Теорема единственности.

Лемма (неравенство Гронуолла). *Пусть неотрицательные функции $\alpha(\cdot)$ и $\omega(\cdot)$ измеримы на отрезке Δ, причем $\alpha(\cdot)\omega(\cdot)$ на Δ интегрируема. Если для некоторых $b>0$ и $\tau\in\Delta$ и для всех $t\in\Delta$ выполняется неравенство*

$$\omega(t)\leqslant\left|\int_\tau^t \alpha(s)\,\omega(s)\,ds\right|+b, \tag{1}$$

то для всех $t\in\Delta$

$$\omega(t)\leqslant be^{\left|\int_\tau^t \alpha(s)\,ds\right|}. \tag{2}$$

Доказательство. Пусть сначала $t\geqslant\tau$. По условию

$$N=\int_\Delta \alpha(s)\,\omega(s)\,ds<\infty$$

и, согласно (1),

$$\omega(t)\leqslant N+b.$$

По индукции находим, что из (1) вытекают неравенства

$$\omega(t)\leqslant b+b\int_\tau^t\alpha(s)\,ds+\ldots$$
$$\ldots+\frac{b}{(m-1)!}\left[\int_\tau^t\alpha(s)\,ds\right]^{m-1}+\frac{b+N}{m!}\left[\int_\tau^t\alpha(s)\,ds\right]^m. \tag{3}$$

Действительно, при $m=0$ (3) верно, а если (3) верно при некотором m, то, подставив эту оценку в (1) и воспользовавшись равенством

$$\int_\tau \alpha(s)\left[\int_\tau^s \alpha(\sigma)\,d\sigma\right]^k ds = \left.\frac{\left[\int_\tau^s \alpha(\sigma)\,d\sigma\right]^{k+1}}{k+1}\right|_{s=\tau}^{s=t} = \frac{\left[\int_\tau^t \alpha(\sigma)\,d\sigma\right]^{k+1}}{k+1},$$

убеждаемся, что (3) верно для $m+1$. Переходя в (3) к пределу при $m\to\infty$, получаем (1).

При $t\leqslant\tau$ рассуждаем аналогично, надо лишь везде поменять местами пределы интегрирования. (Заметим, что функция $\alpha(\cdot)$ не обязательно интегрируема, так что правая часть в (2) может оказаться бесконечной.) ■

Теорема единственности. *Пусть функция* F: $G\to\mathbf{R}^n$ *удовлетворяет условиям* А)—В) п. 2.5.1 *в открытом множестве* $G\subset\mathbf{R}\times\mathbf{R}^n$; Δ_i, $i=1, 2$, — *интервалы в* $\mathbf{R}$ *и* $x_i(\cdot)$: $\Delta_i\to\mathbf{R}^n$ — *два решения задачи Коши* (2) п. 2.5.2, *графики которых содержатся в* G. *Тогда* $x_1(t)\equiv x_2(t)$ *для всех* $t\in\Delta_1\cap\Delta_2$.

Доказательство. Множество $\Delta=\{t\mid x_1(t)=x_2(t)\}$, очевидно, замкнуто в $\Delta_1\cap\Delta_2$ и непусто, поскольку $t_0\in\Delta$. Остается проверить, что оно открыто, и тогда утверждение теоремы будет следствием связности интервала $\Delta_1\cap\Delta_2$.

Пусть $\hat t\in\Delta$, так что $x_1(\hat t)=x_2(\hat t)=\hat x$. Выберем такие $\gamma>0$ и $\beta>0$, чтобы $\mathscr{K}=\{(t, x)\mid |t-\hat t|\leqslant\gamma, |x-\hat x|\leqslant\beta\}\subset G$ и чтобы $(t, x_i(t))\in\mathscr{K}$ при $|t-\hat t|\leqslant\gamma$; функция $k(t)$ отвечает $\mathscr{K}$ в силу условия В) п. 2.5.1.

Применяя теорему о среднем (п. 2.2.3) и неравенство (1) п. 2.5.1, находим, что

$$|x_1(t)-x_2(t)|\leqslant\left|\int_{\hat t}^t [F(s, x_1(s))-F(s, x_2(s))]\,ds\right|\leqslant \leqslant\left|\int_{\hat t}^t k(s)\,|x_1(s)-x_2(s)|\,ds\right|.$$

Если теперь к функции $\omega(t)\doteq|x_1(t)-x_2(t)|$ применить на $[\hat{t}-\gamma, \hat{t}+\gamma]$ только что доказанную лемму ($\alpha(t)=k(t)$, $b=0$), то, согласно (2), $x_1(t)\equiv x_2(t)$.

Следовательно, $(\hat{t}-\gamma, \hat{t}+\gamma)\subset\Delta$ и Δ открыто. ■

2.5.4. Линейные дифференциальные уравнения. В этом пункте Δ — отрезок числовой прямой,

$$A\colon \Delta\longrightarrow\mathscr{L}(\mathbf{R}^n, \mathbf{R}^n),\ b\colon \Delta\longrightarrow\mathbf{R}^n \text{ и } c\colon \Delta\longrightarrow\mathbf{R}^{n*}$$

— интегрируемые матричная функция и две вектор-функции; $x=(x_1, \ldots, x_n)\in\mathbf{R}^n$, $p=(p_1, \ldots, p_n)\in\mathbf{R}^{n*}$. Мы будем рассматривать линейные системы дифференциальных уравнений

$$\dot{x}=A(t)x+b(t) \tag{1}$$

и

$$\dot{p}=-pA(t)+c(t) \tag{2}$$

(иногда систему (2) называют *сопряженной* системе (1)). Хорошо известно, что для линейных систем существование решений удается доказать нелокально, в нашем случае — на всем отрезке Δ.

Лемма. *Если функции $A\colon \Delta\longrightarrow\mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$, $b\colon \Delta\longrightarrow\mathbf{R}^n$ и $c\colon \Delta\longrightarrow\mathbf{R}^{n*}$ измеримы и интегрируемы на отрезке Δ, то любая задача Коши для уравнений* (1) *и* (2), *поставленная в момент $t_0\in\Delta$, имеет единственное решение, и это решение продолжается на весь отрезок Δ.*

Доказательство. Поскольку все рассуждения аналогичны, ограничимся уравнением (1). Для удобства продолжим функции $A(\cdot)$, $b(\cdot)$ и $c(\cdot)$ на все $\mathbf{R}$, положив их равными нулю вне Δ. Тогда функция

$$F(t, x)=A(t)x+b(t)$$

будет удовлетворять условиям А)—В) п. 2.5.1 в области $G=\mathbf{R}\times\mathbf{R}^n$, и потому к дифференциальному уравнению (1) применима теорема единственности п. 2.5.3 и локальная теорема существования п. 2.5.2 в каждой точке $(\hat{t}, \hat{x})$, так что решение уравнения (1), у которого $x(\hat{t})=\hat{x}$, заведомо единственно и определено на некотором отрезке $[\hat{t}-\delta, \hat{t}+\delta]$.

Заметим теперь, что если $|\hat{x}|\leqslant M$, то

$$|F(t, \hat{x})|\leqslant\|A(t)\|M+|b(t)|,$$
$$\|F_x(t, \hat{x})\|=\|A(t)\|.$$

Вернувшись к доказательству теоремы п. 2.5.2, мы можем выбрать достаточно произвольно константы γ, β, θ и ε, например, положив $\theta=1/2$, $\beta=4$, $\varepsilon=\gamma=1$, и тогда неравенства (7) и (8) п. 2.5.2, определяющие δ, приобретут вид

$$M\int_{\hat{t}-\delta}^{\hat{t}+\delta}\|A(s)\|ds+\int_{\hat{t}-\delta}^{\hat{t}+\delta}|b(s)|ds<1,$$
$$\int_{\hat{t}-\delta}^{\hat{t}+\delta}\|A(s)\|ds\leqslant 1/2. \tag{3}$$

Поскольку $\|A(\cdot)\|$ и $|b(\cdot)|$ интегрируемы, то мы можем выбрать $\delta>0$ столь малым (и не превосходящим $\gamma=1$), чтобы неравенства (3) выполнялись при всех $\hat{t}$ [КФ, стр. 301]. Следовательно, решение уравнения (1), у которого $x(\hat{t})=\hat{x}$, определено на отрезке $[\hat{t}-\delta,\ \hat{t}+\delta]$ с гарантированным δ, одним и тем же для любого $\hat{t}$ и любого $\hat{x}$, такого, что $|\hat{x}|\leqslant M$.

Пусть теперь заданы $t_0\in\Delta$ и x_0 и ищется решение уравнения (1), у которого

$$x(t_0)=x_0. \tag{4}$$

Положим

$$M=\left(|x_0|+\int_\Delta|b(s)|ds\right)e^{\int_\Delta\|A(s)\|ds} \tag{5}$$

Задача Коши (1), (4) эквивалентна интегральному уравнению

$$x(t)=x_0+\int_{t_0}^{t}A(s)\,x(s)\,ds+\int_{t_0}^{t}b(s)\,ds,$$

откуда

$$|x(t)|\leqslant|x_0|+\left|\int_{t_0}^{t}A(s)\,x(s)\,ds\right|+\left|\int_{t_0}^{t}b(s)\,ds\right|\leqslant$$
$$\leqslant\left(|x_0|+\int_\Delta|b(s)|ds\right)+\left|\int_{t_0}^{t}\|A(s)\|\,|x(s)|\,ds\right|. \tag{6}$$

Решение $x(\cdot)$ задачи (1), (4) заведомо определено на отрезке $\Delta_1=[t_0-\delta,\ t_0+\delta]$. Применяя на этом отрезке лемму п. 2.5.3 к функции $\omega(t)=|x(t)|$, $\left(\alpha(t)=\|A(t)\|,\right.$ $\left.b=|x_0|+\int_\Delta |b(s)|\,ds\right)$, получаем неравенство

$$|x(t)|\leqslant\left(|x_0|+\int_\Delta |b(s)|\,ds\right)e^{\left|\int_{t_0}^{t}\|A(s)\|ds\right|} \tag{7}$$

и, в частности, $|x(t_0\pm\delta)|\leqslant M$ согласно (5). Поэтому в точках $(t_0\pm\delta,\ x(t_0\pm\delta))$ снова можно воспользоваться локальной теоремой существования и продолжить решение $x(\cdot)$ на отрезок $\Delta_2=[t_0-2\delta,\ t_0+2\delta]$.

Дальше процедура повторяется. Неравенство (6) справедливо на Δ_2, а значит, на Δ_2 верно (7) и $|x(t_0\pm2\delta)|\leqslant M$ и т. д. За конечное число шагов мы продолжим решение задачи Коши (1), (4) на весь отрезок Δ. ■

Явные формулы для решений систем (1) и (2) выражаются через фундаментальную матрицу решений однородной системы

$$\dot{x}=A(t)\,x. \tag{8}$$

Определение. *Фундаментальной матрицей* $\Omega(t,\tau)$ *решений системы* (8) называется матричная функция Ω: $\Delta\times\Delta\to\mathscr{L}(\mathbf{R}^n,\mathbf{R}^n)$, являющаяся решением задачи Коши

$$\frac{\partial\Omega(t,\tau)}{\partial t}=A(t)\,\Omega(t,\tau), \tag{9}$$

$$\Omega(\tau,\tau)=E. \tag{10}$$

Другими словами, каждый столбец матрицы $\Omega(t,\tau)$ является решением системы (8), а при $t=\tau$ эти n столбцов обращаются в набор единичных векторов стандартного базиса в $\mathbf{R}^n$ $e_1=(1,0,\ldots,0)$, $e_2=(0,1,\ldots,0)$, … …, $e_n=(0,0,\ldots,1)$.

Теорема. *Если матричная функция* $A(\cdot)$ *интегрируема на отрезке* Δ, *то фундаментальная матрица* $\Omega(t,\tau)$ *системы* (8) *существует и непрерывна на квадрате* $\Delta\times\Delta$, *причем:*

1) $\Omega(t,s)\,\Omega(s,\tau)\equiv\Omega(t,\tau)$ *для всех* $t,s,\tau\in\Delta$; (11)

2) $\Omega(t,\tau)$ *при каждом* t *является решением дифференциального уравнения*

$$\frac{\partial\Omega(t,\tau)}{\partial\tau}=-\Omega(t,\tau)\,A(\tau). \tag{12}$$

Кроме того:

3) *Если функция* $b\colon \Delta \to \mathbf{R}^n$ *интегрируема на отрезке* Δ *и* $x(\cdot)$ — *решение системы* (1), *то для любых* $t, \tau \in \Delta$

$$x(t) = \Omega(t, \tau)\, x(\tau) + \int_{\tau}^{t} \Omega(t, s)\, b(s)\, ds. \tag{13}$$

4) *Если функция* $c\colon \Delta \to \mathbf{R}^{n*}$ *интегрируема на отрезке* Δ *и* $p(\cdot)$ — *решение системы* (2), *то для любых* $t, \tau \in \Delta$

$$p(t) = p(\tau)\, \Omega(\tau, t) - \int_{t}^{\tau} c(s)\, \Omega(s, t)\, ds. \tag{14}$$

Доказательство. Непосредственной подстановкой с использованием (9) проверяется, что для любого $\xi \in \mathbf{R}^n$ функции

$$x_1(t) = \Omega(t, s)\, \Omega(s, \tau)\, \xi \quad \text{и} \quad x_2(t) = \Omega(t, \tau)\, \xi$$

являются решениями системы (8), а так как в силу (10)

$$x_1(s) = \Omega(s, s)\, \Omega(s, \tau)\, \xi = \Omega(s, \tau)\, \xi = x_2(s),$$

то по теореме единственности

$$\Omega(t, s)\, \Omega(s, \tau)\, \xi = x_1(t) \equiv x_2(t) = \Omega(t, \tau)\, \xi.$$

Поскольку ξ любое, должно выполняться (11). Фиксируя в (11) s, мы представляем функцию пары переменных $\Omega(t, \tau)$ в виде произведения двух функций одной переменной, а так как обе они непрерывны, то Ω непрерывна по (t, τ) на $\Delta \times \Delta$. В частности, она ограничена. Полагая в (11) $t = \tau$, получаем $\Omega(t, s)\, \Omega(s, t) = E$, откуда

$$\Omega(t, s) \equiv [\Omega(s, t)]^{-1}.$$

Согласно предложению 3 п. 2.2.1 матричная функция $f(A) = A^{-1}$ непрерывно дифференцируема на множестве обратимых матриц и

$$f'(A)[H] = -A^{-1} H A^{-1}. \tag{15}$$

Множество $\mathcal{K} = \{A \mid A = \Omega(s, t),\ s, t \in \Delta\}$ является образом компакта $\Delta \times \Delta$ при непрерывном отображении $\Omega\colon \Delta \times \Delta \to \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$, а потому это компакт и на нем функция f ограничена: $\|f(A)\| \leqslant M$ — и удовлетворяет условию Липшица, так как

$$\|f(A) - f(B)\| = \|B^{-1} - A^{-1}\| = \|-B^{-1}(B - A)\, A^{-1}\| \leqslant M^2 \|B - A\|.$$

Согласно предложению 1 п. 2.1.8 функция

$$\varphi(s)=\Omega(t,\ s)=\Omega(s,\ t)^{-1}=f(\Omega(s,\ t))$$

абсолютно непрерывна, и в силу (15) и теоремы о суперпозиции п. 2.2.2 почти всюду

$$\frac{\partial\Omega(t,\ s)}{\partial s}=-[\Omega(s,\ t)]^{-1}\frac{\partial\Omega(s,\ t)}{\partial s}[\Omega(s,\ t)]^{-1}=$$

$$=-[\Omega(s,\ t)]^{-1}A(s)\Omega(s,\ t)[\Omega(s,\ t)]^{-1}=-\Omega(t,\ s)A(s),$$

чем доказано (12).

Представив правую часть (13) в виде

$$\Omega(t,\ \tau)\left[x(\tau)+\int_{\tau}^{t}\Omega(\tau,\ s)b(s)\,ds\right], \tag{16}$$

мы убеждаемся в ее абсолютной непрерывности. Действительно, $\Omega(\tau,\ s)b(s)$ интегрируемо (как функция s) на Δ, а интеграл от интегрируемой функции абсолютно непрерывен (п. 2.1.8). Следовательно, квадратная скобка в (16) абсолютно непрерывна по t. Остается заметить, что произведение двух абсолютно непрерывных функций также абсолютно непрерывно (предложение 1 п. 2.1.8).

Равенство (14) доказывается аналогично. ■

2.5.5. Глобальная теорема о существовании и непрерывной зависимости решения от начальных данных и параметров. Вернемся к рассмотрению задачи Коши:

$$\dot{x}=F(t,\ x), \tag{1}$$

$$x(t_0)=x_0 \tag{2}$$

для дифференциальных уравнений, правая часть которых удовлетворяет условиям А)—В) п. 2.5.1. Используя локальную теорему существования и теорему единственности, мы можем, как это и делается в курсах дифференциальных уравнений, продолжить решение $X(\cdot,\ t_0,\ x_0)$ задачи Коши (1), (2) на максимальный интервал его существования $(a,\ b)$. При этом из теоремы п. 2.5.2 следует, что при $t\downarrow a$ и при $t\uparrow b$ это решеиие должно покидать любой компакт $\mathcal{K}\subset G$, ибо, пока $(t,\ x(t))\in\mathcal{K}$, решение гарантировано продолжается на интервал $(t-\delta,\ t+\delta)$, откуда $a+\delta<t<b-\delta$ ввиду максимальности интервала $(a,\ b)$. Следующая теорема показывает, в каком смысле $X(\cdot\ ,\ t_0,\ x_0)$ непрерывно зависит от $(t_0,\ x_0)$.

Теорема 1. *Пусть на открытом множестве $G\subset\mathbf{R}\times\times\mathbf{R}^n$ функция $F\colon\ G\longrightarrow\mathbf{R}^n$ удовлетворяет условиям* А)—В) п. 2.5.1, *и пусть* $\hat{x}$: $[\hat{t}_0, \hat{t}_1] \to \mathbb{R}^n$ — *решение уравнения* (1), *график которого* $\Gamma = \{(t, \hat{x}(t)) | \hat{t}_0 \leqslant t \leqslant \hat{t}_1\} \subset G$.

Тогда существуют $\hat{\delta} > 0$ *и окрестность* $\hat{G} \supset \Gamma$ *такие, что для любых* $(t_0; x_0) \in \hat{G}$ *решение* $X(\cdot\ , t_0, x_0)$ *задачи Коши* (1)—(2) *определено на отрезке* $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$ *и функция* $(t, t_0, x_0) \mapsto X(t, t_0, x_0)$ *является непрерывной функцией по совокупности аргументов на* $\Delta \times \hat{G}$. *В частности, при* $t_0 \to \tau$, $x_0 \to \hat{x}(\tau)$ *решение* $X(t, t_0, x_0) \to \hat{x}(t)$ *равномерно по* $t \in \Delta$.

Доказательство. А) Применив локальную теорему существования в точках $(\hat{t}_0, \hat{x}(\hat{t}_0))$ и $(\hat{t}_1, \hat{x}(\hat{t}_1))$, продолжим решение $\hat{x}(\cdot)$ на некоторый отрезок $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$ так, чтобы график $\hat{x}(\cdot)$ по-прежнему содержался в G. После этого выберем $\hat{\varepsilon}$ так, чтобы компакт

$$\mathcal{K} = \{(t, x) | |x - \hat{x}(t)| \leqslant \hat{\varepsilon},\ t \in \Delta\} \subset G.$$

Применив к $\mathcal{K}$ теорему п. 2.5.2, найдем такое $\delta > 0$, чтобы при $(t_0, x_0) \in \mathcal{K}$ решение $X(\cdot\ , t_0, x_0)$ было определено на $[t_0 - \delta, t_0 + \delta]$. Наконец, пусть $k(t)$ и $\varkappa(\cdot)$ — локально интегрируемые функции, соответствующие $\mathcal{K}$ согласно условию В) и лемме 1 п. 2.5.1.

Обозначим теперь

$$\hat{G} = \{(t, x) | t \in (\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}),\ |x - \hat{x}(t)| < \varepsilon\}$$

и покажем, что при достаточно малом $\varepsilon > 0$ эта область обладает свойствами, указанными в формулировке теоремы.

Б) Пусть два решения $x_i(\cdot)$, $i = 1, 2$, уравнения (1) определены на отрезке $[\beta, \gamma] \subset \Delta$, причем $(t, x_i(t)) \in \mathcal{K}$, $\beta \leqslant t \leqslant \gamma$, и пусть $\tau \in [\beta, \gamma]$. Применяя теорему о среднем (п. 2.2.3) и неравенство (1) п. 2.5.1, получаем

$$|x_1(t) - x_2(t)| =$$
$$= \left| x_1(\tau) + \int_\tau^t F(s, x_1(s))\,ds - x_2(\tau) - \int_\tau^t F(s, x_2(s))\,ds \right| \leqslant$$
$$\leqslant |x_1(\tau) - x_2(\tau)| + \left| \int_\tau^t [F(s, x_1(s)) - F(s, x_2(s))]\,ds \right| \leqslant$$
$$\leqslant |x_1(\tau) - x_2(\tau)| + \left| \int_\tau^t k(s) |x_1(s) - x_2(s)|\,ds \right|. \quad (3)$$

По лемме п. 2.5.3 ($\alpha(s)=k(s)$, $\omega(s)=|x_1(s)-x_2(s)|$, $b=|x_1(\tau)-x_2(\tau)|$) имеем

$$|x_1(t)-x_2(t)|\leqslant|x_1(\tau)-x_2(\tau)|\,e^{\left|\int\limits_\tau^t k(s)\,ds\right|},\quad \beta\leqslant t\leqslant\gamma. \tag{4}$$

В) Выберем теперь $\varepsilon>0$ так, чтобы

$$\varepsilon e^{\int\limits_\Delta k(s)\,ds}\leqslant\hat{\varepsilon}, \tag{5}$$

и пусть $(t_0,\ x_0)\in\hat{G}$. Поскольку $\hat{G}\subset\mathcal{K}^\circ$, решение $x(\cdot)=X(\cdot,\ t_0,\ x_0)$ определено на $[\beta,\ \gamma]=[t_0-\delta,\ t_0+\delta]\cap\Delta$. Покажем, что оно останется на этом отрезке в $\mathcal{K}$. Действительно, пусть

$$T=\sup\{t\mid t_0\leqslant t\leqslant\min\{t_0+\delta,\ \hat{t}_1+\hat{\delta}\};\ (s,\ x(s))\in\mathcal{K}^\circ,\ \forall s\in[t_0,\ t]\}. \tag{6}$$

На отрезке $[t_0,\ T]$ к решениям $x(\cdot)$ и $\hat{x}(\cdot)$ применимы рассуждения п. Б) доказательства и в силу (4)

$$|x(T)-\hat{x}(T)|\leqslant|x_0-\hat{x}(t_0)|\,e^{\int\limits_{t_0}^T k(s)\,ds}<\varepsilon e^{\int\limits_\Delta k(s)\,ds}\leqslant\hat{\varepsilon}.$$

Но тогда $(T,\ x(T))\in\operatorname{int}\mathcal{K}^\circ$ и точка $(t,\ x(t))$ остается в $\mathcal{K}^\circ$ и при $t>T$, близких к T, что противоречит определению (6). Аналогично рассуждаем и для $t\leqslant t_0$.

Поскольку $(\beta,\ x(\beta))\in\mathcal{K}^\circ$ и $(\gamma,\ x(\gamma))\in\mathcal{K}^\circ$ решение $x(\cdot)$ определено на отрезке $[t_0-2\delta,\ t_0+2\delta]\cap\Delta$. К этому отрезку снова применяем те же рассуждения и убеждаемся, что $(t,\ x(t))$ остается в $\mathcal{K}^\circ$. Продолжая эту процедуру, убеждаемся, что $x(\cdot)$ определено на Δ.

Г) По доказанному любое решение $X(\cdot,\ t_0,\ x_0)$ остается в $\mathcal{K}^\circ$, если $(t_0,\ x_0)\in\hat{G}$. Вспоминая о функции $\varkappa(\cdot)$, получаем неравенство

$$|X(t_1,\ t_0,\ x_0)-X(t,\ t_0,\ x_0)|=\left|\int\limits_t^{t_1}F(s,\ X(s,\ t_0,\ x_0))\,ds\right|\leqslant\left|\int\limits_t^{t_1}\varkappa(s)\,ds\right|. \tag{7}$$

Возьмем теперь два решения, у которых $(t_0, x_0) \in \hat{G}$ и $(\tilde{t}_0, \tilde{x}_0) \in \hat{G}$. Тогда ввиду (4) и (7)

$$
\begin{aligned}
&|X(\tilde{t}, \tilde{t}_0, \tilde{x}_0) - X(t, t_0, x_0)| \leqslant \\
&\leqslant |X(\tilde{t}, \tilde{t}_0, \tilde{x}_0) - X(t, \tilde{t}_0, \tilde{x}_0)| + |X(t, \tilde{t}_0, \tilde{x}_0) - X(t, t_0, x_0)| \leqslant \\
&\qquad \leqslant \left|\int_t^{\tilde{t}} \varkappa(s)\,ds\right| + |\tilde{x}_0 - X(\tilde{t}_0, t_0, x_0)|\, e^{\left|\int_{t_0}^{t} k(s)\,ds\right|} \leqslant \\
&\qquad \leqslant \left|\int_t^{\tilde{t}} \varkappa(s)\,ds\right| + \left\{|\tilde{x} - x_0| + \left|\int_{t_0}^{\tilde{t}_0} \varkappa(s)\,ds\right|\right\} e^{\int_\Delta k(s)\,ds}
\end{aligned}
\tag{8}
$$

что, очевидно, стремится к нулю при $\tilde{t} \to t$, $\tilde{t}_0 \to t_0$, $\tilde{x}_0 \to x_0$, доказывая непрерывность $X(t, t_0, x_0)$. Полагая в (8) $\tilde{t} = t$, $\tilde{t}_0 = \tau$, $\tilde{x}_0 = \hat{x}(\tau)$ и вспоминая, что по теореме единственности $X(t, \tau, \hat{x}(\tau)) \equiv \hat{x}(t)$, имеем

$$
|x(t) - X(t, t_0, x_0)| \leqslant \left\{|\hat{x}(\tau) - x_0| + \int_{t_0}^{\tau} \varkappa(s)\,ds\right\} e^{\int_\Delta k(s)\,ds}
$$

откуда $X(t, t_0, x_0) \to \hat{x}(t)$ равномерно при $x_0 \to \hat{x}(\tau)$, $t_0 \to \tau$. ■

Перейдем теперь к непрерывной зависимости решений от параметра. Предположим, что нам дано семейство дифференциальных уравнений

$$
\dot{x} = F_\alpha(t, x), \tag{9}
$$

зависящих от параметра $\alpha \in \mathfrak{A}$, где $\mathfrak{A}$ — некоторое топологическое пространство.

Теорема 2. *Пусть на открытом множестве* $G \subset \mathbf{R} \times \mathbf{R}^n$ *функции семейства* $\{F_\alpha : G \to \mathbf{R}^n \mid \alpha \in \mathfrak{A}\}$ *удовлетворяют при каждом* $\alpha \in \mathfrak{A}$ *условиям* А) *и* Б) п. 2.5.1; $\hat{x}: [\hat{t}_0, \hat{t}_1] \to \mathbf{R}^n$ — *решение уравнения* (6) *при* $\alpha = \hat{\alpha}$ *и график* $\Gamma = \{(t, \hat{x}(t)) \mid \hat{t}_0 \leqslant t \leqslant \hat{t}_1\} \subset G$, *и пусть еще выполняются условия*

В′) *Для любого компакта* $\mathcal{K} \subset G$ *существуют такие локально интегрируемые функции* $\varkappa(\cdot)$ *и* $k(\cdot)$, *что для* *любых* $(t, x) \in \mathcal{K}$, $\alpha \in \mathfrak{A}$ *имеют место неравенства*

$$|F_\alpha(t, x)| \leqslant \varkappa(t), \quad \|F_{\alpha x}(t, x)\| \leqslant k(t). \qquad (10)$$

Г)
$$\lim_{\alpha \to \hat{\alpha}} \int_{\hat{t}_0}^{\hat{t}_1} |F_\alpha(s, \hat{x}(s)) - F_{\hat{\alpha}}(s, \hat{x}(s))| \, ds = 0. \qquad (11)$$

Тогда существуют $\hat{\delta} > 0$, *окрестность* $\hat{G} \supset \Gamma$ *в* G, *и окрестность* $U \ni \hat{\alpha}$ *в* $\mathfrak{A}$ *такие, что для* $(t_0, x_0) \in \hat{G}$, $\alpha \in U$ *решение* $x_\alpha(t, t_0, x_0)$ *задачи Коши* (9), (2) *определено на отрезке* $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$ *и при* $t_0 \to \tau \in [\hat{t}_0, \hat{t}_1]$, $x_0 \to \hat{x}(\tau)$, $\alpha \to \hat{\alpha}$

$$X_\alpha(t, t_0, x_0) \to \hat{x}(t)$$

равномерно по $t \in [\hat{t}_0, \hat{t}_1]$.

Доказательство этой теоремы проводится по той же схеме, что и доказательство теоремы 1 с небольшими изменениями. В пункте А) мы учитываем, что число δ в локальной теореме существования определяется неравенствами (7) и (8) п. 2.5.2 и в силу условия В′) может быть выбрано одним и тем же для всех $\alpha \in \mathfrak{A}$.

В пункте Б) заменяем $x_1(\cdot)$ решением $x_\alpha(\cdot)$ уравнения (9), а $x_2(\cdot)$ — решением $\hat{x}(\cdot)$. В оценке следует учесть, что теперь правые части уравнений разные, в связи с чем

$$|x_\alpha(t) - \hat{x}(t)| \leqslant$$
$$\leqslant |x_\alpha(\tau) - \hat{x}(\tau)| + \left| \int_\tau^t \{F_\alpha(s, \hat{x}(s)) - F_{\hat{\alpha}}(s, \hat{x}(s))\} \, ds \right| +$$
$$+ \left| \int_\tau^t k(s) |x_\alpha(s) - \hat{x}(s)| \, ds \right|$$

и (4) заменяется неравенством

$$|x_\alpha(t) - \hat{x}(t)| \leqslant \Bigg(|x_\alpha(\tau) - \hat{x}(\tau)| +$$
$$+ \left| \int_\tau^t \{F_\alpha(s, \hat{x}(s)) - F_{\hat{\alpha}}(s, \hat{x}(s))\} \, ds \right| \Bigg) e^{\left| \int_\tau^t k(s)\, ds \right|}. \qquad (4')$$

В неравенстве (5) надо заменить ε на 2ε, а $\hat{\delta}$ и U следует с учетом условий В′) и Г) выбрать так, чтобы выполнялись неравенства

$$\left|\int_{\tau}^{t}\{F_\alpha(s,\ \hat{x}(s))-F_{\hat\alpha}(s,\ \hat{x}(s))\}\,ds\right|\leqslant$$

$$\leqslant\int_{\hat t_0-\hat\delta}^{\hat t_1+\hat\delta}|F_\alpha(s,\ \hat{x}(s)-F_{\hat\alpha}(s,\ \hat{x}(s))|\,ds\leqslant$$

$$\leqslant 2\left(\int_{\hat t_0-\hat\delta}^{\hat t_0}+\int_{\hat t_1}^{\hat t_1+\hat\delta}\varkappa(s)\,ds\right)+$$

$$+\int_{\hat t_0}^{\hat t_1}|F_\alpha(s,\ \hat{x}(s))-F_{\hat\alpha}(s,\ \hat{x}(s))|\,ds\leqslant\varepsilon. \quad (12)$$

Тогда рассуждения пункта В) (с заменой (4) на (4′) остаются в силу одновременно для всех $\alpha\in U$.

В пункте Г) следует в неравенство (8) внести уточнение, вытекающее из замены (4) на (4′), после чего мы получаем

$$|X_\alpha(t,\ t_0,\ x_0)-\hat{x}(t)|\leqslant$$

$$\leqslant\left\{|x_0-\hat{x}(\tau)|+\left|\int_{t_0}^{\tau}\varkappa(s)\,ds\right|+\right.$$

$$\left.+\int_{\hat t_0}^{\hat t_1}|F_\alpha(s,\ \hat{x}(s))-F_{\hat\alpha}(s,\ \hat{x}(s))|\,ds\right\}e^{\int_{\hat t_0}^{\hat t_1}k(s)\,ds}$$

и остальное очевидно. ■

Следствие. *Если выполнены все условия теоремы* 2, *причем условие* Г) *в усиленной форме*:

Г′) *для любой функции* $\bar{x}$: $[\bar t_0,\ \bar t_1]\to\mathbf{R}^n$, *график которой лежит в* G *и любого* $\bar\alpha\in\mathfrak{A}$

$$\lim_{\alpha\to\bar\alpha}\int_{\bar t_0}^{\bar t_1}|F_\alpha(s,\ \bar{x}(s))-F_{\bar\alpha}(s,\ \bar{x}(s))|\,ds=0,$$

то решение $x_\alpha(t,\ t_0,\ x_0)$ *задачи Коши* (9), (2) *является на* $\Delta\times\hat{G}\times U$ *непрерывной функцией от* $(t,\ t_0,\ x_0,\ \alpha)$.

Доказательство. При любых $(\bar t_0,\ \bar x_0)\in\hat{G}$ и $\bar\alpha\in U$ график решения $\bar{x}(\cdot)=X_{\bar\alpha}(\cdot,\ \bar t_0,\ \bar x_0)$: $\Delta\to\mathbf{R}^n$ лежит в G.

Теперь мы можем применить теорему 2, заменив в ее условиях $\hat{x}(\cdot)$ на $\tilde{x}(\cdot)$. ■

Аналогом основного свойства фундаментальной матрицы линейной системы (формула (11) предыдущего пункта) является следующее тождество:

$$X(t, \tau, X(\tau, t_0, x_0)) \equiv X(t, t_0, x_0). \tag{13}$$

Оно автоматически вытекает из теоремы единственности, поскольку каждая из частей является решением уравнения (1) и при $t=\tau$ обе они совпадают. Читателям рекомендуется продумать связь между этим тождеством и принципом Гюйгенса п. 1.1.3.

2.5.6. Теорема о дифференцируемой зависимости решений от начальных данных. До сих пор наше изложение было основано на сравнительно слабых предположениях А)—В) п. 2.5.1. Чтобы продвинуться далее и установить дифференцируемость решения задачи Коши $X(t, t_0, x_0)$ по начальному значению x_0, эти предположения придется несколько усилить. При этом открывается возможность применения классической теоремы о неявной функции. В доказательстве теоремы этого пункта мы не будем останавливаться на существовании решения $X(t, t_0, x_0)$ для $(t_0, x_0) \in \hat{G}$ и его продолжимости на все $t \in \Delta$, поскольку и то и другое следует из теоремы 1 п. 2.5.5. Тем не менее читателю полезно убедиться в том, что оба эти факта также вытекают из теоремы о неявной функции и при желании можно было бы обойтись без ссылки на предыдущий пункт. (Аналогичное рассуждение будет полностью проведено в п. 2.5.7, где при еще более сильных предположениях доказывается дифференцируемость по t_0 и теорема о неявной функции применяется в несколько ином варианте, чем здесь.)

Теорема. *Пусть на открытом множестве* $G \subset \mathbf{R} \times \mathbf{R}^n$ *функция* $F: G \longrightarrow \mathbf{R}^n$ *удовлетворяет условиям* А) *и* В) *п. 2.5.1 и условию*:

Б′) *для любого* t *функция* $x \longmapsto F(t, x)$ *непрерывно дифференцируема на сечении* $G_t = \{x \mid (t, x) \in G\}$. *Пусть* $\hat{x}: [\hat{t}_0, \hat{t}_1] \longrightarrow \mathbf{R}^n$ — *решение уравнения* (1) п. 2.5.5, *график* $\Gamma = \{(t, \hat{x}(t)) \mid \hat{t}_0 \leqslant t \leqslant \hat{t}_1\}$ *которого содержится в* G, *и пусть* $\hat{\delta}$, ε *и* $\hat{G}$ — *те же, что в теореме* 1 п. 2.5.5; $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$.

Тогда для любого фиксированного $t_0 \in \Delta$ *отображение* $x_0 \longmapsto X(\cdot, t_0, x_0)$ *из* $\mathring{B}(\hat{x}(t_0), \varepsilon)$ *в* $C(\Delta, \mathbf{R}^n)$ *непрерывно* *дифференцируемо по Фреше, причем*

$$\frac{\partial X(t, t_0, x_0)}{\partial x_0} = \Omega(t, t_0), \tag{1}$$

где Ω — фундаментальная матрица решений уравнения в вариациях

$$\dot{z} = F_x(t, X(t, t_0, x_0))\, z. \tag{2}$$

Доказательство. Пусть компакт $\mathcal{K}$ и функция $k(t)$ — те же, что и в теореме 1 п. 2.5.5. Множество

$$\mathcal{G} = \{x(\cdot) \in C(\Delta, \mathbf{R}^n) \mid (t, x(t)) \in G,\ t \in \Delta\},$$

очевидно, открыто в $C(\Delta, \mathbf{R}^n)$. Равенством

$$\mathcal{F}(x_0, x(\cdot))(t) = x_0 + \int_{t_0}^{t} F(s, x(s))\, ds - x(t) \tag{3}$$

определим отображение $\mathcal{F}\colon \mathbf{R}^n \times \mathcal{G} \to C(\Delta, \mathbf{R}^n)$. Очевидно, что

$$\mathcal{F}(x_0, x(\cdot)) = 0 \Leftrightarrow x(t) \equiv X(t, t_0, x_0). \tag{4}$$

Пусть $\bar{x}_0 \in \mathring{B}(\hat{x}(t_0), \varepsilon)$ фиксировано и $\bar{x}(t) \equiv X(t, t_0, \bar{x}_0)$. Проверим, что в окрестности точки $(\bar{x}_0, \bar{x}(\cdot))$ применима классическая теорема о неявной функции п. 2.3.4. Действительно, $\mathcal{F}_{x_0}(x_0, x(\cdot))[\xi_0] = \xi_0$, очевидно, непрерывна. Теперь вычислим производную Гато:

$$\mathcal{F}_{x(\cdot)}(x_0, x(\cdot))[\xi(\cdot)](t) =$$
$$= \lim_{\alpha \downarrow 0} \frac{\mathcal{F}(x_0, x(\cdot) + \alpha\xi(\cdot)) - \mathcal{F}(x_0, x(\cdot))}{\alpha}(t) =$$
$$= \lim_{\alpha \downarrow 0} \int_{t_0}^{t} \frac{F(s, x(s) + \alpha\xi(s)) - F(s, x(s))}{\alpha}\, ds - \xi(t) =$$
$$= \int_{t_0}^{t} F_x(s, x(s))\, \xi(s)\, ds - \xi(t). \tag{5}$$

Чтобы убедиться в существовании предела, равномерного по t, заметим, что по теореме о среднем (п. 2.2.3)

$$\left| \int_{t_0}^{t} \frac{F(s, x(s)) + \alpha\xi(s)) - F(s, x(s))}{\alpha}\, ds - \int_{t_0}^{t} F_x(s, x(s))\, \xi(s)\, ds \right| \leqslant \int_{\Delta} r_\alpha(s)\, ds,$$

где

$$r_\alpha(s)=\max_{c\in[x(s),\,x(s)+\alpha\xi(s)]}\|F_x(s,c)-F_x(s,x(s))\|\,|\xi(s)|.$$

Ввиду условия Б′) $r_\alpha(s)\to 0$ при каждом s и $\alpha\downarrow 0$, а в силу (1) п. 2.5.1 $|r_\alpha(s)|\leqslant 2k(s)\|\xi\|$. По теореме Лебега об ограниченной сходимости [КФ, стр. 302] убеждаемся в существовании предела (5).

Оператор

$$\mathscr{F}_{x(\cdot)}(x_0,x(\cdot))[\xi(\cdot)](t)=\int_{t_0}^{t}F_x(s,x(s))\,\xi(s)\,ds$$

ограничен:

$$\|\mathscr{F}_{x(\cdot)}(x_0,x(\cdot))[\xi(\cdot)]\|\leqslant$$
$$\leqslant\max_t\left|\int_{t_0}^{t}\|F_x(s,x(s))\|\,|\xi(s)|\,ds\right|\leqslant\int_\Delta k(s)\,ds\,\|\xi\|$$

и непрерывно зависит от $x(\cdot)$, так как

$$\|\mathscr{F}_{x(\cdot)}(\tilde{x}_0,\tilde{x}(\cdot))-\mathscr{F}_{x(\cdot)}(x_0,x(\cdot))\|=$$
$$=\sup_{\|\xi\|\leqslant 1}\sup_t\left|\int_{t_0}^{t}F_x(s,\tilde{x}(s))\,\xi(s)\,ds-\int_{t_0}^{t}F_x(s,x(s))\,\xi(s)\,ds\right|\leqslant$$
$$\leqslant\int_\Delta\|F_x(s,\tilde{x}(s))-F_x(s,x(s))\|\,ds.$$

При $\|\tilde{x}(\cdot)-x(\cdot)\|\to 0$ имеем в силу Б′)

$$\|F_x(s,\tilde{x}(s))-F_x(s,x(s))\|\to 0,\quad\forall s\in\Delta,$$

и снова в силу (1) п. 2.5.1 $\|F_x(s,\tilde{x}(s))-F_x(s,x(s))\|\leqslant\leqslant 2k(s)$, так что применима теорема Лебега. Сославшись на следствие 2 п. 2.2.3, устанавливаем, что существует непрерывная производная Фреше $\mathscr{F}_{x(\cdot)}$.

Пусть $\eta(\cdot)\in C(\Delta,\mathbf{R}^n)$ произвольна. Согласно (5)

$$\mathscr{F}_{x(\cdot)}(x_0,x(\cdot))[\xi(\cdot)]=\eta(\cdot)\Leftrightarrow$$
$$\Leftrightarrow\int_{t_0}^{t}F_x(s,x(s))\,\xi(s)\,ds-\xi(t)=\eta(t)\quad(6)$$

и для применимости теоремы о неявной функции нужно убедиться в однозначной разрешимости этого уравнения. Подстановкой $\xi(t)+\eta(t)=\zeta(t)$ приведем его к интегральному уравнению

$$\zeta(t)=\int_{t_0}^{t} F_x(s, x(s))[\zeta(s)-\eta(s)]\,ds,$$

которое в свою очередь эквивалентно линейному дифференциальному уравнению

$$\dot{\zeta}(t)=F_x(t, x(t))\zeta(t)-F_x(t, x(t))\eta(t);\quad \zeta(t_0)=0. \tag{7}$$

Остается сослаться на лемму п. 2.5.4. По теореме Банаха (п. 2.1.5) $\mathscr{F}_{x(\cdot)}^{-1}$ ограничен.

По теореме о неявной функции существует непрерывно дифференцируемое в некоторой окрестности $U \ni \bar{x}_0$ отображение φ: $U \to C(\Delta, \mathbf{R}^n)$ такое, что $\mathscr{F}(x_0, \varphi(x_0)) \equiv 0$. Согласно (4) $\varphi(x_0)(t) \equiv X(t, t_0, x_0)$, так что отображение $x_0 \mapsto X(t, t_0, x_0)$ действительно дифференцируемо по Фреше в точке $\bar{x}_0$. По той же теореме о неявной функции

$$\frac{\partial X(t, t_0, x_0)}{\partial x_0}\xi_0=\varphi'(x_0)[\xi_0](t)=-(\mathscr{F}_{x(\cdot)}^{-1}\circ\mathscr{F}_{x_0}[\xi_0])(t)=\xi(t)$$

и, как уже было найдено, $\xi(t)$ — решение интегрального уравнения (6), принимающего вид

$$\int_{t_0}^{t} F_x(s, x(s))\xi(s)\,ds-\xi(t)=-\xi_0.$$

Следовательно,

$$\dot{\xi}=F_x(s, x(s))\xi,\quad \xi(t_0)=\xi_0$$

и по теореме п. 2.5.4 $\xi(t)=\Omega(t, t_0)\xi_0$. Отсюда

$$\frac{\partial X}{\partial x_0}(t, t_0, x_0)=\Omega(t, t_0). \blacksquare$$

2.5.7. Классическая теорема о дифференцируемой зависимости решений от начальных данных. В этом пункте мы снова будем рассматривать задачу Коши

$$\dot{x}=F(t, x), \tag{1}$$
$$x(t_0)=x_0, \tag{2}$$

но уже в классической ситуации, когда $F(t, x)$ и ее частная производная $F_x(t, x)$ непрерывны в G. Для удобства читателей, доказательство формулируемой ниже теоремы будет проведено независимо от доказательств теорем пп. 2.5.5 и 2.5.6. Ссылки на локальную теорему существования (п. 2.5.2) и теорему единственности (п. 2.5.3) можно заменить ссылками на любой учебник по дифференциальным уравнениям.

Теорема. *Пусть на открытом множестве* $G \subset \mathbf{R} \times \mathbf{R}^n$ *функция* $F\colon G \to \mathbf{R}^n$ *и ее частная производная* F_x *непрерывны, и пусть* $\hat{x}\colon [\hat{t}_0, \hat{t}_1] \to \mathbf{R}^n$ — *решение уравнения* (1), *график которого* $\Gamma = \{(t, \hat{x}(t)) \mid \hat{t}_0 \leqslant t \leqslant \hat{t}_1\}$ *содержится в* G.

Тогда существует такое $\hat{\delta} > 0$ *и такая окрестность* $\hat{G} \supset \Gamma$, *что для любых* $(t_0, x_0) \in \hat{G}$ *решение* $X(t, t_0, x_0)$ *задачи Коши* (1), (2) *определено на* $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$, *является непрерывно дифференцируемой функцией по совокупности аргументов в области* $(\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}) \times \hat{G}$ *и при этом*

$$X(t, \tau, \hat{x}(\tau)) = \hat{x}(\tau), \tag{3}$$

$$\frac{\partial X}{\partial t}(t, \tau, \hat{x}(\tau)) = F(t, \hat{x}(t)), \tag{4}$$

$$\frac{\partial X}{\partial t_0}(t, t_0, \hat{x}(\tau))\Big|_{t_0=\tau} = -\Omega(t, \tau) F(\tau, \hat{x}(\tau)), \tag{5}$$

$$\frac{\partial X}{\partial x_0}(t, \tau, x_0)\Big|_{x_0=\hat{x}(\tau)} = \Omega(t, \tau), \tag{6}$$

где $\Omega(t, \tau)$ — *фундаментальная матрица решений системы уравнений в вариациях*

$$\dot{z} = F_x(t, \hat{x}(t)) z.$$

Доказательство. А) Применив локальную теорему существования в точках $(\hat{t}_0, \hat{x}(\hat{t}_0))$ и $(\hat{t}_1, \hat{x}(\hat{t}_1))$, продолжим $\hat{x}(\cdot)$ на отрезок $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$ так, чтобы график $\{(t, \hat{x}(t)) \mid t \in \Delta\}$ по-прежнему содержался в G. После этого выберем $\varepsilon > 0$ так, чтобы

$$\mathscr{K} = \{(t, x) \mid t \in \Delta, |x - \hat{x}(t)| \leqslant \varepsilon\} \subset G.$$

На компакте $\mathscr{K}$ функции $F(t, x)$ и $F_x(t, x)$ равномерно непрерывны.

Б) Теперь в области

$$\mathscr{G} = \{(x(\cdot), t_0, x_0) \mid |x(t) - \hat{x}(t)| < \varepsilon, \forall t \in \Delta; t_0 \in \operatorname{int} \Delta, x_0 \in \mathbf{R}^n\} \subset C^1(\Delta, \mathbf{R}^n) \times \mathbf{R} \times \mathbf{R}^n$$

определим функцию $\mathscr{F}\colon \mathscr{G} \to C(\Delta, \mathbf{R}^n) \times \mathbf{R}^n$ равенством

$$\mathscr{F}(x(\cdot), t_0, x_0)(t) = \begin{pmatrix} \dfrac{dx(t)}{dt} - F(t, x(t)) \\ x(t_0) - x_0 \end{pmatrix}. \tag{7}$$

Очевидно, что равенство $\mathcal{F}(x(\cdot), t_0, x_0)=0$ эквивалентно задаче Коши (1), (2). В частности, по условию теоремы $\mathcal{F}(\hat{x}(\cdot), \tau, \hat{x}(\tau))=0$.

Функция $\mathcal{F}$ непрерывно дифференцируема в области $\mathcal{G}$.

Действительно, в (7) входят линейные отображения $x(\cdot)\mapsto dx(\cdot)/dt$, $x_0\mapsto -x_0$, оператор краевых условий $(x(\cdot), t_0)\mapsto x(t_0)$ и оператор Немыцкого $\{x(t)\}\mapsto\{F(t, x(t))\}$, которые непрерывно дифференцируемы согласно пп. 2.4.1, 2.4.3. При этом

$$\mathcal{F}_{x(\cdot)}(x(\cdot), t_0, x_0)[\xi(\cdot)]=\begin{pmatrix}\dfrac{d\xi}{dt}-F_x(t, x(t))\,\xi\\ \xi(t_0)\end{pmatrix}, \tag{8}$$

$$\mathcal{F}_{t_0}(x(\cdot), t_0, x_0)[\alpha]=\begin{pmatrix}0\\ \dot{x}(t_0)\,\alpha\end{pmatrix}, \tag{9}$$

$$\mathcal{F}_{x_0}(x(\cdot), t_0, x_0)[\xi_0]=\begin{pmatrix}0\\ -\xi_0\end{pmatrix}. \tag{10}$$

В) Пусть $\hat{x}=\hat{x}(\hat{t})$, так что $(\hat{t}, \hat{x})\in\Gamma$. Проверим обратимость оператора $\mathcal{F}_{x(\,)}(\hat{x}(\cdot), \hat{t}, \hat{x})$. Равенство

$$\mathcal{F}_{x(\,)}(\hat{x}(\cdot), \hat{t}, \hat{x})[\xi(\cdot)]=\begin{pmatrix}\zeta(\,)\\ \gamma\end{pmatrix}\in C(\Delta, \mathbf{R}^n)\times\mathbf{R}^n$$

эквивалентно задаче Коши

$$\dot{\xi}-F_x(t, \hat{x}(t))\,\xi=\zeta(t), \quad \xi(\hat{t})=\gamma, \tag{11}$$

которая по теореме п. 2.5.4 имеет решение

$$\xi(t)=\Omega(t, \hat{t})\,\gamma+\int_{\hat{t}}^{t}\Omega(t, s)\,\zeta(s)\,ds.$$

Таким образом, $\mathcal{F}_{x(\,)}$ отображает банахово пространство $C^1(\Delta, \mathbf{R}^n)$ на все банахово пространство $C(\Delta, \mathbf{R}^n)\times\mathbf{R}^n$. Поскольку при $\zeta(\cdot)=0$ и $\gamma=0$, задача (11) имеет только нулевое решение, $\mathcal{F}_{x(\,)}$ взаимно однозначен. По теореме Банаха (п. 2.1.5) оператор $\mathcal{F}_{x(\,)}^{-1}$ существует и ограничен.

Г) По классической теореме о неявной функции (п 2.3.4) существует такое $\delta>0$ и такое непрерывно дифференцируемое отображение

$$\Phi: \{(t_0, x_0)\,|\,|t_0-\hat{t}\,|<\delta, |x_0-\hat{x}|<\delta\}\to C^1(\Delta, \mathbf{R}^n),$$

что

$$\mathcal{F}(\Phi(t_0, x_0), t_0, x_0)\equiv 0.$$

Обозначим

$$X(t,\ t_0,\ x_0)=\Phi(t_0,\ x_0)(t).$$

Тогда $X(t,\ t_0,\ x_0)$ — решение задачи Коши (1), (2) и при этом

$$\frac{\partial X}{\partial t}=F(t,\ X(t,\ t_0,\ x_0)), \tag{12}$$

$$\frac{\partial X}{\partial t_0}=\frac{\partial \Phi}{\partial t_0}(t_0,\ x_0)(t), \tag{13}$$

$$\frac{\partial X}{\partial x_0}=\frac{\partial \Phi}{\partial x_0}(t_0,\ x_0)(t). \tag{14}$$

Если $(\tilde{t}_0,\ \tilde{x}_0)\to(t_0,\ x_0)$, то $\Phi(\tilde{t}_0,\ \tilde{x}_0)\to\Phi(t_0,\ x_0)$ в $C^1(\Delta,\ \mathbb{R}^n)$, т. е. $X(t,\ \tilde{t}_0,\tilde{x}_0)\to X(t,\ t_0,\ x_0)$ равномерно на Δ. Отсюда вытекает непрерывность $X(t,\ t_0,\ x_0)$ по совокупности аргументов. То же рассуждение вместе с формулами (12)—(14) показывает, что и производные функции $X(t,\ t_0,\ x_0)$ непрерывны.

Д) Предыдущие рассуждения были применимы к начальным условиям $(t_0,\ x_0)$, находящимся в δ-окрестности точки $(\hat{t},\ \hat{x})\in\Gamma$. Эта точка выбиралась произвольно, и можно покрыть весь график Γ такими окрестностями. Если $X^{(1)}(t,\ t_0,\ x_0)$ и $X^{(2)}(t,\ t_0,\ x_0)$ определены в двух из этих окрестностей, которые имеют общую точку $(t_0,\ x_0)$, то

$$X^{(1)}(t_0,\ t_0,\ x_0)=x_0=X^{(2)}(t_0,\ t_0,\ x_0)$$

и по теореме единственности п. 2.5.3

$$X^{(1)}(t,\ t_0,\ x_0)=X^{(2)}(t,\ t_0,\ x_0),\quad \forall t\in\Delta.$$

Следовательно, $X(t,\ t_0,\ x_0)$ корректно определено для $(t_0,\ x_0)$, принадлежащих некоторой окрестности Γ и, очевидно, непрерывно дифференцируемо в этой окрестности.

Е) Остается получить формулы (3)—(6). Равенство (3) имеет место по теореме единственности, а тогда (4) является следствием (12). По теореме о неявной функции

$$\frac{\partial \Phi}{\partial t_0}(t_0,\ x_0)=-\mathscr{F}_{x(\ )}^{-1}\circ\mathscr{F}_{t_0},$$

а потому из (13) и (9) имеем

$$\frac{\partial X}{\partial t_0}(t,\ \tau,\hat{x}(\tau))[\alpha]=-\mathscr{F}_{x(\ )}^{-1}(\hat{x}(\cdot),\ \tau,\ \hat{x}(\tau))\begin{pmatrix}0\\ \dot{\hat{x}}(\tau)\alpha\end{pmatrix}.$$

Поэтому $\frac{\partial X}{\partial t_0}(t, \tau, \hat{x}(\tau))$ является (по t) решением задачи Коши (11) с $\zeta=0$ и $\gamma=-\dot{\hat{x}}(\tau)=-F(\tau, \hat{x}(\tau))$, и потому имеет место (5).

Аналогично

$$\frac{\partial \Phi}{\partial x_0}(t_0, x_0)=-\mathcal{F}^{-1}_{x(\cdot)}\circ \mathcal{F}_{x_0},$$

а потому из (14) и (10) имеем

$$\frac{\partial X}{\partial x_0}(t, \tau, \hat{x}(\tau))[\gamma]=-\mathcal{F}^{-1}_{x(\cdot)}(\hat{x}(\cdot), \tau, \hat{x}(\tau))\begin{pmatrix}0\\ -\gamma\end{pmatrix}.$$

Следовательно, $\frac{\partial X}{\partial x_0}(t, \tau, \hat{x}(\tau))$ является (по t) решением задачи Коши (11) с $\zeta=0$, откуда

$$\frac{\partial X}{\partial x_0}(t, \tau, \hat{x}(\tau))[\gamma]=\Omega(t, \tau)\gamma,$$

а это равносильно (6). ■

§ 2.6*. Элементы выпуклого анализа

Выпуклый анализ — это раздел математики, где изучают выпуклые множества и функции. Роль этих понятий была уже продемонстрирована в п. 1.3.3 (теорема Куна — Таккера) и, учитывая, что в большинстве современных работ по экстремальным задачам выпуклость занимает важное место, мы, кроме элементарных сведений, даем здесь набросок теории двойственности (теорема Фенхеля — Моро) и рассказываем о простейших свойствах субдифференциала — понятия, обобщающего на случай выпуклых функций понятие дифференциала. В §§ 3.3, 3.4, 4.3 будет показано, как применяются эти понятия. Изложение этого параграфа опирается по преимуществу на пп. 2.1.3 и 2.1.4.

2.6.1. Основные определения. Пусть X — линейное вещественное пространство.

Определение 1. а) Множество $C\subset X$ называется *выпуклым*, если вместе с любыми двумя своими точками x_1 и x_2 оно содержит весь отрезок

$$[x_1, x_2]=\{x \mid x=\alpha x_1+(1-\alpha)x_2,\ 0\leqslant\alpha\leqslant 1\}.$$

б) Множество $A\subset X$ называется *аффинным многообразием*, если вместе с любыми двумя своими точками x_1 и

x_2 оно содержит всю прямую

$$\{x \mid x = \alpha x_1 + (1-\alpha) x_2, \ \alpha \in \mathbb{R}\}.$$

в) Множество $K \subset X$ называется *конусом* (с вершиной в начале), если вместе с любой своей точкой x_0 оно содержит весь луч

$$\{\alpha x_0 \mid \alpha > 0\}.$$

Пустое множество по определению считается выпуклым, аффинным многообразием и конусом. Совокупность всех выпуклых множеств из X обозначается $\mathfrak{B}(X)$.

Непосредственно из определения 1 вытекает

Предложение 1. а) *Пересечение любого числа выпуклых множеств (соответственно аффинных многообразий или конусов) само является выпуклым множеством (соответственно аффинным многообразием или конусом).*

б) *Образ $f(A)$ и полный прообраз $f^{-1}(B)$ выпуклых множеств (аффинных многообразий, конусов) $A \subset X$, $B \subset Y$ при линейном отображении $f\colon X \to Y$ является выпуклым множеством (аффинным многообразием, конусом).*

в) *Сдвиг $C + \xi$ выпуклого множества C (аффинного многообразия) является выпуклым множеством (аффинным многообразием).*

г) *Конус K является выпуклым множеством тогда и только тогда, когда*

$$x_1, x_2 \in K \Rightarrow (x_1 + x_2) \in K.$$

д) *Аффинное многообразие A является линейным подпространством в X тогда и только тогда, когда $0 \in A$.*

Упражнение 1. Докажите, что аффинное многообразие является сдвигом некоторого линейного подпространства.

Определение 2. а) Пересечение всех выпуклых множеств C, содержащих данное множество M, называется *выпуклой оболочкой* множества M и обозначается $\operatorname{conv} M$;

$$\operatorname{conv} M = \bigcap_{C \supset M} C, \quad C \in \mathfrak{B}(M). \tag{1}$$

б) Если в (1) вместо выпуклых множеств берутся всевозможные выпуклые конусы $K \supset M$ (аффинные многообразия $A \supset M$, линейные подпространства $L \supset M$), то пересечение называется *конической* (соответственно *аффинной* или *линейной*) оболочкой M и обозначается $\operatorname{cone} M$ ($\operatorname{aff} M$, $\operatorname{lin} M$).

Определение 3. Пусть $x_1, \ldots, x_n$ — элементы из X. Элемент

$$x = \sum_{i=1}^{n} \lambda_i x_i \tag{2}$$

называется *линейной*, *аффинной*, *конической* или *выпуклой* комбинацией $x_1, \ldots, x_n$, если в (2) соответственно:

для линейной — λ_i любые,

для аффинной — $\sum_{i=1}^{n} \lambda_i = 1$,

для конической — $\lambda_i \geqslant 0$,

для выпуклой — $\sum_{i=1}^{n} \lambda_i = 1$, $\lambda_i \geqslant 0$.

По индукции доказывается, что если $x_1, \ldots, x_n$ принадлежат выпуклому множеству C, выпуклому конусу K, аффинному многообразию A или линейному подпространству L, то соответственно их выпуклая, коническая, линейная или аффинная комбинация принадлежит C, K, A или L.

Предложение 2. а) *Выпуклая (коническая, аффинная или линейная) оболочка множества M состоит из всех выпуклых (конических, аффинных или линейных) комбинаций элементов из M.*

б) *Множество M выпукло (является выпуклым конусом, аффинным многообразием или линейным подпространством) тогда и только тогда, когда оно совпадает со своей выпуклой (конической, аффинной или линейной) оболочкой.*

Доказательство проведем для выпуклого множества M — в остальных случаях оно аналогично. Обозначим через $\tilde{M}$ множество всех выпуклых комбинаций точек из M. Множество $\tilde{M} \supset M$ и выпукло, ибо если

$$x_1 = \sum_{i=1}^{m_1} \alpha_{i1} x_{i1}, \quad x_2 = \sum_{i=1}^{m_2} \alpha_{i2} x_{i2}, \quad \alpha_{ij} \geqslant 0,$$

$$\sum_{i=1}^{m_j} \alpha_{ij} = 1, \quad j = 1, 2,$$

то для $\alpha \in [0, 1]$

$$\alpha x_1 + (1-\alpha) x_2 = \sum_{i=1}^{m_1} \alpha \alpha_{i1} x_{i1} + \sum_{i=1}^{m_2} (1-\alpha) \alpha_{i2} x_{i2},$$

а эта последняя сумма является выпуклой комбинацией точек $x_{11}, \ldots, x_{m_1 1}, x_{12}, \ldots, x_{m_2 2}$. Значит, $\operatorname{conv} M \subset \tilde{M}$ (ибо $\operatorname{conv} M$ есть пересечение всех выпуклых множеств, содержащих M). С другой стороны, как было замечено выше, каждая точка из $\tilde{M}$ содержится в любом выпуклом множестве, содержащем M, и потому $\tilde{M} \subset \operatorname{conv} M$.

Если M выпукло, то в (1) можно взять $C = M$, и тогда, очевидно, $\operatorname{conv} M = M$. Обратно, если $\operatorname{conv} M = M$, то M выпукло по определению выпуклой оболочки и предложению 1 а). ■

Упражнение 2. Докажите, что алгебраическая сумма

$$M_1 + \ldots + M_n = \left\{ x \mid x = \sum_{i=1}^{n} x_i, \ x_i \in M, \ i = 1, \ldots, n \right\}$$

конечного числа выпуклых множеств (конусов, аффинных многообразий, линейных подпространств) обладает тем же свойством.

Примеры. 1) Выпуклые непустые множества на прямой — это одноточечные множества и промежутки всех видов (отрезки, интервалы, полуинтервалы, лучи замкнутые и открытые и вся прямая).

2) Всякое линейное подпространство или аффинное многообразие — выпуклое множество.

Упражнение 3. Докажите эти утверждения.

3) *Выпуклым многогранником* называется выпуклая оболочка конечного множества точек. Важный частный случай — *n-мерный симплекс*, т. е. выпуклая оболочка $n+1$ аффинно независимых точек (никакая из них не является аффинной комбинацией остальных).

Упражнение 4. Докажите, что выпуклая оболочка трех точек в $\mathbf{R}^2$, не лежащих одной прямой (двумерный симплекс), является треугольником с вершинами в этих точках. Что будет, если точки лежат на одной прямой?

В конечномерном пространстве предложение 2 можно усилить. Соответствующее утверждение мы называем *теоремой Каратеодори*, хотя, строго говоря, это название относится лишь к той его части, в которой идет речь о выпуклой оболочке.

Теорема Каратеодори. *Пусть* $\dim (\operatorname{lin} A) < \infty$. *Каждая точка, принадлежащая выпуклой (конической, аффинной или линейной) оболочке множества A, является выпуклой (конической, аффинной или линейной) комбинацией не более чем s точек $x_1, \ldots, x_s$ из A.*

Для выпуклой и аффинной оболочек $s=n+1$ и точки $x_1, \ldots, x_s$ можно считать аффинно независимыми.

Для конической и линейной оболочек $s=n$ и точки $x_1, \ldots, x_s$ можно считать линейно независимыми.

Доказательство. A) Ограничимся наиболее сложным случаем выпуклой оболочки, указав затем, какие изменения следует внести в доказательство в остальных трех случаях.

Согласно предложению 2 а) каждое $x \in \operatorname{conv} A$ может быть представлено в виде

$$x=\lambda_1 x_1+\ldots+\lambda_k x_k, \quad \lambda_k>0,$$
$$\sum_{i=1}^{k} \lambda_i=1, \quad x_i \in A \tag{3}$$

(в определении выпуклой комбинации $\lambda_i \geqslant 0$, но ясно, что $\lambda_i=0$ можно опустить).

Покажем, что если $x_1, \ldots, x_k$ аффинно зависимы, то число точек в этом наборе можно уменьшить, сохранив представление вида (3). Действительно, если точки $x_1, \ldots, x_k$ аффинно зависимы, то одна из них есть аффинная комбинация остальных; без ограничения общности $x_k=\mu_1 x_1+\ldots+\mu_{k-1} x_{k-1}$, $\sum_{i=1}^{k-1} \mu_i=1$. Положив еще $\mu_k=-1$, имеем $\sum_{i=1}^{k} \mu_i x_i=0$, $\sum_{i=1}^{k} \mu_i=0$. Пусть теперь

$$\alpha=\min \left\{-\frac{\lambda_i}{\mu_i} \middle| \mu_i<0\right\}.$$

Тогда

$$x=\sum_{i=1}^{k} \lambda_i x_i+\alpha \sum_{i=1}^{k} \mu_i x_i=\sum_{i=1}^{k}\left(\lambda_i+\alpha \mu_i\right) x_i=\sum_{i=1}^{k} \lambda_i' x_i,$$

причем все $\lambda_i' \geqslant 0$ в силу выбора α и по крайней мере одно из них равно нулю. Удалив из набора $\{x_1, \ldots, x_k\}$ те x_i, для которых $\lambda_i'=0$, мы получаем представление вида (3) с меньшим числом точек $\left(\sum_{i=1}^{k} \lambda_i'=\sum_{i=1}^{k} \lambda_i+\right.$ $\left.+\alpha \sum_{i=1}^{k} \mu_i=1\right)$.

Обозначим через $s(x)$ наименьшую мощность набора $\{x_1, \ldots, x_k\}$, для которого представление (3) возможно (в каждом непустом множестве натуральных чисел есть наименьшее). Только что проведенное рассуждение показывает, что в наборе $\{x_1, \ldots, x_{s(x)}\}$ точки должны быть аффинно независимы. Остается заметить, что в n-мерном пространстве любые $n+2$ точки $\{x_1, \ldots, x_{n+2}\}$ аффинно зависимы. Действительно, точки $x_i - x_{n+2}$, $i=1, \ldots, n+1$, должны быть линейно зависимыми, и если, например,

$$x_{n+1} - x_{n+2} = \sum_{i=1}^{n} \lambda_i (x_i - x_{n+2}), \text{ то } x_{n+1} = \left(1 - \sum_{i=1}^{n} \lambda_i\right) x_{n+2} + \sum_{i=1}^{n} \lambda_i x_i.$$

Так как $1 - \sum_{i=1}^{n} \lambda_i + \sum_{i=1}^{n} \lambda_i = 1$, то x_{n+1} является аффинной комбинацией остальных.

Б) Для случая аффинной оболочки в приведенном рассуждении опускается все, связанное с неравенствами $\lambda_i > 0$, $\lambda'_i \geqslant 0$, и можно положить $\alpha = -\lambda_i/\mu_i$ для любого i с $\mu_i \neq 0$.

В остальных двух случаях аффинные комбинации заменяются линейными и число линейно независимых точек не может быть больше n. Для конической оболочки сохраняются рассуждения с неравенствами; для линейной оболочки и их можно опустить. ■

Переходим к описанию выпуклых функций. Здесь, как и далее в этом параграфе, будем называть функцией отображение $f\colon X \to \overline{\mathbf{R}}$, где $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$ — расширенная числовая прямая[1]). С каждой функцией можно связать два множества:

$$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\},$$
$$\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbf{R} \mid \alpha \geqslant f(x),\ x \in \operatorname{dom} f\},$$

называемые соответственно *эффективным множеством* и *надграфиком функции* f.

---

[1]) Арифметические операции и неравенства, в которых участвуют несобственные числа $+\infty$ и $-\infty$, определяются так (далее $a \in \mathbf{R}$ любое, $p > 0$):

$$\begin{array}{lll}
(+\infty)+(+\infty)=+\infty, & (+\infty)(\pm\infty)=\pm\infty, & +\infty > a,\\
\pm\infty + a = \pm\infty, & (-\infty)(\pm\infty)=\mp\infty, & -\infty < a,\\
(-\infty)+(-\infty)=-\infty, & (\pm\infty)\cdot p=\pm\infty, & +\infty > -\infty,\\
-(+\infty)=-\infty, & (\pm\infty)\cdot 0=0, & \\
-(-\infty)=+\infty, & (\pm\infty)\cdot(-p)=\mp\infty; &
\end{array}$$

выражения $(+\infty)-(+\infty)$, $(+\infty)+(-\infty)$ и $(-\infty)-(-\infty)$ объявляются лишенными смысла; не определяются и операции деления с несобственными числами.

Упражнение 5. Докажите, что множество $C \subset X \times \mathbf{R}$ является надграфиком некоторой функции $f\colon X \longrightarrow \overline{\mathbf{R}}$ тогда и только тогда, когда $(x, \alpha) \in C$, $\beta \geqslant \alpha \Longrightarrow (x, \beta) \in C$.

Определение 4. Функция f, у которой $\operatorname{dom} f \neq \varnothing$ и всюду $f(x) > -\infty$, называется *собственной*, остальные функции называются *несобственными*. Функция $f\colon X \to \overline{\mathbf{R}}$ на линейном пространстве X называется *выпуклой*, если $\operatorname{epi} f$ — выпуклое множество в $X \times \mathbf{R}$. Совокупность всех выпуклых функций на X обозначим $\mathscr{V}^{o}(X)$.

Прямо из этого определения вытекает

Предложение 3. а) *Для выпуклости собственной функции f необходимо и достаточно, чтобы для любых точек $x_i \in \operatorname{dom} f$ и любых $\alpha_i \geqslant 0$, $i = 1, \ldots, n$, таких, что $\sum_{i=1}^{n} \alpha_i = 1$, выполнялось неравенство Иенссена:*

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leqslant \sum_{i=1}^{n} \alpha_i f_i(x_i). \tag{4}$$

б) *Сумма $\sum_{i=1}^{n} f_i(x)$ конечного числа и верхняя грань $\vee_\alpha f_\alpha(x) = \sup_\alpha \{f_\alpha(x)\}$ любого семейства выпуклых функций выпукла.*

Примеры. 1) Выпуклые функции на прямой. Пусть $f_0(x)\colon (\alpha, \beta) \longrightarrow \mathbf{R}$, причем на интервале $(\alpha, \beta) \subseteq \mathbf{R}$ существует производная $f_0'(x)$, которая не убывает. Положим

$$f(x) = \begin{cases} f_0(x), & x \in (\alpha, \beta), \\ +\infty, & x \notin (\alpha, \beta). \end{cases}$$

Для этой функции $\operatorname{dom} f = (\alpha, \beta)$. Проверим, что она выпукла. Произвольная точка отрезка, соединяющего точки $(x_i, z_i) \in \operatorname{epi} f$, $i = 1, 2$, имеет вид $(\alpha x_1 + (1-\alpha) x_2, \alpha z_1 + (1-\alpha) z_2)$, $\alpha \in [0, 1]$. Применим формулу Лагранжа:

$$\begin{aligned} &\alpha z_1 + (1-\alpha) z_2 - f(\alpha x_1 + (1-\alpha) x_2) \geqslant \\ &\quad \geqslant \alpha f(x_1) + (1-\alpha) f(x_2) - f(\alpha x_1 + (1-\alpha) x_2) = \\ &\quad = \alpha f'(c_1)(x_1 - (\alpha x_1 + (1-\alpha) x_2)) + \\ &\quad + (1-\alpha) f'(c_2)(x_2 - (\alpha x_1 + (1-\alpha) x_2)) = \\ &\quad = \alpha(1-\alpha)(f'(c_1) - f'(c_2))(x_1 - x_2) \geqslant 0, \end{aligned}$$

так как $x_1 \leqslant c_1 \leqslant \alpha x_1 + (1-\alpha) x_2 \leqslant c_2 \leqslant x_2$ и производная по условию не убывает. Следовательно, $(\alpha x_1 + (1-\alpha) x_2, \alpha z_1 + (1-\alpha) z_2) \in \operatorname{epi} f$ и f выпукла. ■

Упражнение 6. Пусть выпуклая функция $f\colon \mathbf{R} \longrightarrow \overline{\mathbf{R}}$ конечна на интервале $I=(\alpha, \beta)$. Докажите, что f непрерывна на I и имеет в каждой точке $x \in I$ левую и правую производные $f'_-(x)$, $f'_+(x)$, которые на I не убывают и совпадают всюду, кроме не более чем счетного множества точек.

2) Пусть на выпуклом открытом множестве U нормированного пространства X функция f_0 принадлежит классу C^2 (п. 2.2.5) и ее второй дифференциал $d^2 f_0 = f''_0(x)[\xi, \xi]$ неотрицателен. Как и в предыдущем примере, положим

$$f(x)=\begin{cases} f_0(x), & x \in U, \\ +\infty, & x \notin U. \end{cases}$$

Ограничение этой функции на любую прямую в X обладает свойствами функции, рассмотренной в предыдущем примере (докажите!). Поэтому f выпукла (для проверки выпуклости $\operatorname{epi} f$ достаточно ограничиться всевозможными сечениями этого множества двумерными вертикальными плоскостями в $X \times \mathbf{R}$).

Эти примеры дают нам большой запас конкретных выпуклых функций. Таковыми будут на $\mathbf{R}$ функции e^x, ax^2+bx+c при $a \geqslant 0$, $-\ln x$ (дополненный значением $+\infty$ при $x \leqslant 0$) и т. д. В евклидовом пространстве выпуклой является квадратичная функция $f(x)=(Ax \mid x)+(b \mid x)+c$, где A — положительно определенный симметричный оператор.

3) *Индикаторная функция.* Так называется функция

$$\delta A(x)=\begin{cases} +\infty, & x \notin A, \\ 0, & x \in A. \end{cases}$$

Ясно, что $\delta A(x) \in \mathscr{V}^{\circ}(X)$ тогда и только тогда, когда $A \in \mathfrak{B}(X)$.

4) *Функция Минковского* (см. п. 2.1.3). Неравенство Йенсена (4) является источником многих важных неравенств, используемых в разных разделах математики. В частности, применив его (с $\alpha_i = 1/n$) к функциям

$$f_1(x)=\begin{cases} -\ln x, & x>0, \\ +\infty, & x \leqslant 0, \end{cases}$$

$$f_2(x)=\begin{cases} x \ln x, & x>0, \\ 0, & x=0, \\ +\infty, & x<0, \end{cases}$$

мы получаем в первом случае известное неравенство Коши между средним арифметическим и средним геометрическим, а во втором — важное в теории вероятностей неравенство для энтропии распределения $(p_1, \ldots, p_n)$, где $p_i = x_i \Big/ \sum\limits_{i=1}^{n} x_i$, $x_i \geqslant 0$,

$$H(p) = -\sum_{i=1}^{n} p_i \ln p_i \leqslant \ln n.$$

2.6.2. Выпуклые множества и функции в линейных топологических пространствах. Предположим теперь, что X — локально выпуклое линейное топологическое пространство [КФ, гл. III, § 5]. Как известно, в этом случае сопряженное пространство X^*, состоящее из всех линейных непрерывных функционалов, является достаточно богатым (см. там же, гл. IV, §§ 1—2), и это нам доставляет большой запас выпуклых множеств и функций.

Простейшими выпуклыми множествами в X являются гиперплоскости и полупространства (здесь $x^* \in X^*$)

$H(x^*, \alpha) = \{x \mid \langle x^*, x\rangle = \alpha\}$ — *гиперплоскость*,

$H_+(x^*, \alpha) = \{x \mid \langle x^*, x\rangle \leqslant \alpha\}$, $H_-(x^*, \alpha) = \{x \mid \langle x^*, x\rangle \geqslant \alpha\}$ — *замкнутые полупространства*,

$\mathring{H}_-(x^*, \alpha) = \{x \mid \langle x^*, x\rangle < \alpha\}$, $\mathring{H}_-(x^*, \alpha) = \{x \mid \langle x^*, x\rangle > \alpha\}$ — *открытые полупространства*.

Упражнение 1. Проверьте, что эти множества выпуклы.

Далее, назовем *выпуклым полиэдром* пересечение конечного числа замкнутых полупространств (подчеркнем, что это пересечение не обязано быть ограниченным множеством).

Упражнение 2. Докажите, что в $\mathbf{R}^n$ всякий выпуклый многогранник (п. 2.6.1) является выпуклым полиэдром и что всякий ограниченный выпуклый полиэдр является выпуклым многогранником.

Приведем сводку нужных в дальнейшем топологических свойств выпуклых множеств (по поводу встречающихся топологических терминов см. [КФ]).

Предложение 1. *Пусть A — выпуклое множество. Тогда:*

а) *все точки полуинтервала* $[x_1, x_2)$, *где* $x_1 \in \operatorname{int} A$, $x_2 \in A$, *принадлежат* $\operatorname{int} A$;

б) *внутренность* $\operatorname{int} A$ *и замыкание* $\bar{A}$ *множества* A *выпуклы; если* $\operatorname{int} A \neq \varnothing$, *то* $\bar{A} = \overline{\operatorname{int} A}$.

Доказательство. а) Пусть $V \subset A$ — выпуклая окрестность точки x_1. Произвольная точка $x \in [x_1, x_2)$ имеет вид $x = \alpha x_1 + (1-\alpha) x_2$, $0 < \alpha \leqslant 1$ и $\alpha V + (1-\alpha) x_2$ — ее окрестность, содержащаяся в A, т. е. $x \in \operatorname{int} A$.

Из утверждения а) следует, что $A \subset \overline{\operatorname{int} A}$, если $\operatorname{int} A \neq \varnothing$, откуда $\bar{A} \subset \overline{\operatorname{int} A}$. Обратное включение очевидно.

б) Если $x_i \in \operatorname{int} A$, $i = 1, 2$, то, согласно а), $[x_1, x_2] \in \in \operatorname{int} A$, т. е. $\operatorname{int} A$ выпукла. Пусть теперь $x_i \in \bar{A}$, $i = 1, 2$. Возьмем любую выпуклую окрестность нуля V. По определению замыкания существуют $x'_i \in (x_i + V) \cap A$, $i = 1, 2$. Для произвольной точки $x = \alpha x_1 + (1-\alpha) x_2 \in [x_1, x_2]$ положим $x' = \alpha x'_1 + (1-\alpha) x'_2$. Тогда $x' \in A$ и $x' \in \alpha (x_1 + V) + (1-\alpha)(x_2 + V) = x + V$, т. е. каждая окрестность точки x пересекается с A, откуда $x \in \bar{A}$ и $\bar{A}$ выпукло. ■

Определение 1. Пересечение всех замкнутых выпуклых множеств, содержащих множество A, называется *выпуклым замыканием* множества A и обозначается $\overline{\operatorname{conv}} A$.

Ясно, что $\overline{\operatorname{conv}} A$ выпукло и замкнуто.

Предложение 2. а) $\overline{\operatorname{conv}} A = A \Leftrightarrow$ *множество* A *выпукло и замкнуто*;

б) $\overline{\operatorname{conv}} A$ *совпадает с пересечением всех замкнутых полупространств, содержащих* A;

в) $\overline{\operatorname{conv}} A = \overline{\operatorname{conv}} \bar{A}$;

Доказательство. Первое утверждение непосредственно следует из определения. Включение $\overline{\operatorname{conv}} A \subset \subset \overline{\operatorname{conv}} \bar{A}$ также очевидно (не всякое выпуклое множество, содержащее A, еще и замкнуто), а потому $\overline{\operatorname{conv}} \bar{A} \subset \subset \overline{\operatorname{conv}} A$. С другой стороны, $\overline{\operatorname{conv}} A$, будучи выпуклым (предложение 1б)), замкнутым и содержащим A, должно по определению содержать также и $\overline{\operatorname{conv}} A$, откуда $\overline{\operatorname{conv}} \bar{A} = \overline{\operatorname{conv}} A$.

Далее, обозначим через B пересечение всех замкнутых полупространств, содержащих A. Тогда $\overline{\operatorname{conv}} A \subset B$ (не всякое замкнутое выпуклое множество, содержащее A, является еще и полупространством). Пусть точка $x_0 \notin \notin \overline{\operatorname{conv}} A$. По второй теореме отделимости (п. 2.1.4) существует линейный функционал $x^* \in X^*$, строго разделяющий x_0 и $\overline{\operatorname{conv}} A$

$$\langle x^*, x_0\rangle > \sup_{x \in \overline{\operatorname{conv}}\, A} \langle x^*, x\rangle = \alpha.$$

Отсюда $x_0 \notin H_+(x^*, \alpha) \supset \overline{\operatorname{conv}}\, A \supset A$, и потому $x_0 \notin B$. Следовательно, $\overline{\operatorname{conv}}\, A = B$. ■

Упражнение 3. В гильбертовом пространстве l_2 рассмотрим бесконечномерный эллипсоид $B = \left\{ x = (x_1, x_2, \ldots) \,\middle|\, \sum_{n=1}^{\infty} (x_n/a_n)^2 \leqslant 1 \right\}$ с полуосями a_n.

а) Докажите, что множество B выпукло и замкнуто.

б) Найдите условия на a_n, при которых $\operatorname{int} B \neq \varnothing$.

в) Пусть $\operatorname{int} B \neq \varnothing$ и точка $y = (y_1, y_2, \ldots)$ лежит на границе эллипсоида, т. е. $\sum_{n=1}^{\infty} (y_n/a_n)^2 = 1$. Докажите, что через y можно провести гиперплоскость так, что B будет лежать по одну ее сторону.

г) Всегда ли верно предыдущее утверждение, если $\operatorname{int} B = \varnothing$?

Подобно тому как простейшими выпуклыми множествами являются полупространства, простейшими выпуклыми функциями будут *аффинные функции*

$$a(x) = \langle x^*, x\rangle - b,\ x^* \in X, \quad b \in \mathbf{R}.$$

Упражнение 4. Докажите, что $a(\cdot)$ выпукла; найдите $\operatorname{epi} a$.

Предложения 1 и 2 имеют свои аналоги. К их формулировке мы и перейдем.

Определение 2. Пусть $f\colon X \to \mathbf{R}$. Функция $\bar f$, определяемая условием $\operatorname{epi} \bar f = \overline{\operatorname{epi} f}$, называется *замыканием* функции f; если $f = \bar f$, то функция называется *замкнутой*. Функция $\overline{\operatorname{conv}} f$, определяемая условием $\operatorname{epi}(\overline{\operatorname{conv}} f) = \overline{\operatorname{conv}}(\operatorname{epi} f)$, называется *выпуклым замыканием* f.

Упражнение 5. Проверьте, что $\overline{\operatorname{epi} f}$ и $\overline{\operatorname{conv}} f$ суть надграфики некоторых функций (см. упражнение 5 п. 2.6.1).

Ясно, что $\overline{\operatorname{conv}} f = f \Leftrightarrow f$ выпукла и замкнута.

Напомним также, что функция $f\colon X \to \overline{\mathbf{R}}$ называется *полунепрерывной снизу в точке* x_0, если $\varliminf_{x \to x_0} f(x) \geqslant f(x_0)$, и просто *полунепрерывной снизу*, если то же самое верно для любого x_0.

Упражнение 6. Докажите, что $f\colon X \longrightarrow \overline{\mathbf{R}}$ полунепрерывна снизу тогда и только тогда, когда для любого $c \in \mathbf{R}$ множество $\mathscr{L}_c f = \{x \mid f(x) \leqslant c\}$ замкнуто в X.

Предложение 3. а) *Для того чтобы собственная функция была замкнутой, необходимо и достаточно, чтобы она была полунепрерывна снизу.*

б) *Для того чтобы выпуклая функция f была непрерывной на $\operatorname{int}\operatorname{dom} f$ достаточно, чтобы f была ограничена в окрестности U некоторой точки $\bar{x}$ и конечна в точке $\bar{x}$.*

При этом f собственная, $\operatorname{int}\operatorname{dom} f \neq \varnothing$,

$$\operatorname{int}\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbf{R} \mid x \in \operatorname{int}\operatorname{dom} f,\quad \alpha > f(x)\} \neq \varnothing$$

и

$$f(x) = (\overline{\operatorname{conv}} f)(x), \quad \forall x \in \operatorname{int}\operatorname{dom} f.$$

В соответствии с упражнением 6 будем проверять замкнутость множеств $\mathscr{L}_c f$.

Доказательство. А) Пусть f замкнутая функция. Тогда $\operatorname{epi} f$ — замкнутое множество в $X \times \mathbf{R}$. Гиперплоскость $H = \{(x, \alpha \mid \alpha = c\}$ также замкнута. Значит, замкнуто и множество $\mathscr{L}_c f = \{x \mid f(x) \leqslant c\}$, ибо $\operatorname{epi} f \cap H = \{(x, c) \mid x \in \mathscr{L}_c f\}$ и отображение $x \longmapsto (x, c)$ является гомеоморфизмом.

Пусть наоборот, все $\mathscr{L}_c f$ замкнуты и $(x_0, \alpha_0) \notin \operatorname{epi} f$. Тогда $f(x_0) > \alpha_0$, а следовательно, $f(x_0) > \alpha_0 + \varepsilon$ для некоторого $\varepsilon > 0$ и $x_0 \notin \mathscr{L}_{\alpha_0 + \varepsilon} f$. Ввиду замкнутости $\mathscr{L}_{\alpha_0 + \varepsilon} f$ существует окрестность $V \ni x_0$ такая, что $V \cap \mathscr{L}_{\alpha_0 + \varepsilon} f = \varnothing$, т. е. $f(x) > \alpha_0 + \varepsilon$, $x \in V$. Открытое множество $\{(x, \alpha) \mid x \in V, \alpha < \alpha_0 + \varepsilon\}$ содержит (x_0, α_0) и не пересекается с $\operatorname{epi} f$. Следовательно, дополнение к $\operatorname{epi} f$ открыто, а сам $\operatorname{epi} f$ замкнут.

Б) Лемма. *Пусть X — локально выпуклое пространство, V — выпуклая окрестность точки $\bar{x}$ в X, f — выпуклая функция на X, принимающая в точке $\bar{x} \in V$ конечное значение ($|f(\bar{x})| < \infty$) и ограниченная в V сверху. Тогда f непрерывна в точке $\bar{x}$.*

Доказательство леммы. Произведя сдвиг $x \longmapsto x + \bar{x}$ и вычитая из f константу $f(\bar{x})$, сведем дело к случаю, когда $\bar{x} = 0$, $0 \in V$, $f(0) = 0$, $\sup\limits_{x \in V} f(x) \leqslant C$. При этом V можно считать симметричной окрестностью нуля (иначе мы рассмотрели бы $W=V\cap(-V)$). Пусть $0<\alpha<1$ и $x\in\alpha V$. Тогда $x/\alpha\in V$ и из-за выпуклости f мы имеем

$$f(x)=f((1-\alpha)\,0+\alpha x/\alpha)\leqslant(1-\alpha)\,f(0)+\alpha f(x/\alpha)\leqslant\alpha C.$$

С другой стороны, $-x/\alpha\in V$ ввиду симметрии V и, используя равенство $0=x/(1+\alpha)+\alpha/(1+\alpha)(-x/\alpha)$ и выпуклость f, получаем

$$0=f(0)=f(x/(1+\alpha)+\alpha/(1+\alpha)(-x/\alpha))\leqslant$$
$$\leqslant 1/(1+\alpha)\,f(x)+\alpha/(1+\alpha)\,f(-x/\alpha),$$

а, значит, $f(x)\geqslant-\alpha f(-x/\alpha)\geqslant-\alpha C$.

Итак, если $x\in\alpha V$, то $|f(x)|\leqslant\alpha C$, т. е. f непрерывна в нуле. ■

Следствие. *В условиях леммы* $\operatorname{int}\operatorname{dom}f\neq\varnothing$.

В) Возвращаемся к доказательству пункта б) предложения 3. По лемме и следствию из нее f непрерывна в $\bar{x}$ и $\operatorname{int}\operatorname{dom}f\neq\varnothing$.

Предположим, что $f(y)=-\infty$ для некоторого y, т. е. для любого $\alpha\in\mathbf{R}$ точка $(y,\alpha)\in\operatorname{epi}f$. Поскольку $\operatorname{epi}f$ выпукло,

$$((1-\lambda)\,\bar{x}+\lambda y,\ (1-\lambda)\,f(\bar{x})+\lambda\alpha)\in\operatorname{epi}f,$$

для любого $\alpha\in\mathbf{R}$, откуда $f((1-\lambda)\,\bar{x}+\lambda y)=-\infty$, что при $\lambda\downarrow 0$ противоречит непрерывности f в $\bar{x}$. Таким образом, $f(y)>-\infty$ всюду, т. е. f является собственной.

Пусть теперь $y\in\operatorname{int}\operatorname{dom}f$. Найдем $\rho>1$ такое, что $z=\bar{x}+\rho\,(y-\bar{x})\in\operatorname{int}\operatorname{dom}f$ (это можно сделать, ибо пересечение открытого множества $\operatorname{int}\operatorname{dom}f$ с прямой, проходящей через $\bar{x}$ и y, есть открытый интервал этой прямой). Гомотетия G с центром в z и коэффициентом $(\rho-1)/\rho$ переводит $\bar{x}$ в y и окрестность V в окрестность V' точки y. При этом, если $\zeta\in G(V)=V'$, то $\zeta=(\rho-1)/\rho\tilde{x}+z/\rho$, $\tilde{x}\in V$ и

$$f(\zeta)\leqslant\frac{\rho-1}{\rho}f(\tilde{x})+\frac{1}{\rho}f(z)\leqslant\frac{\rho-1}{\rho}C+\frac{1}{\rho}f(z),$$

следовательно, f непрерывна в точке y по лемме, т. е. f непрерывна на $\operatorname{int}\operatorname{dom}f$.

Если $(x_0,\alpha_0)\in\operatorname{int}\operatorname{epi}f$, то по определению найдутся окрестность W точки x_0 и $\varepsilon>0$ такие, что

$$\{(x,\alpha)\,|\,x\in W,\ |\alpha-\alpha_0|<\varepsilon\}\subset\operatorname{epi}f,$$

откуда $x_0 \in \operatorname{int}\operatorname{dom} f$ и $\alpha_0 > f(x_0)$. Обратное включение:

$$\{(x, \alpha) \mid x \in \operatorname{int}\operatorname{dom} f,\ \alpha > f(x)\} \subset \operatorname{int}\operatorname{epi} f$$

очевидно; в частности, $\operatorname{int}\operatorname{epi} f \neq \varnothing$.

Поскольку f выпукла, $\operatorname{conv}(\operatorname{epi} f) = \operatorname{epi} f$. Вспоминая определение 2 и предложение 2в), имеем

$$\operatorname{epi}(\overline{\operatorname{conv}} f) = \overline{\operatorname{conv}}(\operatorname{epi} f) = \overline{\operatorname{conv}(\operatorname{epi} f)} = \overline{\operatorname{epi}(f)} \supset \operatorname{epi}(f).$$

Следовательно, всегда $(\overline{\operatorname{conv}} f)(x) \leqslant f(x)$. Если же f непрерывна в точке x (или хотя бы полунепрерывна снизу) и $f(x) > \alpha$, то это же неравенство сохранится и в целой окрестности точки x, а потому $(x, \alpha) \notin \overline{\operatorname{epi} f} = \operatorname{epi}(\overline{\operatorname{conv}} f)$ и $(\overline{\operatorname{conv}} f)(x) > \alpha$. Поэтому $(\overline{\operatorname{conv}} f)(x) = f(x)$. В частности, в условиях пункта б) доказываемого утверждения это равенство верно для $x \in \operatorname{int}(\operatorname{dom} f)$. ■

Упражнение 7. Пусть f выпукла на X и $f(\bar{x}) = -\infty$ в некоторой точке $\bar{x} \in \operatorname{int}\operatorname{dom} f$. Тогда $f(x) \equiv -\infty$, $\forall x \in \operatorname{int}\operatorname{dom} f$.

Определение 3. Аффинная функция $a(x) = \langle x^*, x\rangle - b$ называется *опорной* для функции f, если:

а) $a(x) \leqslant f(x)$ для всех x;

б) для всякого $\varepsilon > 0$ найдется такое x, что $a(x) > f(x) - \varepsilon$. Другими словами,

$$b = \sup_x \{\langle x^*, x\rangle - f(x)\}. \tag{1}$$

Теорема Минковского. *Собственная функция f выпукла и замкнута тогда и только тогда, когда она является верхней гранью множества всех своих опорных аффинных функций.*

Доказательство. 1) «Тогда». Аффинная функция выпукла и замкнута, так как ее надграфик — замкнутое полупространство. Далее, для любого семейства функций $\{f_\alpha\}$ имеем

$$\operatorname{epi}\left\{\sup_\alpha f_\alpha\right\} = \left\{(x, z) \mid z \geqslant \sup_\alpha f_\alpha(x)\right\} = \\ = \bigcap_\alpha \{(x, z) \mid z \geqslant f_\alpha(x)\} = \bigcap_\alpha \operatorname{epi} f_\alpha, \tag{2}$$

и если все f_α выпуклы и замкнуты, то $\operatorname{epi} f_\alpha$ — выпуклые замкнутые множества и их пересечение обладает тем же свойством.

2) «Только тогда». По условию $B = \operatorname{epi} f$ — замкнутое выпуклое непустое множество в $X \times \mathbf{R}$. Согласно

предложению 2 B является пересечением всех содержащих его замкнутых полупространств. Так как всякий линейный непрерывный функционал на $X \times \mathbf{R}$ имеет вид $\langle x^*, x\rangle + \lambda z$, $x^* \in X^*$, $\lambda \in \mathbf{R}$ (см. п. 2.1.2), замкнутое полупространство определяется неравенством

$$\langle x^*, x\rangle + \lambda z \leqslant b. \tag{3}$$

Поскольку $B = \operatorname{epi} f$ непусто и вместе со всякой своей точкой (x_0, z_0) содержит все точки (x_0, z), $z > z_0$, B может содержаться в полупространстве (3), только если $\lambda \leqslant 0$. При $\lambda = 0$ полупространство (3) будем называть *вертикальным*. Очевидно, что

$$\begin{aligned}\operatorname{epi} f = \{(x, z) \mid f(x) \leqslant z\} \subset \{(x, z) \mid \langle x^*, x\rangle \leqslant b\} \Leftrightarrow \\ \Leftrightarrow \operatorname{dom} f \subset \{x \mid \langle x^*, x\rangle \leqslant b\} = H_+(x^*, b) \Rightarrow \\ \Rightarrow \operatorname{epi} f \subset H_+(x^*, b) \times \mathbf{R}.\end{aligned}$$

При $\lambda < 0$ все члены неравенства (3) можно разделить на $|\lambda|$, так что можно считать $\lambda = -1$ и полупространство (3) совпадает с надграфиком аффинной функции $a(x) = \langle x^*, x\rangle - b$. Но

$$B = \operatorname{epi} f \subset \operatorname{epi} a \Leftrightarrow f(x) \geqslant a(x),\ \forall x.$$

Поэтому

$$\operatorname{epi} f = \bigcap_{a \leqslant f} \operatorname{epi} a \cap \bigcap_{\operatorname{dom} f \subset H_+(x^*, b)} [H_+(x^*, b) \times \mathbf{R}]. \tag{4}$$

Теперь заметим, что хотя бы одна аффинная $a_0 \leqslant f$ существует (в противном случае $\operatorname{epi} f$ вместе с точкой (x_0, z_0) содержало бы все точки (x_0, z), $z \in \mathbf{R}$, откуда $f(x_0) = -\infty$, вопреки тому, что f — собственная). Но

$$\begin{aligned}\operatorname{epi} a_0 \cap H_+(x^*, b) \times \mathbf{R} = \{(x, z) \mid a_0(x) \leqslant z,\ \langle x^*, x\rangle \leqslant b\} = \\ = \bigcap_{\lambda \geqslant 0} \{(x, z) \mid a_0(x) + \lambda(\langle x^*, x\rangle - b) \leqslant z\} = \\ = \bigcap_{\lambda \geqslant 0} \operatorname{epi}\{a_0(x) + \lambda(\langle x^*, x\rangle - b)\},\end{aligned}$$

и если $a_0(x) \leqslant f(x)$ и $\operatorname{dom} f \subset H_+(x^*, b)$, то

$$f(x) < +\infty \Rightarrow \langle x^*, x\rangle - b \leqslant 0 \Rightarrow a_0(x) + \lambda(\langle x^*, x\rangle - b) \leqslant f(x).$$

Следовательно, $\operatorname{epi} a_0 \cap [H_+(x^*, b) \times \mathbf{R}] \subset \bigcap_{a \leqslant f} \operatorname{epi} a$, так что в (4) мы можем избавиться от всех вертикальных полупространств $H_+(x^*, b)$. Итак,

$$\operatorname{epi} f = \bigcap_{a \leqslant f} \operatorname{epi} a$$

и, согласно (2),

$$f(x) = \sup \{a(x) \mid a(x) \text{ аффинная и } \leqslant f(x)\}. \qquad (5)$$

Остается заметить, что в (5) можно брать только опорные к f функции, поскольку верхняя грань не изменится, если из всех функций $a(x) = \langle x^*, x\rangle - b \leqslant f(x)$ с одним и тем же x^* мы оставим ту, у которой b определяется равенством (1), т. е. опорную. ■

Упражнение 8. Если $f\colon X \to \overline{\mathbf{R}}$ выпукла и замкнута и $f(\overline{x}) = -\infty$, то $f(x) \equiv -\infty$, $\forall x \in \operatorname{dom} f$.

Для гладких функций мы можем воспользоваться следующим утверждением.

Предложение 4. *Если функция f выпукла и дифференцируема по Гато в точке x_0, то функция $a(x) = \langle f'_\Gamma(x_0), x - x_0\rangle + f(x_0)$ является для нее опорной.*

Доказательство. Условие б) определения 3 следует из равенства $a(x_0) = f(x_0)$. Предположим, что $f(x_1) < a(x_1)$ для некоторого x_1. Тогда $f(x_1) < a(x_1) - \varepsilon$ для некоторого $\varepsilon > 0$, и потому при $0 < \alpha < 1$

$$f(x_0 + \alpha(x_1 - x_0)) = f((1-\alpha)x_0 + \alpha x_1) \leqslant$$
$$\leqslant (1-\alpha) f(x_0) + \alpha f(x_1) < (1-\alpha) f(x_0) + \alpha(a(x_1) - \varepsilon) =$$
$$= (1-\alpha) f(x_0) + \alpha(f(x_0) + \langle f'_\Gamma(x_0), x_1 - x_0\rangle - \varepsilon) =$$
$$= f(x_0) + \alpha\langle f'_\Gamma(x_0), x_1 - x_0\rangle - \alpha\varepsilon.$$

Следовательно,

$$\langle f'_\Gamma(x_0), x_1 - x_0\rangle = \lim_{\alpha \downarrow 0} \frac{f(x_0 + \alpha(x_1 - x_0)) - f(x_0)}{\alpha} \leqslant$$
$$\leqslant \langle f'_\Gamma(x_0), x_1 - x_0\rangle - \varepsilon,$$

что противоречиво. Тем самым выполнено и условие а). ■

Пусть теперь функция $f\colon X \to \mathbf{R}$ дифференцируема по Гато в каждой точке $x_0 \in X$. Составим для нее функцию Вейерштрасса (ср. п. 1.4.4):

$$\mathscr{E}(x, x_0) = f(x) - f(x_0) - \langle f'_\Gamma(x_0), x - x_0\rangle.$$

Предложение 5. *Для того чтобы f была выпуклой, необходимо и достаточно, чтобы выполнялось неравенство $\mathscr{E}(x, x_0) \geqslant 0$, $\forall x, x_0$.*

Доказательство. а) Если f выпукла, то $\mathscr{E}(x, x_0) \geqslant 0$, $\forall x, x_0$ в силу только что доказанного предложения 4.

б) Пусть $x_0=\alpha x_1+(1-\alpha)x_2$, $0\leqslant\alpha\leqslant 1$. Неравенства $\mathscr{E}(x_1, x_0)\geqslant 0$, $\mathscr{E}(x_2, x_0)\geqslant 0$ означают, что точки $(x_1, f(x_1))$, $(x_2, f(x_2))$ лежат в полупространстве $\{(x, z)\,|\,z\geqslant f(x_0)+$ $+\langle f'_{\Gamma}(x_0), x-x_0\rangle\}$. Следовательно, и отрезок, их соединяющий, лежит в том же полупространстве, а потому

$$\alpha f(x_1)+(1-\alpha)f(x_2)\geqslant$$
$$\geqslant f(x_0)+\langle f'(x_0), \alpha x_1+(1-\alpha)x_2-x_0\rangle=f(x_0). \blacksquare$$

Таким образом, выполнение условия Вейерштрасса из п. 1.4.4 для всех точек $(t, x, \dot{x})$ эквивалентно выпуклости лагранжиана $L(t, x, \dot{x})$ по $\dot{x}$.

2.6.3. Преобразование Лежандра — Юнга — Фенхеля. Теорема Фенхеля — Моро. По-прежнему пусть X — локально выпуклое линейное топологическое пространство, а X^* — его сопряженное. Зададимся теперь произвольной функцией $f\colon X\to\overline{\mathbb{R}}$ и исследуем подробнее множество ее опорных функций. Естественно считать при этом функцию собственной, так как в соответствии с определением $f(x)\geqslant a(x)>-\infty$ всюду, если f имеет хотя бы одну опорную, а случай $f(x)\equiv+\infty$ тривиален: любая аффинная функция является опорной.

Определение. Функция на X^*, определяемая равенством

$$f^*(p)=\sup_{x\in X}\{\langle p, x\rangle-f(x)\} \qquad (1)$$

называется *сопряженной к f функцией* или ее *преобразованием Юнга — Фенхеля* (иногда *преобразованием Лежандра*), а функция

$$f^{**}(x)=\sup_{p\in X^*}\{\langle p, x\rangle-f^*(p)\} \qquad (2)$$

— *второй сопряженной* к f функцией.

Из равенства (1) п. 2.6.2 видно, что множество опорных для f функций находится во взаимно однозначном соответствии с множеством тех $p\in X^*$, для которых $f^*(p)$ конечно:

$$(p, f^*(p)\neq\pm\infty)\leftrightarrow a_p(x)=\langle p, x\rangle-f^*(p). \qquad (3)$$

Замечание. Несмотря на кажущуюся симметрию формул (1) и (2), $f^{**}\neq(f^*)^*$. Дело в том, что $(f^*)^*$ по определению следует рассматривать на $(X^*)^*$, а не на X. Только для рефлексивных пространств $(f^*)^*$ и f^{**} естественно отождествляются.

Предложение 1. 1) *Функции f^* и f^{**} выпуклы и замкнуты.*

2) *Для любых $x \in X$ и $p \in X^*$ выполняется неравенство Юнга:*

$$f(x) + f^*(p) \geqslant \langle p, x \rangle. \tag{4}$$

3) $f^{**}(x) \leqslant f(x)$, $\forall x$.

4) *Если $f(x) \leqslant g(x)$, то $f^*(p) \geqslant g^*(p)$ и $f^{**}(x) \leqslant g^{**}(x)$.*

Доказательство. 1) Из (1) и (2) видно, что обе функции получаются как верхняя грань некоторого множества аффинных функций, а потому их надграфики являются (по формуле (2) п. 2.6.2) пересечением замкнутых полупространств и, значит, выпуклы и замкнуты.

2) Из (1) вытекает, что при любых $x \in X$ и $p \in X^*$ должно быть $f^*(p) \geqslant \langle p, x \rangle - f(x)$, что равносильно (4).

3) По уже доказанному $f(x) \geqslant \langle p, x \rangle - f^*(p)$, откуда

$$f(x) \geqslant \sup_p \{\langle p, x \rangle - f^*(p)\} = f^{**}(x).$$

4) Очевидное следствие определений. ■

Примеры. 1) Пусть $f(x) \doteq \langle x_0^*, x \rangle - b$ — аффинная функция. Такая функция имеет только одну опорную — себя самое. Поэтому

$$f^*(p) = \sup_x \{\langle p, x \rangle - \langle x_0^*, x \rangle + b\} = \begin{cases} +\infty, & p \neq x_0^*, \\ b, & p = x_0^*, \end{cases} \tag{5}$$

и

$$f^{**}(x) = \sup_p \{\langle p, x \rangle - f^*(p)\} = \langle x_0^*, x \rangle - b = f(x). \tag{6}$$

2) Пусть, как и в примере 1 п. 2.6.1,

$$f(x) = \begin{cases} f_0(x), & \alpha < x < \beta, \\ +\infty, & x \notin (\alpha, \beta), \end{cases}$$

причем $f_0'(x)$ на (α, β) непрерывна и монотонно возрастает. Обозначим

$$A = \lim_{x \to \alpha + 0} f_0'(x), \quad B = \lim_{x \to \beta - 0} f_0'(x).$$

Для любого $p \in (A, B)$ равенство $p = f_0'(x_0)$ выполняется при некотором $x_0 \in (\alpha, \beta)$. Так как $f(x)$ — выпуклая функция, то (предложение 4 п. 2.6.2)

$$a(x) = f(x_0) + f'(x_0)(x - x_0) = px - (px_0 - f(x_0))$$

— опорная функция, и, значит, равенства

$$f^*(p)=px_0-f_0(x_0), \quad p=f_0'(x_0) \tag{7}$$

задают $f^*(p)$ параметрически в (A, B) (функция f^*, определяемая равенствами (7), называется в классическом анализе *преобразованием Лежандра* f_0).

Упражнения.

1. Что можно сказать в этом примере о $f^*(p)$ при $p \notin (A, B)$?

2. Для следующих функций f на $\mathbf{R}$ указать $\operatorname{dom} f$, $\operatorname{epi} f$, проверить выпуклость и замкнутость и вычислить f^* и f^{**}:

а) ax^2+bx+c, $a \geqslant 0$, б) $|x|+|x-a|$,

в) $||x|-1|$, г) $\delta[\alpha, \beta](x)=\begin{cases} 0, & \alpha \leqslant x \leqslant \beta, \\ \infty, & x \notin [\alpha, \beta], \end{cases}$

д) $\begin{cases} -\sqrt{1+x^2}, & |x| \leqslant 1, \\ +\infty, & |x|>1, \end{cases}$ е) $\begin{cases} +\sqrt{1-x^2}, & |x|<1, \\ +\infty, & |x| \geqslant 1, \end{cases}$

ж) e^x, з) $\begin{cases} -\ln x, & x>0, \\ 0, & x \leqslant 0, \end{cases}$

и) $\begin{cases} x \ln x, & x>0, \\ 0, & x=0, \\ +\infty, & x<0, \end{cases}$ к) $\begin{cases} x \ln x, & x>0, \\ +\infty, & x \leqslant 0, \end{cases}$

л) $\begin{cases} \dfrac{x^a}{a}, & x \geqslant 0, \\ +\infty, & x<0 \end{cases} \quad (a \geqslant 1)$, м) $\begin{cases} \dfrac{x^a}{a}, & x \geqslant 0, \\ 0, & x<0 \end{cases} \quad (a \geqslant 1)$,

н) $|x|^a/a \quad (a \geqslant 1)$.

3) Следующий пример мы используем далее в § 3.3.

Предложение 2. *Пусть функция* $f\colon \mathbf{R}^s \to \mathbf{R}$ *определена равенством* $f(x)=f(x_1, \ldots, x_s)=\max\{x_1, \ldots, x_s\}$. *Тогда*

$$f^*(p)=f^*(p_1, \ldots, p_s)=\begin{cases} 0, \text{ если } p_i \geqslant 0, \ \sum_{i=1}^{s} p_i=1, \\ +\infty \text{ в остальных случаях.} \end{cases}$$

Доказательство. Поскольку f является максимумом конечного числа линейных (а значит, выпуклых и непрерывных) функций, ее выпуклость и непрерывность очевидны.

Если $p_i \geqslant 0$, $\sum_{i=1}^{s} p_i=1$, то $\langle p, x\rangle \leqslant \max\{x_1, \ldots, x_s\}$, и потому

$$f^*(p)=\sup_x \{\langle p, x\rangle-\max\{x_1, \ldots, x_s\}\}=0.$$

С другой стороны, если $f^*(p) < +\infty$, то обязательно $f^*(p) = 0$ (в силу однородности функций $x \mapsto \langle p, x\rangle$ и $x \mapsto \max\{x, \ldots, x_s\}$), т. е. для таких p выполняется неравенство

$$\sum_{i=1}^{s} p_i x_i - \max\{x_1, \ldots, x_s\} \leqslant 0, \ \forall x.$$

Подставляя в это неравенство $x_j = 0$, $j \neq i$, $x_i = \xi$, получаем $p_i \xi \leqslant \max\{\xi, 0\}$, откуда $p_i \geqslant 0$. Положив в том же неравенстве $x_i = a$, $i = 1, \ldots, s$, получаем

$$a \sum_{i=1}^{s} p_i \leqslant a, \ \forall a \in \mathbf{R} \Leftrightarrow \sum_{i=1}^{s} p_i = 1. \ \blacksquare$$

4) Сопряженная функция к норме в нормированном пространстве.

Предложение 3. *Пусть X — нормированное пространство, $N(x) = \|x\|$. Тогда N^* совпадает с индикаторной функцией δB^* единичного шара сопряженного пространства X^*, а $N^{**}(x) = N(x)$.*

Доказательство. Если $\|p\| > 1$, то $\langle p, x_0\rangle > \|x_0\|$ для некоторого x_0, и тогда

$$f^*(p) = \sup_x \{\langle p, x\rangle - \|x\|\} \geqslant \sup_{\alpha \geqslant 0} \{\langle p, \alpha x_0\rangle - \alpha \|x_0\|\} = +\infty.$$

Если же $\|p\| \leqslant 1$, то $\langle p, x\rangle \leqslant \|x\|$ и

$$f^*(p) = \sup_x (\langle p, x\rangle - \|x\|) = 0.$$

Далее,

$$f^{**}(x) = \sup_p \{\langle p, x\rangle - f^*(p)\} = \sup_{\|p\| \leqslant 1} \{\langle p, x\rangle\} = \|x\| = N(x). \ \blacksquare$$

5) Пусть $X = \mathbf{R}$, $f(x) = 1/(1 + x^2)$. Тогда, как легко понять,

$$f^*(p) = \begin{cases} 0, & p = 0, \\ +\infty, & p \neq 0, \end{cases} \quad f^{**}(x) \equiv 0,$$

и, таким образом, равенство $f^{**}(x) = f(x)$, которое мы наблюдали в примерах 1) и 4), здесь не выполняется.

Следующая теорема, которая принадлежит к числу важнейших в выпуклом анализе, показывает, что совпадение f и f^{**} отнюдь не является случайным.

Теорема Фенхеля—Моро. *Пусть X — локально выпуклое топологическое линейное пространство, $f\colon X \to \overline{\mathbf{R}}$ — функция, всюду бо́льшая $-\infty$. Тогда* а) $f^{**}(x) \equiv f(x)$ *тогда и только тогда, когда* f *выпукла и замкнута.*

б) $f^{**}(x) = \sup\{a(x) \mid a(x)$ — *аффинная и* $\leqslant f(x)\}$.

в) *Если существует хотя бы одна аффинная функция* $a(x) \leqslant f(x)$ (*эквивалентные условия:* $f^*(p) \not\equiv +\infty$ *или* $f^{**}(x) > -\infty$ *всюду*), *то* $f^{**}(x)$ — *наибольшая из замкнутых выпуклых функций, не превосходящих* $f(x)$, *т. е.* $f^{**} = \overline{\operatorname{conv}} f$.

г) $(f^{**})^* = f^*$.

Доказательство. А) «Только тогда» следует из 1) предложения 1. Предположим теперь, что $a(x) \leqslant f(x)$ — аффинииая функцня. В примере 1) мы установили, что $a^{**}(x) = a(x)$. Воспользовавшись сиова предложеиием 1, имеем

$$a(x) = a^{**}(x) \leqslant f^{**}(x) \leqslant f(x),$$

а потому

$$\sup\{a(x) \mid a(x) \text{ — аффинная и } \leqslant f(x)\} \leqslant f^{**}(x) \leqslant f(x). \quad (8)$$

Отсюда прежде всего следует «тогда» в а), ибо когда f выпукла и замкиута, левая и правая части в (8) совпадают (для собственной f по теореме Минковского, а если $f \equiv +\infty$, то и левая часть равиа $+\infty$). Тем самым а) доказаио.

Б) Для произвольиой f возможиы три случая. Если аффииной фуикции $a(x) \leqslant f(x)$ не существует, то $f^*(p) = \sup\limits_x \{\langle p, x\rangle - f(x)\} = +\infty$ при всех p и, значит, $f^{**}(x) = -\infty$. Но в этом случае и левая часть в (8) равна $-\infty$ ($\sup \varnothing = -\infty$ по определению), т. е. б) имеет место.

Если же существует аффинная $a(x) \leqslant f(x)$, то из (8) вытекает, что $f^{**}(x) > -\infty$ всюду. Остаются две возможности. Если $f^{**}(x) \equiv +\infty$, то $f(x) \equiv +\infty$ и обе функции равны левой части в (8), так как аффииная фуикция здесь может быть любой. Если же f^{**} — собственная, то по теореме Минковского

$$f^{**}(x) = \sup\{a(x) \mid a(x) \text{ опорная для } f^{**}(x)\}.$$

Сопоставляя с (8), иаходим $\sup\{a(x) \mid a(x)$ аффинна и $\leqslant f(x)\} \leqslant f^{**}(x) = \sup\{a(x) \mid a(x)$ опорная для $f^{**}(x)\} \leqslant \sup\{a(x) \mid a(x)$ аффиииая и $\leqslant f(x)\}$. Таким образом, и в этих случаях б) имеет место.

В) Пусть далее существует аффинная фуикция $a(x) \leqslant f(x)$ и $g(x) \leqslant f(x)$ — произвольная замкиутая выпуклая функция. Положим

$$h(x)=\sup\{g(x),\ f^{**}(x)\}.$$

Эта функция замкнута, выпукла и всюду $f(x)\geqslant h(x)>>-\infty$. Следовательно,

$$h(x)=h^{**}(x)\leqslant f^{**}(x)\leqslant h(x),$$

а потому $h(x)=f^{**}(x)$ и $g(x)\leqslant f^{**}(x)$, чем доказано в).

Г) Поскольку $f^{**}(x)\leqslant f(x)$, то $(f^{**})^*(p)\geqslant f^*(p)$. С другой стороны, в силу неравенства Юнга $f(x)\geqslant a(x)= =\langle p,\ x\rangle-f^*(p)$, откуда при конечном $f^*(p)$в силу (8) $f^{**}(x)\geqslant a(x)$ и $(f^{**})^*(p)=$

$$=\sup_x(\langle p,\ x\rangle-f^{**}(x))\leqslant\sup_x(\langle p,\ x\rangle-a(x))=f^*(p).$$

Случай $f^*(p)=+\infty$ тривиален, а при $f^*(p)=-\infty$ получаем

$$f(x)=f^{**}(x)\equiv+\infty\ \text{ и }\ (f^{**})^*(p)=-\infty=f^*(p). \qquad \blacksquare$$

Упражнение 3. Пусть f выпукла и ограничена сверху в окрестности некоторой точки. Тогда $f^{**}(x)=f(x)$, $\forall x\in\operatorname{int}\operatorname{dom}f$.

2.6.4. Субдифференциал. Теорема Моро — Рокафеллара. Теорема Дубовицкого — Милютина. Пусть X — локально-выпуклое линейное топологическое пространство, f — функция на X; $f\colon X\to\overline{\mathbb{R}}$.

Определение. *Субдифференциалом* f в точке x_0 называется подмножество X^*, состоящее из таких элементов $x^*\in X^*$, для которых выполнено неравенство

$$f(x)-f(x_0)\geqslant\langle x^*,\ x-x_0\rangle,\quad\forall x\in X. \tag{1}$$

Субдифференциал функции f в точке x_0 обозначается $\partial f(x_0)$.

Таким образом, $\partial f(x_0)$ — это множество «угловых коэффициентов» аффинных функций $a(x)=\langle x^*,\ x\rangle-b$, *опорных к f в точке* x_0, т. е. таких, что в формуле (1) п. 2.6.2 верхняя грань достигается в x_0:

$$b=\langle x^*,\ x_0\rangle-f(x_0). \tag{2}$$

Предложение 1. *Субдифференциал $\partial f(x_0)$ является выпуклым (возможно пустым) множеством в X^*.*

Доказательство. Пусть $x_i^*\in\partial f(x_0)$, $i=1, 2$. Тогда по определению $f(x)-f(x_0)\geqslant\langle x_1^*,\ x-x_0\rangle$, $f(x)-f(x_0)\geqslant\geqslant\langle x_2^*,\ x-x_0\rangle$. Умножив первое неравенство на α, второе — на $(1-\alpha)$, где $\alpha\in[0,\ 1]$, и складывая, получаем

$$f(x)-f(x_0)\geqslant\langle\alpha x_1^*+(1-\alpha)x_2^*,\ x-x_0\rangle,\quad\forall x.$$

Примеры. 1) Субдифференциалы некоторых функций на прямой и плоскости. Предоставим читателю проверить, что:

а) $f_1(x)=|x|, \quad (x\in\mathbf{R})\Rightarrow \partial f_1(x)=\begin{cases}[-1,\ 1], & x=0,\\ \operatorname{sign} x, & x\neq 0.\end{cases}$

б) $f_2(x)=\sqrt{x_1^2+x_2^2}, \quad (x\in\mathbf{R}^2)\Rightarrow$

$$\Rightarrow \partial f_2(x)=\begin{cases}\{y\mid \sqrt{y_1^2+y_2^2}\leqslant 1\}, & x=0,\\ \dfrac{x}{|x|}=\left(\dfrac{x_1}{\sqrt{x_1^2+x_2^2}},\ \dfrac{x_2}{\sqrt{x_1^2+x_2^2}}\right), & x\neq 0.\end{cases}$$

в) $f_3(x)=\max(|x_1|,\ |x_2|), \quad (x\in\mathbf{R}^2)\Rightarrow$

$$\Rightarrow \partial f_3(x)=\begin{cases}\{y\mid |y_1|+|y_2|\leqslant 1\}, & x=0,\\ (\operatorname{sign} x_1,\ 0), & |x_1|>|x_2|,\\ (0,\ \operatorname{sign} x_2), & |x_2|>|x_1|,\\ \operatorname{conv}((\operatorname{sign} x_1, 0), (0, \operatorname{sign} x_2)), & |x_1|=|x_2|=z,\ z\neq 0.\end{cases}$$

2) Субдифференциал нормы в нормированном пространстве.

Предложение 2. *Пусть X — нормированное пространство, X^* — его сопряженное, $N(x)=\|x\|$. Тогда*

$$\partial N(x)=\begin{cases} B^*, \text{ где } B^* \text{ — единичный шар сопряженного пространства, если } x=0,\\ \{x^*\in X^*\mid \|x^*\|=1,\ \langle x^*,\ x\rangle=\|x\|\}, \text{ если } x\neq 0.\end{cases}$$

Доказательство. Как было отмечено выше (соотношение (3) п. 2.6.3), множество опорных функций к f находится во взаимно однозначном соответствии с множеством $\{x^*\in X^*,\ f^*(x^*)\neq\pm\infty\}$, причем

$$a_{x^*}(x)=\langle x^*,\ x\rangle - f^*(x^*).$$

Согласно предложению 2 п. 2.6.3

$$N^*(x^*)=\delta B^*(x^*)=\begin{cases}0, & \|x^*\|\leqslant 1,\\ +\infty, & \|x^*\|>1,\end{cases}$$

так что опорными к N являются функции $\langle x^*, x\rangle$, $\|x^*\|\leqslant 1$. Равенство (2) превращается в

$$0=\langle x^*,\ x_0\rangle-\|x_0\|, \tag{3}$$

так что любая функция $x\to\langle x^*,\ x\rangle$, $\|x^*\|\leqslant 1$ является опорной к N в точке $x_0=0$ и $\partial N(0)=B^*$. Если же $x_0\neq 0$, то множество $x^* \in B^*$, для которых имеет место (3), совпадает с указанным в формулировке. ■

В приведенных выше примерах субдифференциал существовал в каждой точке. Разумеется, для невыпуклых функций (например, для $-|x|$) его может не быть ни в одной точке. Однако и выпуклые функции могут не иметь субдифференциалов даже в точках из эффективной области.

Вот простейший пример:

$$f(x)=\begin{cases}-\sqrt{1-x^2}, & |x|\leqslant 1,\\ +\infty, & |x|>1.\end{cases}$$

В точках $x_1=+1$ и $x_2=-1$ субдифференциал пуст. Достаточное условие существования субдифференциала будет получено далее (см. следствие из теоремы Моро—Рокафеллара).

Следующая теорема является аналогом в выпуклом анализе теоремы о линейном свойстве дифференциала.

Теорема Моро—Рокафеллара. *Пусть f_i, $i=1, \ldots, n$,—выпуклые собственные функции на X. Тогда*

$$\partial\left(\sum_{i=1}^{n} f_i\right)(x) \supseteq \sum_{i=1}^{n} \partial f_i(x). \tag{4}$$

Если же в некоторой точке $\bar{x}$ все функции, кроме, быть может, одной, непрерывны, а эта последняя в $\bar{x}$ конечна, то в любой точке x имеет место равенство

$$\partial\left(\sum_{i=1}^{n} f_i\right)(x) = \sum_{i=1}^{n} \partial f_i(x). \tag{5}$$

Доказательство. Ограничимся случаем $n=2$; большее число слагаемых рассматривается по индукции. Включение $\partial f_1(x)+\partial f_2(x) \subset \partial(f_1+f_2)(x)$ сразу следует из определения субдифференциала.

Пусть $x_0^* \in \partial(f_1+f_2)(x_0)$. Без ограничения общности можно считать, что $x_0=0$, $x_0^*=0$, $f_i(0)=0$, $i=1, 2$. Действительно, если эти соотношения не выполнены, то вместо f_1 и f_2 можно рассмотреть функции

$$g_1(x)=f_1(x_0+x)-f_1(x_0)-\langle x_0^*, x\rangle,$$
$$g_2(x)=f_2(x_0+x)-f_2(x_0).$$

Итак, пусть $0 \in \partial(f_1+f_2)(0)$. Согласно (1) это означает,

что
$$\min_x (f_1(x)+f_2(x)) = f_1(0)+f_2(0) = 0.$$

Пусть $\bar{x}$ — та точка, где f_1 непрерывна, а f_2 конечна. Рассмотрим два выпуклых множества
$$C_1 = \{(x, \alpha) \mid \alpha > f_1(x),\ x \in \operatorname{int} \operatorname{dom} f_1\} = \operatorname{int} \operatorname{epi} f_1,$$
$$C_2 = \{(x, \alpha) \mid -\alpha \geqslant f_2(x)\}.$$

Ясно, что множества C_1 и C_2 выпуклы, непусты ($(\bar{x}, -f_2(\bar{x})) \in C_2$), C_1 открыто и $\neq \varnothing$ в силу предложения 3 б) п. 2.6.2: непрерывная в точке $\bar{x}$ функция f_1 ограничена в ее окрестности) и $C_1 \cap C_2 = \varnothing$. Действительно, если $(\alpha_1, x_1) \in C_1 \cap C_2$, то $f_1(x_1) < \alpha \leqslant -f_2(x_1)$, т. е.
$$0 = \min_x (f_1(x)+f_2(x)) \leqslant f_1(x_1)+f_2(x_1) < 0,$$
чего не может быть. По первой теореме отделимости (см. п. 2.1.4) C_1 отделяется от C_2 ненулевым линейным функционалом (x_1^*, β):
$$\inf_{(x, \alpha) \in C_2} (\langle x_1^*, x\rangle + \beta\alpha) \geqslant \sup_{(x, \alpha) \in C_1} (\langle x_1^*, x\rangle + \beta\alpha). \quad (6)$$

Ясно, что $\beta \leqslant 0$, ибо иначе верхняя грань равнялась бы $+\infty$. Если допустить, что $\beta = 0$, то
$$\sup_{x \in \operatorname{int} \operatorname{dom} f_1} \langle x_1^*, x\rangle \leqslant \inf_{x \in \operatorname{dom} f_2} \langle x_1^*, x\rangle.$$

Но максимум линейной функции не может достигаться во внутренней точке, а потому
$$\langle x_1^*, \bar{x}\rangle < \sup_{x \in \operatorname{int} \operatorname{dom} f_1} \langle x_1^*, x\rangle \leqslant \inf_{x \in \operatorname{dom} f_2} \langle x_1^*, x\rangle \leqslant \langle x_1^*, \bar{x}\rangle.$$

Противоречие показывает, что $\beta \neq 0$. Следовательно, можно в (6) поделить все члены на $|\beta|$, и тогда, если обозначить $\tilde{x}_1^* = |\beta|^{-1} x_1^*$ и воспользоваться тем, что $\operatorname{epi} f_1 \subseteq \overline{\operatorname{int} \operatorname{epi} f_1} = \overline{C}_1$ (предложение 1 п. 2.6.2), мы получим
$$\sup_x \{\langle \tilde{x}_1^*, x\rangle - f_1(x)\} = \sup_{\left\{\substack{x \in \operatorname{dom} f_1 \\ \alpha \geqslant f_1(x)}\right.} \{\langle \tilde{x}_1^*, x\rangle - \alpha\} =$$
$$= \sup_{(x, \alpha) \in C_1} \{\langle \tilde{x}_1^*, x\rangle - \alpha\} \leqslant \inf_{(x, \alpha) \in C_2} \{\langle \tilde{x}_1^*, x\rangle - \alpha\} =$$
$$= \inf_x \{\langle \tilde{x}_1^*, x\rangle + f_2(x)\}.$$

При $x=0$ значения функций в фигурных скобках совпадают и равны нулю ($f_1(0)=f_2(0)=0$). Значит,

$$f_1(x)-f_1(0)\geqslant\langle\bar{x}_1^*, x\rangle, \quad f_2(x)-f_2(0)\geqslant\langle-\bar{x}_1^*, x\rangle,$$

т. е.

$$\bar{x}_1^*\in\partial f_1(0), \ -\bar{x}_1^*\in\partial f_2(0)$$

и

$$0\in\partial f_1(0)+\partial f_2(0). \blacksquare$$

Следствие 1. *Пусть f—выпуклая функция, непрерывная в точке x_0. Тогда $\partial f(x_0)\neq\varnothing$.*

Доказательство. Рассмотрим функцию

$$f+\delta\{x_0\}=\begin{cases} f(x_0), & x=x_0,\\ \infty, & x\neq x_0.\end{cases}$$

Функции f и $\delta\{x_0\}$ удовлетворяют условиям теоремы Моро—Рокафеллара. Равенство

$$X^*=\partial(f+\delta\{x_0\})\ (x_0)=\partial f(x_0)+\partial\delta\{x_0\}(x_0)=\partial f(x_0)+X^*$$

означает, что $\partial f(x_0)\neq\varnothing$. $\blacksquare$

На самом деле, субдифференциал $\partial f(x_0)$ выпуклой функции, непрерывной в точке x_0, является выпуклым компактом в $*$-слабой топологии.

Пусть K—конус (п. 2.6.1). Конус K^*, состоящий из тех элементов x^*, для которых $\langle x^*, x\rangle\geqslant 0$, $\forall x\in K$, называется *конусом, сопряженным с* K. Если $0\in K$, то из определения (1) вытекает сразу равенство $\partial\delta K(0)=-K^*$.

Следствие 2. *Пусть $K_1, \ldots, K_n$—выпуклые открытые конусы, имеющие непустое пересечение. Тогда*

$$\left(\bigcap_{i=1}^{n} K_i\right)^*=\sum_{i=1}^{n} K_i^*.$$

Действительно, добавим к K_i начало координат и рассмотрим функции δK_i. Применив к ним теорему Моро—Рокафеллара (в точке $\bar{x}\in\bigcap_{i=1}^{n} K_i$, отличной от нуля, все δK_i непрерывны), получим, что

$$\left(\bigcap_{i=1}^{n} K_i\right)^*=-\partial\delta\left(\bigcap_{i=1}^{n} K_i\right)(0)=-\partial\left(\sum_{i=1}^{n}\delta K_i\right)(0)=$$
$$=-\sum_{i=1}^{n}\partial\delta K_i(0)=\sum_{i=1}^{n} K_i^*.$$

Следствие 3. Теорема Дубовицкого—Милютина о пересечении конусов. *Для того чтобы выпуклые конусы $K_1, \ldots, K_n, K_{n+1}$, из которых первые n открыты, не пересекались, необходимо и достаточно, чтобы нашлись функционалы $x_i^* \in K_i^*$, $i=1, \ldots, n+1$, не равные одновременно нулю и такие, что* $\sum_{i=1}^{n+1} x_i^*=0$.

Доказательство. «Необходимо». Не ограничивая себя в общности, можно считать, что $K=\bigcap_{i=1}^{n} K_i \neq \varnothing$. Тогда K—открытый конус, не пересекающийся по условию с K_{n+1}. По первой теореме отделимости можно отделить K от K_{n+1} ненулевым линейиым функционалом $y^* \in X^*$:

$$\inf_{x \in K} \langle y^*, x\rangle \geqslant \sup_{x \in K_{n+1}} \langle y^*, x\rangle.$$

Последнее соотношение означает, что $y^* \in K^*$, а $(-1) y^* \in K_{n+1}^*$ (точка $x=0$—предельная как для K, так и для K_{n+1}). По следствию 2 разложим y^* в сумму $y^*= \sum_{i=1}^{n} x_i^*$, $x_i^* \in K_i^*$, $1 \leqslant i \leqslant n$; обозначив $(-1) y^*$ через x_{n+1}^*, получим требуемое.

«Достаточно». Пусть не равны одновременно нулю $x_i^* \in K_i^*$ и $\sum_{i=1}^{n+1} x_i^*=0$. Пусть $x \in \bigcap_{i=1}^{n+1} K_i$, $x \neq 0$, и $x_{i_0}^* \neq 0$, $1 \leqslant i_0 \leqslant n$. Тогда $x \in \operatorname{int} K_{i_0} \Rightarrow \langle x_{i_0}^*, x\rangle>0$ и, значит, $0=\langle\sum x_i^*, x\rangle>0$. Противоречие доказывает теорему. ■

Теорема Дубовицкого—Милютина о субдифференциале максимума. *Пусть $f_1, \ldots, f_n$—выпуклые функции на локально выпуклом топологическом пространстве X, непрерывные в точке x_0; $f(x)=\max \{f_1(x), \ldots, f_n(x)\}$; $I=\{i_1, \ldots, i_s\}$—набор индексов такой, что*

$$f_i(x_0)\begin{cases} =f(x_0) & \text{при } i \in I, \\ <f(x_0) & \text{при } i \notin I. \end{cases}$$

Тогда

$$\partial f(x_0)=\operatorname{conv}\{\partial f_{i_1}(x_0) \cup \ldots \cup \partial f_{i_s}(x_0)\}=$$
$$=\left\{x^* \mid x^*=\sum_{k=1}^{s} \lambda_k x_k^*,\ x_k^* \in \partial f_{i_k}(x_0),\quad \lambda_k \geqslant 0, \quad \sum_{k=1}^{s} \lambda_k=1\right\}.$$

Доказательство. А) Если $i \in I$ и $x^* \in \partial f_i(x_0)$, то

$$f(x) - f(x_0) = f(x) - f_i(x_0) \geqslant f_i(x) - f_i(x_0) \geqslant$$
$$\geqslant \langle x^*, x - x_0 \rangle \Rightarrow x^* \in \partial f(x_0) \Rightarrow \partial f(x_0) \supset \bigcup_{i \in I} \partial f_i(x_0) \Rightarrow$$
$$\Rightarrow \partial f(x_0) \supset \operatorname{conv} \left\{ \bigcup_{i \in I} \partial f_i(x_0) \right\},$$

причем последняя импликация является следствием выпуклости $\partial f(x_0)$ (предложение 1).

Обратное включение будет доказано индукцией по числу функций. При $n = 1$ утверждение очевидно. В дальнейшем будем предполагать, что для $(n-1)$ функции оно верно.

Б) Лемма. *Если функция $f(x)$ выпукла и неравенство*

$$\varphi(x) = f(x) - f(x_0) - \langle x^*, x - x_0 \rangle \geqslant 0 \tag{7}$$

выполняется в некотором открытом выпуклом множестве V, причем $\varphi(\bar{x}) = 0$ для некоторого $\bar{x} \in V$, то $x^* \in \partial f(x_0)$.

Доказательство. Покажем, что неравенство (7) имеет место при всех x; утверждение леммы следует отсюда по определению субдифференциала.

Пусть $\tilde{x}$ произвольно. Функция φ выпукла вместе с f, и если $\varphi(\tilde{x}) < 0$, то для любого $\alpha \in (0, 1)$

$$\varphi((1-\alpha)\bar{x} + \alpha\tilde{x}) \leqslant (1-\alpha)\varphi(\bar{x}) + \alpha\varphi(\tilde{x}) = \alpha\varphi(\tilde{x}) < 0.$$

При достаточно малом α точка $(1-\alpha)\bar{x} + \alpha\tilde{x} \in V$, и мы пришли к противоречию. Следовательно, $\varphi(x) \geqslant 0$ при всех x. ■

В) Пусть $x^* \in \partial f(x_0)$. Покажем, что тогда либо x_0 является решением следующей задачи выпуклого программирования:

$$\varphi_0(x) = f_{i_1}(x) - f(x_0) - \langle x^*, x - x_0 \rangle \to \inf, \tag{8}$$
$$\varphi_i(x) = f_i(x) - f(x_0) - \langle x^*, x - x_0 \rangle \leqslant 0, \quad i \neq i_1, \tag{9}$$

либо утверждение теоремы верно. Действительно, если первое не имеет места, то для некоторого $\bar{x}$

$$\varphi_0(\bar{x}) < 0 = \varphi_0(x_0), \quad \varphi_i(\bar{x}) \leqslant 0, \quad i \neq i_1. \tag{10}$$

По непрерывности неравенство $\varphi_0(x) < 0$ сохраняется в некоторой выпуклой окрестности $V \ni \bar{x}$, т. е.

$$f_{i_1}(x) - f(x_0) < \langle x^*, x - x_0 \rangle, \quad x \in V. \tag{11}$$

Теперь вспомним, что $x^* \in \partial f(x_0)$, а следовательно, должно выполняться неравенство

$$\max\{f_1(x), \ \ldots, \ f_n(x)\} - f(x_0) \geqslant \langle x^*, \ x - x_0\rangle.$$

Сопоставив его с неравенством (11), заключаем, что при $x \in V$ имеет место неравенство

$$\varphi(x) = \max\{f_i(x) \mid i \neq i_1\} - f(x_0) - \langle x^*, \ x - x_0\rangle \geqslant 0. \quad (12)$$

В то же время, согласно (10),

$$\varphi(\bar{x}) = \max\{f_i(\bar{x}) \mid i \neq i_1\} - f(x_0) - \langle x^*, \ \bar{x} - x_0\rangle \leqslant 0,$$

так что $\varphi(\bar{x}) = 0$ и по лемме $x^* \in \partial\tilde{f}(x_0)$, где

$$\tilde{f}(x) = \max\{f_i(x) \mid i \neq i_1\}.$$

Поскольку $\tilde{f}$ образовано $(n-1)$ функцией, то в силу индукционной гипотезы

$$x^* \in \operatorname{conv}\left\{\bigcup_{k \geqslant 2} \partial f_{i_k}(x_0)\right\} \subset \operatorname{conv}\left\{\bigcup_{k \geqslant 1} \partial f_{i_k}(x_0)\right\}.$$

Г) Зная, что x_0 — решение задачи (8)—(9), мы можем воспользоваться теоремой Куна — Таккера (п. 1.3.3); следует обратить внимание на замечание, сделанное после доказательства этой теоремы: хотя в силу предложения 3 п. 2.6.2 наши функции непрерывны на $\operatorname{int}\operatorname{dom}$, вне него они могут обращаться в $+\infty$.

По теореме Куна — Таккера существуют не равные нулю одновременно множители Лагранжа λ_0, $\lambda_i \geqslant 0$, $i \neq i_1$, для которых x_0 — точка минимума функции Лагранжа

$$\mathscr{L}(x; \lambda) = \lambda_0\varphi_0(x) + \sum_{i \neq i_1} \lambda_i\varphi_i(x).$$

Кроме того, должны выполняться условия дополняющей нежесткости, согласно которым $\lambda_i = 0$, если $\varphi_i(x_0) = f_i(x_0) - f(x_0) \neq 0$, т. е. при $i \notin I$. Перенумеровав λ_0 в λ_{i_1} и φ_0 в φ_{i_1}, имеем $\mathscr{L}(x, \lambda) = \sum_{i \in I} \lambda_i\varphi_i(x)$, а так как множители Лагранжа определены с точностью до положительного множителя, можно считать, что $\sum_{i \in I} \lambda_i = 1$.

Теперь имеем

$$\mathscr{L}(x, \lambda) \geqslant \mathscr{L}(x_0, \lambda) \Leftrightarrow \sum_{i \in I} \lambda_i\varphi(x) \geqslant 0 \Leftrightarrow$$

$$\Leftrightarrow \sum_{i \in I} \lambda_i f_i(x) - \left(\sum_{i \in I} \lambda_i\right) f(x_0) - \left(\sum_{i \in I} \lambda_i\right) \langle x^*, \ x - x_0\rangle \geqslant 0 \Leftrightarrow$$

$$\Leftrightarrow x^* \in \partial\left(\sum_{i \in I} \lambda_i f_i(\cdot)\right)(x_0).$$

Применяя теорему Моро — Рокафеллара и учитывая, что из определения субдифференциала вытекает очевидное равенство

$$\partial(\lambda f)(x_0) = \lambda \partial f(x_0),$$

справедливое при $\lambda > 0$, мы убеждаемся в том, что

$$x^* = \sum_{i \in I} \lambda_i x_i^*, \quad x_i^* \in \partial f_i(x_0),$$

$$\lambda_i \geqslant 0, \quad \sum_{i \in I} \lambda_i = 1.$$

ГЛАВА III

ПРИНЦИП ЛАГРАНЖА ДЛЯ ГЛАДКИХ ЗАДАЧ С ОГРАНИЧЕНИЯМИ

Основная цель этой главы — обоснование принципа Лагранжа для гладких задач с ограничениями типа равенств и неравенств, а также вывод достаточных условий экстремума для таких задач. Общие результаты этой главы применяются затем в гл. 4 к задачам классического вариационного исчисления и оптимального управления. Поскольку на элементарном уровне часть материала излагалась уже в гл. 1, читателю рекомендуется параллельно просматривать соответствующие места в §§ 1.3—1.4.

§ 3.1. Элементарные задачи

3.1.1. Элементарные задачи без ограничений. Пусть X — топологическое пространство, U — окрестность в X, $f\colon U \to \mathbf{R}$. Задача

$$f(x) \to \mathrm{extr} \qquad (\mathfrak{z})$$

называется элементарной задачей без ограничений (экстремальной — здесь и далее подразумевается). В случае, если X — линейное нормированное пространство и f обладает гладкостью в каком-либо смысле, то задачу $(\mathfrak{z})$ называют *элементарной гладкой задачей*; если X — линейное топологическое пространство и f — выпуклая функция, то задачу $(\mathfrak{z})$ на минимум называют *элементарной выпуклой задачей*.

Определение локального экстремума для задачи $(\mathfrak{z})$ см. в п. 1.2.1. Если $\hat{x}$ доставляет локальный минимум (максимум, экстремум), в задаче $(\mathfrak{z})$, мы пишем кратко $\hat{x} \in \mathrm{locmin}\,\mathfrak{z}$ ($\hat{x} \in \mathrm{locmax}\,\mathfrak{z}$, $\hat{x} \in \mathrm{locextr}\,\mathfrak{z}$). Определения различных терминов, связанных с понятием гладкости см. в § 2.2.

Начнем с простейшего — одномерного случая, когда $X=\mathbf{R}$.

Лемма. *Пусть функция* $f\colon U\to\mathbf{R}$ *определена на некотором интервале* $U\subseteq\mathbf{R}$, *содержащем точку* $\hat{x}$.

а) *Если* f *имеет в точке* $\hat{x}$ *производные справа и слева и* $\hat{x}\in\operatorname{locmin}\mathfrak{z}$ $(\operatorname{locmax}\mathfrak{z})$, *то*

$$f'(\hat{x}+0)\geqslant 0,\quad f'(\hat{x}-0)\leqslant 0\ \ (f'(\hat{x}+0)\leqslant 0,\ f'(\hat{x}-0)\geqslant 0). \tag{1}$$

Если же в (1) *выполнены строгие неравенства, то* $\hat{x}\in\operatorname{locmin}\mathfrak{z}$ $(\operatorname{locmax}\mathfrak{z})$.

Пусть далее функция f *дифференцируема* k *раз в точке* $\hat{x}$.

б) *Если* $\hat{x}\in\operatorname{locmin}\mathfrak{z}\,(\operatorname{locmax}\mathfrak{z})$, *то либо* $f^{(i)}(\hat{x})=0$, $1\leqslant i\leqslant k$, *либо существует такое* s, $1\leqslant s\leqslant[k/2]$, *что*

$$f'(\hat{x})=\ldots=f^{(2s-1)}(\hat{x})=0,\quad f^{(2s)}(\hat{x})>0\qquad (f^{(2s)}(\hat{x})<0). \tag{2}$$

в) *Если существует такое* s, $1\leqslant s\leqslant[k/2]$, *что выполнены соотношения* (2), *то* $\hat{x}\in\operatorname{locmin}\mathfrak{z}$ $(\operatorname{locmax}\mathfrak{z})$.

Утверждение б) при $k=1$ $(\hat{x}\in\operatorname{locextr}\mathfrak{z}\Rightarrow f'(\hat{x})=0)$ известно в анализе как *теорема Ферма*. Эту теорему мы доказали в п. 1.4.1.

Доказательство. Утверждения а) сразу следуют из определений односторонних производных и локального экстремума. Утверждение б) при $k=1$ (теорема Ферма) вытекает из а), если учесть, что существование производной влечет за собой равенства $f'(\hat{x})=f'(\hat{x}+0)=$ $=f'(\hat{x}-0)$. Пусть далее $k>1$ и для определенности речь идет о минимуме. По формуле Тейлора

$$f(\hat{x}+h)=\sum_{j=0}^{k}\frac{f^{(j)}(\hat{x})}{j!}h^k+o(h^k). \tag{3}$$

Пусть $f'(\hat{x})=\ldots=f^{(m-1)}(\hat{x})=0$, $1<m\leqslant k$, $f^{(m)}(\hat{x})\neq 0$. Возможно одно из двух: m — нечетно и m — четно. В первом случае пусть $\varphi(\xi)=f\left(\hat{x}+\sqrt[m]{\xi}\right)$, $\xi\in\mathbf{R}$. Из (3) следует тогда, что $\varphi\in D^1(0)$ и $\varphi'(0)=f^{(m)}(\hat{x})/m!\neq 0$, а по теореме Ферма должно быть $\varphi'(0)=0$. Противоречие показывает,

что m должно быть четным. Но тогда, полагая $\psi(\xi)=f(\hat{x}+\sqrt[m]{\xi})$, $\xi\geqslant 0$, получаем из (1) $\psi'(+0)=f^{(m)}(\hat{x})/m!>0$, доказывающее утверждение б). Утверждение в) сразу вытекает из формулы Тейлора $f(\hat{x}+h)=f(\hat{x})+f^{(2s)}(\hat{x})h^s/(2s)!+o(h^{2s})$, из которой видно, что если $f^{(2s)}(\hat{x})>0$, то $\hat{x}\in\operatorname{locmin}\mathfrak{z}$, если же $f^{(2s)}(\hat{x})<0$, то $\hat{x}\in\operatorname{locmax}\mathfrak{z}$. ■

Следующие утверждения являются почти непосредственными следствиями определений.

Теорема 1 (необходимое условие экстремума первого порядка). *Пусть в* ($\mathfrak{z}$) X — *нормированное линейное пространство и функция* f *имеет в точке* $\hat{x}\in U$ *производную по направлению* h.

Если $\hat{x}\in\operatorname{locmin}\mathfrak{z}$ ($\hat{x}\in\operatorname{locmax}\mathfrak{z}$), *то*

$$f'(\hat{x};h)\geqslant 0 \quad (f'(\hat{x};h)\leqslant 0). \tag{4}$$

Следствие 1. *Пусть* f *имеет первую вариацию (первую вариацию по Лагранжу) в точке* $\hat{x}$. *Если* $\hat{x}\in\operatorname{locmin}\mathfrak{z}$ ($\hat{x}\in\operatorname{locmax}\mathfrak{z}$), *то*

$$\delta_+f(\hat{x},h)\geqslant 0,\quad \forall h\in X \qquad (\delta_+f(\hat{x},h)\leqslant 0,\quad \forall h\in X) \tag{5}$$
$$(\delta f(\hat{x},h)=0,\ \forall h\in X).$$

Следствие 2. *Пусть* f *имеет производную в смысле Фреше (Гато) в точке* $\hat{x}$. *Тогда, если* $\hat{x}\in\operatorname{locextr}\mathfrak{z}$, *то*

$$f'(\hat{x})=0 \quad (f'_\Gamma(\hat{x})=0). \tag{6}$$

Утверждение следствия 2 также называют теоремой Ферма. Точки $\hat{x}$, в которых выполнено равенство (6), называют *стационарными* точками задачи ($\mathfrak{z}$).

Теорема 2 (необходимое условие экстремума второго порядка). *Пусть в* ($\mathfrak{z}$) X — *нормированное линейное пространство и* f *имеет вторую вариацию по Лагранжу в точке* $\hat{x}\in U$. *Тогда, если* $\hat{x}\in\operatorname{locmin}\mathfrak{z}$ ($\hat{x}\in\operatorname{locmax}\mathfrak{z}$), *то выполнены следующие соотношения:*

$$\delta f(\hat{x},h)=0,\quad \forall h\in X, \tag{7}$$

$$\delta^2f(\hat{x},h)\geqslant 0,\quad \forall h\in X \quad (\delta^2f(\hat{x},h)\leqslant 0,\ \forall h\in X). \tag{8}$$

Доказательство. Равенство (7) содержится в следствии 1. Неравенства (8) немедленно вытекают из определения второй вариации по Лагранжу и утверждения б) леммы. ■

Следствие 3. *Пусть f имеет вторую производную в смысле Фреше в точке $\hat{x}$. Тогда, если $\hat{x} \in \operatorname{locmin} \mathfrak{z}$ ($\hat{x} \in \operatorname{locmax} \mathfrak{z}$), то выполнены такие соотношения:*

$$f'(\hat{x}) = 0, \tag{9}$$

$$f''(\hat{x}) \geqslant 0 \qquad (f''(\hat{x}) \leqslant 0). \tag{10}$$

(Эти неравенства означают, что квадратичная форма $d^2 f = f''(\hat{x})[h, h]$ неотрицательна (неположительна).)

Доказательство. Воспользовавшись формулой Тейлора (см. п. 2.2.5), имеем

$$\varphi(\alpha) = f(\hat{x} + \alpha h) = f(\hat{x}) + \alpha f'(\hat{x})[h] + \frac{\alpha^2}{2} f''(\hat{x})[h, h] + o(\alpha^2),$$

откуда

$$\delta f(\hat{x}, h) = \varphi'(0) = f'(\hat{x})[h],$$

$$\delta^2 f(\hat{x}, h) = \varphi''(0) = f''(\hat{x})[h, h],$$

и остается сослаться на (7) и (8). ■

Следствие 2 утверждает, что локальные экстремумы являются стационарными точками. Мы упоминали уже в п. 1.4.1, что обратное неверно. Чтобы получить достаточные условия, приходится привлекать к рассмотрению производные более высоких порядков, как это мы уже делали в лемме.

Теорема 3 (достаточное условие экстремума второго порядка). *Пусть в ($\mathfrak{z}$) X — нормированное пространство и f имеет вторую производную в смысле Фреше. Тогда, если выполнены соотношения*

$$f'(\hat{x}) = 0, \tag{11}$$

$$f''(\hat{x})[h, h] \geqslant \alpha \|h\|^2, \quad \forall h \in X \ (f''(\hat{x})[h, h] \leqslant -\alpha \|h\|^2, \ \forall h \in X) \tag{12}$$

для некоторого $\alpha > 0$, то $\hat{x} \in \operatorname{locmin} \mathfrak{z}$ ($\hat{x} \in \operatorname{locmax} \mathfrak{z}$).

Доказательство. Если выполнено первое из неравенств (12), то, снова обратившись к формуле Тейлора, имеем

$$f(\hat{x} + h) = f(\hat{x}) + f'(\hat{x})[h] + \frac{f''(\hat{x})}{2}[h, h] + o(\|h\|^2) \geqslant$$

$$\geqslant f(\hat{x}) + \frac{\alpha}{2}\|h\|^2 + o\|h\|^2) > f(\hat{x}),$$

если $h \neq 0$ и $\|h\|$ достаточно мала. Следовательно, $\hat{x} \in \operatorname{locmin} \mathfrak{z}$. Случай максимума рассматривается аналогично. ■

Условие (12) называют условием *строгой положительности (отрицательности)* второго дифференциала f.

Замечания. 1. В конечномерном случае, когда $X = \mathbf{R}^n$, утверждения теорем 1—3 хорошо известны из курса математического анализа (см. об этом также в п. 1.4.1). Теорема Ферма означает следующее:

$$f'(\hat{x}) = 0 \Leftrightarrow \partial f(\hat{x})/\partial x_1 = \ldots = \partial f(\hat{x})/\partial x_n = 0. \qquad (13)$$

Из необходимого условия второго порядка вытекает, что в конечномерной задаче на экстремум утверждение $\hat{x} \in \operatorname{locmin} \mathfrak{z}$ влечет (помимо (13)) неотрицательную определенность матрицы $\left(\frac{\partial^2 f(\hat{x})}{\partial x_i \partial x_j}\right)$:

$$\sum_{i,\, j=1}^{n} \frac{\partial^2 f(\hat{x})}{\partial x_i \partial x_j} h_i h_j \geqslant 0, \quad \forall h \in \mathbf{R}^n. \qquad (14)$$

Условие положительной определенности матрицы $\left(\frac{\partial^2 f(\hat{x})}{\partial x_i \partial x_j}\right)$ гарантирует, как нетрудно убедиться, строгую положительность второго дифференциала (и, значит, является достаточным условием минимума в стационарной точке). В бесконечномерных пространствах это не так.

Пример. Пусть $X = l_2$, $f(x) = \sum_{j=1}^{\infty} ((x_j^2/n^3) - x_j^4)$. Тогда точка $\hat{x} = 0$ является стационарной, $f'(0) = 0$, а второй дифференциал f в нуле — положительно определенным

$$f''(0)[h, h] = \sum_{j=1}^{\infty} h_j^2/n^3.$$

Но вместе с тем $0 \notin \operatorname{locmin} \mathfrak{z}$, ибо на последовательности $\{e_n/n\}_{n=1}^{\infty}$ (e_i — i-й базисный вектор пространства l_2) функционал f принимает отрицательные значения: $f(e_n/n) = (1/n^5) - (1/n^4) < 0$, а сама последовательность стремится к нулю.

2. В одномерном случае лемма, с которой начинается этот пункт, дает в некотором отношении исчерпывающий ответ на вопрос о том, является ли заданная точка точкой локального минимума бесконечно дифференцируемой функции f или нет: за исключением тривиального случая,

когда все производные в даниой точке равны нулю, лемма дает ответ на поставленный вопрос. Даже для функций двух переменных такого рода исчерпывающей процедуры по-видимому нет.

3. Неравенства (12) накладывают серьезные ограничения на структуру нормированного пространства X. Если выполнено первое из них, то симметричная билинейная функция $(\xi \mid \eta) = f''(\hat{x})[\xi, \eta]$ задает в X скалярное произведение, а неравенства

$$\|f''(\hat{x})\| \|\xi\|^2 \geqslant (\xi \mid \xi) = f''(\hat{x})[\xi, \xi] \geqslant \alpha \|\xi\|^2$$

показывают, что норма, порожденная этим скалярным произведением, эквивалентна исходной. Следовательно, X линейно гомеоморфно евклидову (предгильбертову) пространству. В частности, неравенства (12) не могут иметь места в пространствах C и C^1.

3.1.2. Элементарная задача линейного программирования. В п. 1.2.6 мы в общих чертах описали класс задач линейного программирования. Здесь будет рассмотрена самая простая из задач этого класса. Для нас она интересна тем, что при ее рассмотрении мы познакомимся с важными условиями экстремума — условиями *соответствия* и *согласования* знаков и условием *дополняющей нежесткости.*

Пусть $X = \mathbf{R}^n$, $a \in \mathbf{R}^{n*}$ — элемент сопряженного пространства, $a = (a_1, \ldots, a_n)$. Рассмотрим задачу

$$ax \to \inf(\sup); \quad \Leftrightarrow \sum_{i=1}^{n} a_i x_i \to \inf(\sup); \qquad (\mathfrak{z})$$
$$x_i \gtreqless 0.$$

Символ $\gtreqless$ здесь означает, что в соответствующем ограничении может стоять один из знаков $\geqslant$, $\leqslant$ или $=$. Задачу ($\mathfrak{z}$) будем называть *элементарной задачей линейного программирования.*

Будем говорить, что для вектора a выполнено условие *соответствия* знаков (с ограничениями $x_i \gtreqless 0$), если $a_i \geqslant 0$ при ограничении $x_i \geqslant 0$, $a_i \leqslant 0$ при ограничении $x_i \leqslant 0$ и a_i может иметь любой знак, если ограничение имеет вид равенства $x_i = 0$. В случае, если неравенства «оборачиваются», т. е. если $a_i \geqslant 0$ при ограничении $x_i \leqslant 0$, $a_i \leqslant 0$ при ограничении $x_i \geqslant 0$ (и снова a_i — любого знака при $x_i = 0$), то будем говорить, что выполнено условие *согласования* знаков. В первом случае мы пишем $a_i \gtrless 0$, во втором $a_i \lesseqgtr 0$.

Теорема (необходимое и достаточное условие экстремума в элементарной задаче линейного программирования). *Для того, чтобы точка $\hat{x} \in \mathbf{R}^n$ доставляла абсолютный минимум (максимум) задаче* ($\mathfrak{z}$) *необходимо и достаточно, чтобы были выполнены:*

а) *условия соответствия (согласования) знаков*

$$a_i \gtrless 0 \qquad (a_i \lesseqgtr 0), \tag{1}$$

б) *условия дополняющей нежесткости*

$$a_i \hat{x}_i = 0, \quad i = 1, \ldots, n. \tag{2}$$

Доказательство. Допустим, что ($\mathfrak{z}$) — задача минимизации и i_0-е ограничение имеет вид $x_{i_0} \geqslant 0$. Если допустить, что $a_{i_0} < 0$, то нижняя грань в задаче ($\mathfrak{z}$) равна $-\infty$, ибо x_{i_0} может принимать сколь угодно большие положительные значения. Поэтому условие $a_{i_0} \geqslant 0$ необходимо для того, чтобы значение задачи (искомый inf) было конечно. Если допустить теперь, что $a_{i_0}\hat{x}_{i_0} \neq 0$, то вследствие условия $\hat{x}_{i_0} \geqslant 0$ и доказанного неравенства $a_{i_0} \geqslant 0$ на самом деле $a_{i_0}\hat{x}_{i_0} > 0$. Но тогда точка $\hat{x}$ не будет минималью задачи, ибо на допустимом векторе $\tilde{x} = (\hat{x}_1, \ldots, \hat{x}_{i_0-1}, 0, \hat{x}_{i_0+1}, \ldots, \hat{x}_n)$ имеем $a\tilde{x} < a\hat{x}$. Достаточность условий (1), (2) очевидна: если для определенности ($\mathfrak{z}$) — по-прежнему задача минимизации и выполнены условия соответствия знаков и дополняющей нежесткости, то для любого допустимого вектора $x \in \mathbf{R}^n$ получаем $ax \geqslant 0$, в то время, как для $\hat{x}$ в силу (2) $a\hat{x} = 0$. ■

Упражнение. Вывести доказанную теорему из теоремы Куна — Таккера.

3.1.3. Задача Больца. Пусть Δ — конечный замкнутый отрезок вещественной прямой. Рассмотрим банахово пространство

$$\Xi = C^1(\Delta, \mathbf{R}^n) \times \mathbf{R} \times \mathbf{R},$$

состоящее из элементов $\xi = (x(\cdot), t_0, t_1)$. В этом пространстве рассмотрим задачу

$$\mathscr{B}(\xi) = \mathscr{B}(x(\cdot), t_0, t_1) =$$
$$= \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt + l(t_0, x(t_0), t_1, x(t_1)) \to \text{extr}. \tag{$\mathfrak{z}$}$$

При этом в (з) функции $(t, x, \dot{x}) \mapsto L(t, x, \dot{x})$ и $(t_0, x_0, t_1, x_1) \to \to l(t_0, x_0, t_1, x_1)$ определены в открытых подмножествах V и W пространств $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^n$ и $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}\times\mathbf{R}^n$ соответственно и $t_0, t_1 \in \Delta$. Обе функции будем предполагать по меньшей мере непрерывными. Задача (з) называется *задачей Больца с незакрепленным временем*, а функционал в (з) — *функционалом Больца*. В п. 1.4.2 был рассмотрен ее частный вариант, когда $n=1$ (о случае $n>1$ было лишь кратко упомянуто) и моменты времени t_0 и t_1 были закрепленными.

Лемма. *Пусть функции L и l и их частные производные L_x, $L_{\dot{x}}$, l_{t_i} и l_{x_i}, $i=0,1$, непрерывны в областях V и W соответственно. Тогда функционал $\mathscr{B}$ является непрерывно дифференцируемым (по Фреше) на следующем открытом подмножестве пространства Ξ:*

$$\mathscr{U} = \{\xi = (x(\cdot),\ t_0,\ t_1) \mid x(\cdot) \in C^1(\Delta,\ \mathbf{R}^n);\ t \in \Delta,$$
$$(t,\ x(t),\ \dot{x}(t)) \in V;\ (t_0,\ x(t_0),\ t_1,\ x(t_1)) \in W,\ t_0,\ t_1 \in \operatorname{int}\Delta\},$$

и при этом

$$\mathscr{B}'(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)[h(\cdot),\ \tau_0,\ \tau_1] = \int_{\hat{t}_0}^{\hat{t}_1} (\hat{L}_x(t)h(t) +$$
$$+\hat{L}_{\dot{x}}(t)\dot{h}(t))dt + \hat{L}(\hat{t}_1)\tau_1 - \hat{L}(\hat{t}_0)\tau_0 + \hat{l}_{t_0}\tau_0 + \hat{l}_{t_1}\tau_1 +$$
$$+\hat{l}_{x_0}(h(\hat{t}_0) + \dot{\hat{x}}(\hat{t}_0)\tau_0) + \hat{l}_{x_1}(h(\hat{t}_1) + \dot{\hat{x}}(\hat{t}_0)\tau_1), \quad (1)$$

где

$$\hat{L}(t) = L(t,\ \hat{x}(t),\ \dot{\hat{x}}(t)),\ \hat{L}_x(t) = L_x(t,\ \hat{x}(t),\ \dot{\hat{x}}(t)),$$
$$\hat{L}_{\dot{x}}(t) = L_{\dot{x}}(t,\ \hat{x}(t),\ \dot{\hat{x}}(t)), \quad \hat{l}_{t_i} = l_{t_i}(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \hat{t}_1,\ \hat{x}(\hat{t}_1)),$$
$$\hat{l}_{x_i} = l_{x_i}(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \hat{t}_1,\ \hat{x}(\hat{t}_1)), \quad i = 0,\ 1.$$

Доказательство. Функционал $\mathscr{B}$ является суммой интегрального и терминального функционалов. Их непрерывная дифференцируемость была доказана в пп. 2.4.2 и 2.4.3. (Терминальный функционал $\xi \to l(t_0,\ x(t_0),\ t_1,\ x(t_1))$ является частным случаем оператора краевых условий.) Чтобы получить (1), остается воспользоваться формулами (9) п. 2.4.2 и (3) п. 2.4.3. ■

В соответствии с этой леммой (з) является элементарной гладкой задачей без ограничений. Поэтому мы можем применять к ней необходимые и достаточные условия экстремума п. 3.1.1. Расшифровка утверждения теоремы Ферма приведет нас к следующей теореме.

Теорема 1 (необходимое условие экстремума первого порядка в задаче Больца). *Пусть выполнены условия леммы. Если элемент* $\hat{\xi}=(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)\in\Xi$ *принадлежит множеству* $\mathcal{U}$ *и доставляет локальный минимум задаче* ($\mathfrak{z}$), *то*

1) *выполнены условия стационарности по* $x(\cdot)$:

а) *уравнение Эйлера*

$$-\frac{d}{dt}\hat{L}_{\dot{x}}(t)+\hat{L}_x(t)=0, \tag{2}$$

б) *условия трансверсальности по* x

$$\hat{L}_{\dot{x}}(\hat{t}_i)=(-1)^i\hat{l}_{x_i},\quad i=0,\ 1; \tag{3}$$

2) *выполнены условия трансверсальности по* t:

$$\hat{H}(\hat{t}_i)=(-1)^{i+1}\hat{l}_{t_i},\quad i=0,\ 1, \tag{4}$$

где

$$\hat{H}(t)=\hat{L}_{\dot{x}}(t)\dot{\hat{x}}(t)-\hat{L}(t).$$

Соотношения (2) и (3) уже встретились нам в п. 1.4.2. Условия трансверсальности по t встречаются здесь впервые, но впоследствии в гл. 4 мы постоянно будем иметь с ними дело.

Доказательство. Было сказано уже, что задача ($\mathfrak{z}$) является гладкой элементарной задачей, и к ней применима, следовательно, теорема Ферма (следствие 2 п. 3.1.1), согласно которой

$$\hat{\xi}\in\operatorname{locextr}\mathfrak{z}\Rightarrow\mathcal{B}'(\hat{\xi})=0\Leftrightarrow\mathcal{B}'(\hat{\xi})[\eta]=0,$$
$$\forall\eta=(h(\cdot),\ \tau_0,\ \tau_1)\in\Xi. \tag{5}$$

Вспоминая (1), получаем из (5)

$$0=\mathcal{B}'(\hat{\xi})[\eta]=$$
$$=\int_{\hat{t}_0}^{\hat{t}_1}(\hat{L}_x(t)h(t)+\hat{L}_{\dot{x}}(t)\dot{h}(t))\,dt+\alpha_0h(\hat{t}_0)+$$
$$+\alpha_1h(\hat{t}_1)+\beta_0\tau_0+\beta_1\tau_1, \tag{6}$$

где для краткости введены такие обозначения:

$$\alpha_i=\hat{l}_{x_i},\quad \beta_i=\hat{l}_{t_i}+\hat{l}_{x_i}\dot{\hat{x}}(\hat{t}_i)+(-1)^{i+1}\hat{L}(\hat{t}_i),\quad i=0,\ 1. \tag{7}$$

Соотношение $\mathscr{B}'(\hat{\xi})[\eta]=0$ имеет место для любого элемента $\eta=(h(\cdot), \tau_0, \tau_1)$. Рассмотрим сначала элементы вида $\eta=(h(\cdot), 0, 0)$, $h\in C^1(\Delta, \mathbf{R}^n)$, $h(t_i)=0$, $i=1, 2$. Тогда из (6)

$$\int_{\hat{t}_0}^{\hat{t}_1}(\hat{L}_x(t)h(t)+\hat{L}_{\dot{x}}(t)\dot{h}(t))\,dt=0 \qquad (8)$$

для любой вектор-фуикции $h(\cdot)\in C^1([\hat{t}_0, \hat{t}_1], \mathbf{R}^n)$ такой, что $h(\hat{t}_0)=h(\hat{t}_1)=0$. Но с такой ситуацией нам пришлось уже встречаться в п. 1.4.1. Из леммы Дюбуа-Реймона, доказанной там, мы сразу получаем соотношение (2) (в п. 1.4.2 лемма Дюбуа-Реймона была доказана для скалярных функций, и здесь ее следует применять поочередно к каждой компоненте вектор-функции $h(\cdot)$; полагая $h_1(\cdot)=\ldots=h_{i-1}(\cdot)=h_{i+1}(\cdot)=\ldots=h_n(\cdot)=0$, мы сводим (8) к аналогичному скалярному соотношению с одной компонентой $h_i(\cdot)$ и получаем i-е уравнение системы (2)). При этом заодно доказывается непрерывная дифференцируемость $\hat{L}_{\dot{x}}(t)$. Но тогда в выражении (6) можно произвести интегрирование по частям, что (с учетом (2)) дает

$$0=\mathscr{B}'(\hat{\xi})[\eta]=(\alpha_0-\hat{L}_{\dot{x}}(\hat{t}_0))h(t_0)+$$
$$+(\alpha_1+\hat{L}_{\dot{x}}(\hat{t}_1))h(t_1)+\beta_0\tau_0+\beta_1\tau_1 \qquad (9)$$

для произвольных векторов $h(t_0)$, $h(t_1)$ и чисел τ_0 и τ_1. Отсюда и из (7) немедленно следуют соотношения (3) и (4). ■

Как мы видим, доказательство теоремы, в сущности, такое же, как и в гл. I. Только там нам пришлось вычислять производные кустарно, а здесь мы воспользовались общими теоремами дифференциального исчисления §§ 2.2 и 2.4.

3.1.4. Элементарная задача оптимального управления. Пусть $\mathfrak{U}$—топологическое пространство и $\varphi:[t_0, t_1]\times\times\mathfrak{U}\to\mathbf{R}$. Рассмотрим задачу

$$f(u(\cdot))=\int_{t_0}^{t_1}\varphi(t, u(t))\,dt\to\inf \qquad (\mathfrak{z})$$

в пространстве $KC([t_0, t_1], \mathfrak{U})$ кусочно непрерывных функций $u(\cdot):[t_0, t_1]\to\mathfrak{U}$ со значениями в $\mathfrak{U}$. Функцию φ

будем предполагать непрерывной в $[t_0, t_1]\times\mathfrak{U}$. Задачу ($\mathfrak{z}$) назовем *элементарной задачей оптимального управления.*

Теорема (необходимое и достаточное условие минимума в элементарной задаче оптимального управления). *Пусть функция* $u(\cdot)\in \in KC([t_0, t_1], \mathfrak{U})$ *и доставляет абсолютный минимум в задаче* ($\mathfrak{z}$). *Тогда для любой точки непрерывности функции* $\hat{u}(\cdot)$ *выполнено соотношение*

$$\min_{u\in\mathfrak{U}} f(t, u)=f(t, \hat{u}(t)). \tag{1}$$

Доказательство. Допустим, что соотношение (1) не выполнено и существуют точка τ (где $\hat{u}(\cdot)$ непрерывна) и элемент $v\in\mathfrak{U}$ такие, что $f(\tau, v)<f(\tau, \hat{u}(\tau))$. Вследствие непрерывности функций $t\to f(t, v)$ и $t\to f(t, \hat{u}(t))$ в окрестности точки τ найдется такой интервал $\Delta=[\tau-\delta, \tau+\delta]$, что $f(t, v)<f(t, \hat{u}(t))$ при $t\in\Delta$. Положим $\tilde{u}(t)=\hat{u}(t)$ при $t\notin\Delta$ и $\tilde{u}(t)=v$ при $t\in\Delta$. Тогда $f(\tilde{u}(\cdot))<f(\hat{u}(\cdot))$ вопреки минимальности $\hat{u}(\cdot)$.

3.1.5. Принцип Лагранжа для задач с равенствами и неравенствами. В гл. 1 уже несколько раз обсуждался принцип, которому соответствуют необходимые условия для задач с ограничениями. Здесь, после того как иами выделены некоторые «элементарные» задачн, можно подвести кое-какие итоги.

Рассматрнвавшиеся экстремальные задачи были формализованы так, что ограннчения делились на две группы, первая нз которых имела вид равенств. По ограиичениям этой группы составлялась функцня Лагранжа. Далее мысленио ставилась задача о соответствующем экстремуме фуикции Лагранжа по второй группе ограничений—по той группе, которая не участвовала в формировании функции Лагранжа. При этом оказывалось, что полученная экстремальная задача либо сама оказывалась элементарной, лнбо элементарными оказывались «частичные» задачи, получаемые фиксированием всех неизвестных, кроме одного. Необходимые условия для полученных элементарных задач в своей совокупности и составляли искомый набор необходимых условий экстремума. Предоставим читателю убедиться, что все необходимые условия, о которых речь шла в гл. 1, и все необходимые условия, о которых речь пойдет далее в гл. 4, соответствуют описанной процедуре. Исключение составляют задачи с неравенствами, в которых появляются дополнительные условия. Этими задачами мы займемся в следующем параграфе, а здесь покажем, что и они также могут быть включены в рамки описанного «принципа Лагранжа».

Пусть (для определенности) у нас имеется задача минимизации

$$f_0(x) \to \inf; \quad F(x) = 0, \quad f_i(x) \leqslant 0, \quad i = 1, \ldots, m, \tag{$\mathfrak{z}$}$$

где наряду с равенствами встречаются и неравенства (X и Y в $(\mathfrak{z})$ — линейные топологические пространства, $f_i\colon X \to \mathbf{R}$, $F\colon X \to Y$). Введением новых переменных u_i приведем задачу $(\mathfrak{z})$ к виду

$$\begin{gathered} f_0(x) \to \inf; \quad F(x) = 0, \quad f_i(x) + u_i = 0, \\ u_i \geqslant 0, \quad i = 1, \ldots, m, \end{gathered} \tag{$\tilde{\mathfrak{z}}$}$$

и разобьем ограничения на две группы, первая из которых состоит из равенств $(F(x) = 0,\ f_i(x) + u_i = 0)$, а вторая группа — из ограничений вида $u_i \geqslant 0$. Для задачи $(\tilde{\mathfrak{z}})$ составим функцию Лагранжа, игнорируя ограничения второй группы:

$$\begin{aligned} \tilde{\mathscr{L}}(x, u, y^*, \lambda, \lambda_0) = \lambda_0 f_0(x) &+ \\ + \sum_{i=1}^{m} \lambda_i (f_i(x) + u_i) &+ \langle y^*, F(x)\rangle, \quad \lambda = (\lambda_1, \ldots, \lambda_m), \end{aligned}$$

условившись о знаке множителя λ_0 (в задаче на минимум $\lambda_0 \geqslant 0$, на максимум $\lambda_0 \leqslant 0$).

В задаче о минимизации функции Лагранжа (при фиксированных миожителях)

$$\tilde{\mathscr{L}}(x, u, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0) \to \inf$$

имеется две группы переменных x и u. Если закрепить переменные $u = \hat{u}$, то получается элементарная задача без ограничений (гладкая, выпуклая и т. п.), и здесь можно написать нужное необходимое условие экстремума, причем, как легко видеть, $\hat{u}_i$ в это условие не войдут. Если же фиксировать переменные $x = \hat{x}$, то получится элементарная задача линейного программирования. Условия экстремума, написанные в соответствии с п. 3.1.2, дают нам условия *соответствия знаков* множителей

Лагранжа $\hat{\lambda}$

$$\hat{\lambda}_i \gtreqless 0 \tag{1}$$

и условия *дополняющей нежесткости*

$$\hat{\lambda}_i \hat{u}_i = 0 \Leftrightarrow \hat{\lambda}_i f_i(\hat{x}) = 0. \tag{2}$$

Далее мы будем каждый раз пользоваться этой процедурой, но, если у нас имеются ограничения типа неравенств, составляя функцию Лагранжа, сразу будем писать ее укороченной: $\mathscr{L}(x, y^*, \lambda, \lambda_0) = \sum_{i=0}^{m} \lambda_i f_i(x) + \langle y^*, F(x) \rangle$ — и именно такую функцию будем называть *функцией Лагранжа* задачи ($\mathfrak{z}$). Нужно помнить только, что к условиям ее экстремума по x следует присоединить соотношения, вытекающие из (1), (2), а именно, условия *согласования знаков*

$$f_i(x) \lesseqgtr 0 \Rightarrow \hat{\lambda}_i \gtreqless 0 \tag{1'}$$

и дополняющей нежесткости

$$\hat{\lambda}_i f_i(\hat{x}) = 0.$$

Заметим теперь, что принцип Лагранжа (как, например, и теорема Ферма для гладкой экстремальной задачи) дает только необходимые условия экстремума, т. е. выделяет множество «подозреваемых» объектов, но не доказывает их «виновности». Поэтому для полного решения задачи мы должны либо иметь в распоряжении набор достаточных условий экстремума, либо иметь уверенность в существовании решения.

Первое позволит «провести экспертизу»: каждый из выделенных необходимыми условиями объектов мы подвергаем проверке на достаточность. Найдя среди них тот, который удовлетворяет достаточным условиям, мы считаем задачу решенной.

Во втором случае искомое решение (существование которого заранее известно или доказано) обязано попасть в число подозреваемых объектов, удовлетворяющих необходимым условиям экстремума. Вычисляя для каждого из них значение функционала, мы объявляем решением тот, где это значение экстремально.

Разумеется, ни тому, ни другому приему нельзя отдать предпочтения. Достаточные условия обычно не совпадают с необходимыми, и потому могут оставаться «подозреваемые», «вину» которых установить не удается (например, теоремы п. 3.1.1 не дают ответа о существовании экстремума в точке $\hat{x}$ для «плоской» функции $f(x)$, у которой $f^{(k)}(\hat{x})=0$ при всех k). С другой стороны, даже зная, что решение существует, мы можем оказаться в затруднении, если необходимыми условиями выделяется бесконечное (или просто очень большое) множество подозреваемых объектов.

В большинстве вопросов существования решения оказывается возможным обойтись усовершенствованным вариантом классической теоремы Вейерштрасса, которую мы уже упоминали в п. 1.6.1: существование является следствием компактности множества допустимых элементов и полунепрерывности функционала.

Определение. Функция $f\colon X \to \mathbf{R}$, определенная на топологическом пространстве X, называется *полунепрерывной снизу* в точке x_0, если $\underline{\lim}_{x \to x_0} f(x) \geqslant f(x_0)$, и просто *полунепрерывной снизу*, если она полунепрерывна снизу в каждой точке (ср. п. 2.6.2).

Теорема Вейерштрасса. *Полунепрерывная снизу функция $f\colon X \to \mathbf{R}$ достигает минимума на всяком счетно-компактном подмножестве топологического пространства X.*

Доказательство. Пусть $A \subset X$ — счетно-компактное подмножество и S — значение задачи

$$f(x) \to \inf, \quad x \in A, \tag{3}$$

т. е.

$$S = \inf_{x \in A} f(x).$$

По определению нижней грани мы можем выбрать минимизирующую последовательность задачи (3), т. е. такую последовательность точек $x_n \in A$, что $f(x_n) \to S$. По определению счетной компактности из x_n можно выбрать подпоследовательность x_{n_k}, сходящуюся к некоторой точке $\hat{x} \in A$. Ввиду полунепрерывности

$$f(\hat{x}) \leqslant \underline{\lim}_{k \to \infty} f(x_{n_k}) = \lim_{n \to \infty} f(x_n) = S$$

и так как, с другой стороны, $f(\hat{x})$ не может быть меньше значения задачи (3), $f(\hat{x}) = S$, т. е. $\hat{x}$ — точка минимума. ■

Следствие. *Пусть f полунепрерывна снизу на топологическом пространстве X. Если некоторое лебеговское* *множество $\mathscr{L}_\alpha f=\{x \mid f(x)\leqslant\alpha\}$ функции f непусто и счетно компактно, то f достигает на X своего минимума.*

§ 3.2. Принцип Лагранжа для гладких задач с ограничениями типа равенств и неравенств

3.2.1. Формулировка теоремы. Рассмотрим экстремальную задачу:

$$f_0(x)\to\operatorname{extr};\quad F(x)=0,\quad f_i(x)\lesseqgtr 0,\quad i=1,\ \ldots,\ m. \tag{1}$$

Символ $f_i(x)\lesseqgtr 0$ означает, что i-е ограничение имеет вид либо $f_i(x)=0$, либо $f_i(x)\leqslant 0$, либо $f_i(x)\geqslant 0$.

Задачи типа (1), где X и Y — нормированные пространства, f_i — гладкие функции на X, а F — гладкое отображение из X в Y, называются *гладкими экстремальными задачами с ограничениями типа равенств и неравенств*.

Функцией Лагранжа задачи (1) называется функция

$$\mathscr{L}(x,\ y^*,\ \lambda,\ \lambda_0)=\sum_{i=0}^{m}\lambda_i f_i(x)+\langle y^*,\ F(x)\rangle, \tag{2}$$

где в (2) $\lambda=(\lambda_1,\ \ldots,\ \lambda_m)\in\mathbb{R}^{m*}$, $\lambda_0\in\mathbb{R}$, $y^*\in Y^*$ — *множители Лагранжа*.

Теорема (правило множителей Лагранжа для гладких задач с равенствами и неравенствами). *Пусть X и Y — банаховы пространства, U — открытое множество в X, функции f_i: $U\to\mathbb{R}$, $i=0,\ 1,\ \ldots,\ m$, и отображение F: $U\to Y$ строго дифференцируемы в точке $\hat{x}$.*

Если $\hat{x}$ доставляет локальный экстремум задаче (1) *и если*

$$\textit{образ } \operatorname{Im} F'(\hat{x}) \textit{ есть замкнутое подпространство в } Y, \tag{3}$$

то найдутся такие множители Лагранжа $\hat{y}^*$, $\hat{\lambda}$, $\hat{\lambda}_0$, для которых выполняются:

а) *условие стационарности функции Лагранжа по x*

$$\mathscr{L}_x(\hat{x},\ \hat{y}^*,\ \hat{\lambda},\ \hat{\lambda}_0)=0. \tag{4}$$

б) *условие согласования знаков: $\hat{\lambda}_0\geqslant 0$, если задача на минимум, $\hat{\lambda}_0\leqslant 0$, если задача на максимум,*

$$\hat{\lambda}_i\gtreqless 0,\quad i=1,\ \ldots,\ m, \tag{5}$$

в) *условия дополняющей нежесткости*

$$\hat{\lambda}_i f_i(\hat{x}) = 0, \quad i = 1, \ldots, m. \tag{6}$$

Напомним еще, что запись $\hat{\lambda}_i \gtrless 0$ означает следующее: если в формуле (1) $f_i(x) \geqslant 0$, то $\hat{\lambda}_i \leqslant 0$, если $f_i(x) \leqslant 0$, то $\hat{\lambda}_i \geqslant 0$, наконец, если $f_i(x) = 0$, то $\hat{\lambda}_i$ может иметь любой знак.

Утверждение о существовании множителей Лагранжа, удовлетворяющих совокупности условий а)—в), находится в точном соответствии с общим приемом снятия ограничений—об этом как раз и говорилось в последнем пункте предыдущего параграфа. Поэтому сформулированный результат мы называем еще *принципом Лагранжа для гладких задач с равенствами и неравенствами*.

Доказательство общей теоремы базируется, с одной стороны, на теореме о неявной функции из § 2.3, а с другой—на теореме Куна—Таккера, т. е. в конечном счете—на конечномерной теореме отделимости. Теорема Куна—Таккера была доказана нами в § 1.3, и ссылка на нее—единственная по существу содержательная ссылка на первую главу в этой части книги. Привлечение теоремы Куна—Таккера связано с наличием в (1) неравенств, и оно несколько осложняет доказательство. Случай же, когда неравенства отсутствуют, совсем прост, но вместе с тем и содержателен и поучителен. Поэтому мы разберем его отдельно, хотя, конечно, этот случай является автоматическим следствием общего результата. При чтении следующего пункта мы рекомендуем сравнивать параллельно бесконечномерный вариант доказательства с конечномерным, разобранным в п. 1.3.2.

3.2.2. Правило множителей для гладких задач с равенствами.

Теорема (принцип Лагранжа для гладких задач с равенствами).

а) *Пусть X и Y—банаховы пространства, U—открытое множество в X, f: $U \to \mathbf{R}$, F: $U \to Y$—функция и отображение, строго дифференцируемые в точке $\hat{x}$.*

Если точка $\hat{x}$ является точкой локального экстремума в задаче

$$f(x) \to \text{extr}, \quad F(x) = 0 \tag{1}$$

и если образ $\operatorname{Im} F'(x)$*—замкнутое подпространство в* Y, *то найдутся такие множители Лагранжа* $\hat{\lambda}_0 \in \mathbf{R}$ *и* $\hat{y}^* \in Y^*$, *для которых выполнено условие стационарности функции Лагранжа*:

$$\mathscr{L}_x(\hat{x}, \hat{y}^*, \hat{\lambda}_0)=0 \Leftrightarrow \langle\hat{\lambda}_0 f'(\hat{x}), x\rangle + \langle\hat{y}^*, F'(\hat{x})[x]\rangle=0, \forall x. \quad (2)$$

б) *Если выполнено условие регулярности отображения* F, *т. е. если* $\operatorname{Im} F'(\hat{x})$ *совпадает со всем пространством* Y, *то множитель* $\hat{\lambda}_0$ *отличен от нуля.*

Доказательство проведем для задачи на минимум. Определим отображение

$$\mathscr{F}(x) = (f(x) - f(\hat{x}), F(x)), \quad \mathscr{F}: U \to \mathbf{R} \times Y.$$

Отображение $\mathscr{F}$, очевидно, строго дифференцируемо в $\hat{x}$ и $\mathscr{F}'(\hat{x}) = (f'(\hat{x}), F'(\hat{x}))$.

Возможно одно из двух: образ $\operatorname{Im}\mathscr{F}'(\hat{x})$ может совпадать или не совпадать с пространством $\mathbf{R}\times Y$.

А) Разберем сначала случай $\operatorname{Im}\mathscr{F}'(\hat{x}) \neq \mathbf{R}\times Y$. К отображению $\mathscr{F}'(\hat{x})$ применим лемму о замкнутости образа (п. 2.1.6). Образ $\operatorname{Im} F'(\hat{x})$ замкнут по условию, образ $f'(\hat{x})(\operatorname{Ker} F'(\hat{x}))$ есть либо $\{0\}$, либо $\mathbf{R}$, т. е. замкнутое подмножество $\mathbf{R}$. Значит, по цитированной лемме образ $\operatorname{Im}\mathscr{F}'(\hat{x})$ замкнут в $\mathbf{R}\times Y$. Так как он не совпадает с $\mathbf{R}\times Y$, $\operatorname{Im}\mathscr{F}'(\hat{x})$ — собственное замкнутое подпространство. Но по лемме о нетривиальности аннулятора (п. 2.1.4) найдутся число $\hat{\lambda}_0$ и элемент $\hat{y}^*$ (см. п. 2.1.2, где говорится об общем виде линейного функционала в произведении пространств), не равные нулю одновременно ($|\hat{\lambda}_0| + \|\hat{y}^*\| \neq 0$) и такие, что

$$\langle\hat{\lambda}_0 f'(\hat{x}), x\rangle + \langle\hat{y}^*, F'(\hat{x})[x]\rangle = 0, \quad \forall x.$$

Но это соотношение совпадает с (2).

Б) Пусть теперь $\operatorname{Im}\mathscr{F}'(\hat{x}) = \mathbf{R}\times Y$. Тогда мы можем применить теорему о существовании неявной функции (см. пп. 2.3.1—2.3.3).

Согласно этой теореме[1]) существуют константа $K > 0$, окрестность $\mathscr{U}$ точки $(0, 0)$ в пространстве $\mathbf{R}\times Y$ и

[1]) Приведем сопоставление обозначений теоремы п. 2.3.1 с обозначениями этого пункта. X — топологическое пространство в теореме о неявной функции (н ф) здесь состоит из одной точки $\{x_0\}$, Y (н.ф) $\Leftrightarrow X$, Z(н.ф) $\Leftrightarrow \mathbf{R}\times Y$, $\Psi(x, y) = \Psi(x_0, y)$ (н ф) $\Leftrightarrow \mathscr{F}(x)$, Λ (н ф.) $\Leftrightarrow$ $\Leftrightarrow \mathscr{F}'(\hat{x})$, y_0 (н.ф.) $\Leftrightarrow \hat{x}$, z_0 (н.ф) $\Leftrightarrow 0 \in \mathbf{R}\times Y$. Условие 1) теоремы

отображение $\varphi\colon \mathcal{U} \to X$ такие, что

$$\mathcal{F}(\varphi(\alpha, y)) = (\alpha, y) \quad \text{и} \quad \|\varphi(\alpha, y) - \hat{x}\| \leqslant K \|\mathcal{F}(\hat{x}) - (\alpha, y)\|. \tag{3}$$

Положим $x(\varepsilon) = \varphi(-\varepsilon, 0) = \varphi(z(\varepsilon))$. Тогда из (3) имеем

$$\mathcal{F}(x(\varepsilon)) = z(\varepsilon) \Leftrightarrow f(x(\varepsilon)) - f(\hat{x}) = -\varepsilon, \quad F(x(\varepsilon)) = 0, \tag{4}$$

$$\|x(\varepsilon) - \hat{x}\| = \|\varphi(z(\varepsilon)) - \hat{x}\| \leqslant K \|(0, 0) - (-\varepsilon, 0)\| = K|\varepsilon|. \tag{5}$$

Из (4), (5) следует, что $x(\varepsilon)$ — допустимый элемент в задаче (1), сколь угодно близкий к $\hat{x}$ и вместе с тем $f(x(\varepsilon)) < f(\hat{x})$, т. е. $\hat{x} \notin \operatorname{locmin}(1)$. Противоречие доказывает, что равенство $\operatorname{Im}\mathcal{F}'(\hat{x}) = \mathbf{R} \times Y$ невозможно. Тем самым верно утверждение а) теоремы.

В) Пусть F — регулярно в точке $\hat{x}$ и $\hat{\lambda}_0 = 0$. Тогда $\hat{y}^* \neq 0$ и соотношение (2) приобретает вид: $\langle y^*, F'(\hat{x})[x]\rangle = 0$, $\forall x$. Выберем элемент $\tilde{y}$ такой, что $\langle \hat{y}^*, \tilde{y}\rangle \neq 0$ (это возможно, ибо $\hat{y}^* \neq 0$), и найдем элемент $\tilde{x}$ такой, что $F'(\hat{x})[\tilde{x}] = \tilde{y}$ (он существует ввиду равенства $F'(\hat{x})X = Y$). Тогда $0 \neq \langle \hat{y}^*, \tilde{y}\rangle = \langle \hat{y}^*, F'(\hat{x})[\tilde{x}]\rangle = 0$. Этим противоречием доказано второе утверждение. ■

З а м е ч а н и е. Мы построили доказательство по той же схеме, что и доказательство конечномерного правила множителей Лагранжа (п. 1.3.2), и оно оказалось столь же простым и кратким. Правда, это потребовало некоторой подготовки: были использованы три факта из функционального анализа: теорема о нетривиальности аннулятора (т. е. в конечном счете — теорема Хана — Банаха), лемма о замкнутости образа и теорема о неявной функции. Для завершения общей теоремы — доказательства правила множителей для задач с неравенствами — нам, помимо этих трех фактов, понадобится еще лишь теорема Куна — Таккера, являющаяся прямым следствием конечномерной теоремы отделимости (см. п. Е) доказательства теоремы в п. 3.2.4). Удивительно, что столь малыми средствами, относящимися к общим математическим структурам, получается результат, в качестве прямого следствия из которого с помощью незначительных вспомогательных средств также общего плана (теорема Рисса о линейных функционалах в пространстве C и стандартные теоремы о дифференциальных уравнениях) в следующей главе будут выведены содержательные конкретные результаты: необходимые условия экстремума в задачах классического вариационного исчисления и оптимального управления.

удовлетворяется тривиально, условие 2) выполнено вследствие строгой дифференцируемости $\mathcal{F}$, ибо

$$\Psi(x_0, y') - \Psi(x_0, y'') - \Lambda(y' - y'') \Leftrightarrow \mathcal{F}(x') - \mathcal{F}(x'') - \mathcal{F}'(\hat{x})[x' - x''],$$

и, наконец, условие 3) выполнено, поскольку $\operatorname{Im}\mathcal{F}'(\hat{x}) = \mathbf{R} \times Y$.

3.2.3. Редукция задачи. Перед тем, как приступить к доказательству теоремы, сформулированной в п. 3.2.1, произведем некоторое преобразование ее условий. Сначала, заменив, если нужно, f_0 на $(-1)f_0$, *сведем задачу к задаче на минимум*. Если среди ограничений вида $f_i(x) \leqslant 0$ или $f_i(x) \geqslant 0$ есть такие, что $f_i(\hat{x}) < 0$ или соответственно $f_i(\hat{x}) > 0$, то мы их отбросим, поскольку с локальной точки зрения они несущественны, ибо выполняются во всех точках некоторой окрестности точки $\hat{x}$. Далее, если в (1) п. 3.2.1 имеются неравенства $f_i(x) \geqslant 0$, $f_i(\hat{x}) = 0$, мы *заменим их на неравенства* $(-1)f_i(x) \leqslant 0$. Равенства же $f_i(x) = 0$ *присоединим к равенству* $F(x) = 0$. Перенумеровав теперь заново все неравенства, задающие ограничения, получим следующую задачу, эквивалентную (1):

$$\tilde{f}_0(x) \to \inf; \quad \tilde{F}(x) = 0, \quad \tilde{f}_j(x) \leqslant 0, \quad j = 1, \ldots, \tilde{m}, \tag{1'}$$

где $\tilde{F}(x) = (F(x), \tilde{f}_{\tilde{m}+1}(x), \ldots, \tilde{f}_\mu(x))$ и для каждого j, $0 \leqslant j \leqslant \mu$, существует такое $\varepsilon_j = +1$ или -1, что

$$\tilde{f}_j(x) \equiv \varepsilon_j f_{i_j}(x). \tag{2'}$$

В задаче (1′) $\hat{x}$ доставляет локальный минимум и при этом $\tilde{f}_j(\hat{x}) = 0$, $j = 1, \ldots, \tilde{m}$.

Покажем теперь, что для задачи (1′) выполнены все условия теоремы. Действительно, X и $Y \times \mathbf{R}^{\mu-\tilde{m}}$ — банаховы пространства. Функции $\tilde{f}_j$ и отображение $\tilde{F}$ строго дифференцируемы в точке $\hat{x}$ (в силу того, что F и f_j были таковыми). Осталось показать лишь, что $L = \operatorname{Im} \tilde{F}'(\hat{x})$ замкнуто в $Y \times \mathbf{R}^{\mu-\tilde{m}}$. Имеем $\tilde{F}'(\hat{x}) = (F'(\hat{x}), \Phi'(\hat{x}))$, где $\Phi'(\hat{x})[x] = (\langle \tilde{f}'_{\tilde{m}+1}(\hat{x}), x\rangle, \ldots, \langle \tilde{f}'_\mu(\hat{x}), x\rangle)$, $\Phi'(\hat{x}): X \to \mathbf{R}^{\mu-\tilde{m}}$. Образ $\operatorname{Im} F'(\hat{x})$ замкнут по условию, а подпространство $\Phi'(\hat{x})[\operatorname{Ker} F'(\hat{x})] \subset \mathbf{R}^{\mu-\tilde{m}}$ замкнуто, как и любое подпространство конечномерного пространства. Значит, по лемме о замкнутости образа (п. 2.1.6) $\operatorname{Im} \tilde{F}'(\hat{x})$ замкнут в $Y \times \mathbf{R}^{\mu-\tilde{m}}$.

Предположим теперь, что для задачи (1′) доказываемая теорема верна и, следовательно, существуют множители Лагранжа $\tilde{y}^* = (\hat{y}^*, \tilde{\lambda}_{\tilde{m}+1}, \ldots, \tilde{\lambda}_\mu)$, $\tilde{\lambda}_j$, $j = 0, 1, \ldots, \tilde{m}$, для которых верны а) и б) п. 3.2.1 (условия в) выполняются автоматически, ибо $\tilde{f}_j(\hat{x}) = 0$, $j = 1, \ldots, \tilde{m}$).

В силу б) $\tilde{\lambda}_j \geqslant 0$, $j=0, 1, \ldots, \tilde{m}$, а в силу а)

$$\tilde{\mathscr{L}}_x(\hat{x}, \tilde{y}^*, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_{\tilde{m}}, \tilde{\lambda}_0) = \sum_{j=0}^{\tilde{m}} \tilde{\lambda}_j \tilde{f}'_j(\hat{x}) + (\tilde{y}^* \circ \tilde{F})'(\hat{x}) = 0, \quad (3')$$

где $\tilde{\mathscr{L}} = \sum_{j=0}^{\tilde{m}} \tilde{\lambda}_j \tilde{f}_j(x) + (\tilde{y}^* \circ \tilde{F})(x)$.

Теперь остается положить $\hat{\lambda}_0 = \tilde{\lambda}_0$, $\hat{\lambda}_{i_j} = \varepsilon_j \tilde{\lambda}_j$, $j=1, \ldots, \tilde{m}$, $\hat{\lambda}_{i_j} = \tilde{\lambda}_j$; $j = \tilde{m}+1, \ldots, \mu$ и $\hat{\lambda}_i = 0$, если $f_i(\hat{x}) \neq 0$. Полученный набор будет удовлетворять, очевидно, условиям дополняющей нежесткости и условию согласования знаков (ибо в (2') $\varepsilon_j = +1$, когда ограничение имеет вид $f_{i_j}(x) \leqslant 0$ и $\varepsilon_j = -1$ при $f_{i_j} \geqslant 0$). Кроме того, поскольку

$$\tilde{\mathscr{L}}(x, \hat{y}^*, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_\mu, \tilde{\lambda}_0) = \mathscr{L}(x, \hat{y}^*, \hat{\lambda}_1, \ldots, \hat{\lambda}_m, \hat{\lambda}_0)$$

условие стационарности функции Лагранжа выполняется или не выполняется для обеих функций одновременно.

Итак, в дальнейшем можно ограничиться лишь рассмотрением задачи (1'). Однако для простоты мы больше не будем ставить тильду ($\sim$) над F, f_j и m.

3.2.4. Доказательство теоремы. Итак, пусть нам дана задача:

$$f_0(x) \to \inf; \quad F(x) = 0, \quad f_j(x) \leqslant 0, \quad j = 1, \ldots, m, \quad (1)$$

причем выполнены все условия теоремы п. 3.2.1 и $f_j(\hat{x}) = 0$, $j = 1, \ldots, m$. Без ограничения общности можно считать также, что и $f_0(\hat{x}) = 0$.

Разобьем доказательство на несколько этапов.

А) Линейный случай. Рассмотрим сначала простейшую ситуацию, когда $f_0 = x^*$ — линейный функционал, $f_i = 0$, $i = 1, \ldots, m$, а $F = \Lambda$ — *линейный непрерывный сюръективный оператор*. Точка $\hat{x} = 0$ решение задачи

$$\langle x^*, x \rangle \to \inf; \ \Lambda x = 0 \quad (\Lambda \in \mathscr{L}(X, Y), \ \Lambda(X) = Y) \quad (2)$$

тогда и только тогда, когда $x^* \in (\operatorname{Ker} \Lambda)^\perp$. По лемме об аннуляторе ядра (п. 2.1.7) существует элемент $\hat{y}^* \in Y^*$ такой, что $x^* + \Lambda^* \hat{y}^* = 0$. Но это соотношение в точности есть принцип Лагранжа для задачи (2):

$$x^* + \Lambda^* \hat{y}^* = 0 \Leftrightarrow \mathscr{L}_x(0, \hat{y}^*, 1) = 0,$$

где $\mathscr{L}(x, y^*, \lambda_0) = \lambda_0 \langle x^*, x\rangle + \langle y^*, \Lambda x\rangle = \langle \lambda_0 x^* + \Lambda^* y^*, x\rangle$ — функция Лагранжа задачи (2). Таким образом, в этой ситуации принцип Лагранжа верен.

Б) Вырожденный случай. $\operatorname{Im} F'(\hat{x})$ есть собственное подпространство Y. Вследствие леммы о нетривиальности аннулятора (п. 2.1.4) найдется элемент $\hat{y}^* \in (\operatorname{Im} F'(\hat{x}))^\perp \Leftrightarrow (\hat{y}^* \circ F')(\hat{x}) = 0$. Остается положить $\hat{\lambda}_i = 0$, $i = 0, \ldots, m$, и убедиться в том, что принцип верен и в рассматриваемом вырожденном случае.

Далее считаем, что F — регулярный оператор в точке $\hat{x}$, т. е. $\operatorname{Im} F'(\hat{x}) = Y$.

Положим для $0 \leqslant k \leqslant m$

$$A_k = \{x \mid \langle f_i'(\hat{x}), x\rangle < 0,\ \ i = k, k+1, \ldots, m,\ F'(\hat{x})[x] = 0\}.$$

В) Лемма 1 (основная). *Если $\hat{x}$ есть локальное решение в задаче* (1), *то множество A_0 пусто. Иначе говоря,*

$$\max_{0 \leqslant i \leqslant m} \langle f_i'(\hat{x}), x\rangle \geqslant 0, \quad \forall x \in \operatorname{Ker} F'(\hat{x}).$$

Доказательство. Предположим, что A_0 непусто, т. е. существует элемент ξ такой, что $F'(\hat{x})[\xi] = 0$, $\langle f_i'(\hat{x}), \xi\rangle = \beta_i$, $\beta_i < 0$, $0 \leqslant i \leqslant m$. Тогда по теореме Люстерника (п. 2.3.5) существует отображение $r\colon [-\alpha, \alpha] \to X$, $\alpha > 0$, такое, что

$$F(\hat{x} + \lambda\xi + r(\lambda)) \equiv 0, \quad \lambda \in [-\alpha, \alpha], \quad r(\lambda) = o(\lambda). \tag{3}$$

При малых $\lambda > 0$ имеем для $i = 0, 1, \ldots, m$ неравенства

$$f_i(\hat{x} + \lambda\xi + r(\lambda)) = f_i(\hat{x}) + \lambda \langle f_i'(\hat{x}), \xi\rangle + o(\lambda) = \lambda\beta_i + o(\lambda) < 0. \tag{4_i}$$

Соотношения (3) и (4_i) $i = 1, \ldots, m$, означают, что для малых $\lambda > 0$ элемент $\hat{x} + \lambda\xi + r(\lambda)$ является допустимым в задаче (1). Но при этом неравенство (4_0) противоречит локальной минимальности $\hat{x}$. ■

Г) Лемма 2. *Если A_m есть пустое множество, то для задачи* (1) *верен принцип Лагранжа.*

Доказательство. Пустота A_m означает, что $x = 0$ является решением задачи

$$\langle f_m'(\hat{x}), x\rangle \to \inf, \quad F'(\hat{x})[x] = 0.$$

В силу п. А) для этой задачи справедлив принцип Лагранжа, а значит, он справедлив и для задачи (1) (надо только положить $\hat{\lambda}_0 = \ldots = \hat{\lambda}_{m-1} = 0$).

Таким образом, из пп. В) и Г) вытекает, что либо принцип Лагранжа уже обоснован ($A_m = \varnothing$), либо существует такое k, $0 \leqslant k < m$, что

$$A_k = \varnothing, \quad A_{k+1} \neq \varnothing. \tag{5}$$

Д) Лемма 3. *Если выполнены соотношения* (5), *то нуль является решением следующей задачи линейного программирования:*

$$\langle f'_k(\hat{x}), x\rangle \to \inf; \ \langle f'_i(\hat{x}), x\rangle \leqslant 0,$$
$$i = k+1, \ldots, m, \ F'(\hat{x})[x] = 0. \tag{6}$$

Доказательство. Предположим, что η — такой допустимый в задаче (6) элемент (т. е. $\langle f'_i(\hat{x}), \eta\rangle \leqslant 0$, $i \geqslant k+1$, $F'(\hat{x})[\eta] = 0$), что $\langle f'_k(\hat{x}), \eta\rangle < 0$. Пусть ζ — элемент, принадлежащий A_{k+1}, т. е. $\langle f'_i(\hat{x}), \zeta\rangle < 0$, $i \geqslant k+1$, $F'(\hat{x})[\zeta] = 0$. Тогда при малом $\varepsilon > 0$ элемент $\eta + \varepsilon\zeta$ принадлежит A_k в противоречии с (5). ■

Е) Завершение доказательства. Применяем к задаче (6) теорему Куна—Таккера (п. 1.3.3), учтя при этом, что условие Слейтера для этой задачи выполнено (из-за непустоты A_{k+1}). По этой теореме найдутся такие неотрицательные числа $\hat{\lambda}_{k+1}, \ldots, \hat{\lambda}_m$, что точка нуль является решением задачи

$$\langle f'_k(\hat{x}), x\rangle + \sum_{i=k+1}^{m} \langle \hat{\lambda}_i f'_i(\hat{x}), x\rangle \to \inf, \quad F'(\hat{x})[x] = 0. \tag{7}$$

(Это последнее утверждение есть не что иное, как принцип минимума для задачи (6), причем множитель Лагранжа при функционале взят равным 1 вследствие выполнимости условия Слейтера.) Но задача (7) есть опять-таки задача, о которой говорилось в п. А). Для нее верен принцип Лагранжа или, что в данной ситуации одно и то же, применима лемма об аннуляторе ядра оператора $F'(\hat{x})$. Иначе говоря, существует элемент $\hat{y}^*$, для которого

$$f'_k(\hat{x}) + \sum_{i=k+1}^{m} \hat{\lambda}_i f'_i(\hat{x}) + (\hat{y}^* \circ F')(\hat{x}) = 0.$$

Но это и есть условие стационарности функции Лагранжа, если положить $\hat{\lambda}_0 = \ldots = \hat{\lambda}_{k-1} = 0$, $\hat{\lambda}_k = 1$. ■

По ходу доказательства при $\hat{A}_1 \neq \varnothing$ оказалось, что $\hat{\lambda}_0 = 1$. Покажем, что если F регулярно (т. е. $\operatorname{Im} F'(\hat{x}) = Y$), то в этом случае вообще никакой набор множителей Лагранжа, для которого выполнено условие стационарности функции Лагранжа, не может содержать множитель $\hat{\lambda}_0 = 0$. Действительно, при $A_1 \neq \varnothing$ существует элемент h такой, что $F'(\hat{x})[h] = 0$, $\langle f_i'(\hat{x}), h\rangle < 0$, $i = 1, \ldots, \tilde{m}$. Допустим теперь, что нашлись множители Лагранжа $\tilde{\lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m)$, $\tilde{y}^*$, не равные нулю одновременно и такие, что $\mathscr{L}_x(\hat{x}, \tilde{y}^*, \tilde{\lambda}, 0) = 0$. Тогда ввиду неравенств $\tilde{\lambda}_i \langle f_i'(\hat{x}), h\rangle \leqslant 0$ имеем

$$0 = \mathscr{L}_x(\hat{x}, \tilde{y}^*, \tilde{\lambda}, 0)[h] = \sum_{i=1}^{m} \tilde{\lambda}_i \langle f_i'(\hat{x}), h\rangle + \langle \tilde{y}^*, F'(\hat{x})[h]\rangle =$$

$$= \sum_{i=1}^{m} \tilde{\lambda}_i \langle f_i'(\hat{x}), h\rangle \Rightarrow \tilde{\lambda}_1 = \ldots = \tilde{\lambda}_m = 0 \Rightarrow \tilde{y}^* \neq 0,$$

и теперь

$$0 = \mathscr{L}_x(\hat{x}, \tilde{y}^*, \tilde{\lambda}, 0)[x] = \langle \tilde{y}^*, F'(\hat{x})[x]\rangle, \quad \forall x \Rightarrow \operatorname{Im} F'(\hat{x}) \neq Y$$

вопреки предположению.

Выделим сказанное в отдельное предложение.

Предложение. *Для того чтобы в теореме* п. 3.2.1 $\hat{\lambda}_0 \neq 0$, *достаточно добавить к ее условиям, что* $\operatorname{Im} F'(\hat{x}) = Y$ *и существует элемент* $h \in \operatorname{Ker} F'(x)$, *для которого* $\langle f_i'(\hat{x}), h\rangle < 0$, $i = 1, \ldots, m$.

Дополнительные допущения, о которых здесь говорится, будем называть *условиями усиленной регулярности задачи* (1).

В формулировке доказанного нами принципа Лагранжа участвует важное (и вдобавок единственное, помимо требований гладкости и банаховости) условие замкнутости образа $\operatorname{Im} F'(\hat{x})$. Необходимо отметить, что без условий такого типа принцип Лагранжа может оказаться неверным.

Прежде всего при отказе от требований сюръективности оператора $\Lambda \in \mathscr{L}(X, Y)$ (X и Y — банаховы пространства) формула $(\operatorname{Ker} \Lambda)^{\perp} = \operatorname{Im} \Lambda^*$ может оказаться неверной, точнее, может оказаться, что $\operatorname{Im} \Lambda^*$ есть собственное подпространство $(\operatorname{Ker} \Lambda)^{\perp}$. Например, если $X = Y = l_2$, $x = (x_1, \ldots, x_n, \ldots) \in l_2$, $\Lambda x = (x_1, x_2/2, \ldots, x_n/n, \ldots)$, то $\operatorname{Ker} \Lambda = \{0\}$ и, значит, $(\operatorname{Ker} \Lambda)^{\perp} = l_2$, в то время как $\operatorname{Im} \Lambda = \operatorname{Im} \Lambda^* \neq l_2$ (скажем, элемент $\hat{y} = (1, 1/2, \ldots, 1/n, \ldots)$

принадлежит l_2, но решения уравнения $\Lambda x = \hat{y}$, очевидно, не существует). Теперь мы можем уже привести пример задачи, для которой принцип Лагранжа неверен.

Пример. Пусть X и Y — банаховы пространства и оператор $\Lambda \in \mathscr{L}(X, Y)$ таков, что $\operatorname{Ker}\Lambda^* = \{0\}$, а $\operatorname{Im}\Lambda^*$ есть собственное подпространство $(\operatorname{Ker}\Lambda)^\perp$. Выбрав $x^* \in (\operatorname{Ker}\Lambda)^\perp \setminus \operatorname{Im}\Lambda^*$, рассмотрим задачу

$$\langle x^*, x\rangle \to \inf; \quad \Lambda x = 0.$$

Для этой задачи принцип Лагранжа неверен. Действительно, $\hat{x} = 0$ является точкой минимума, и если бы нашлись $\hat{\lambda}_0$ и $\hat{y}^* \in Y^*$ такие, что

$$\mathscr{L}_x(0, \hat{y}^*, \hat{\lambda}_0)[h] = 0, \quad \forall h \in X \Leftrightarrow \hat{\lambda}_0 \langle x^*, h\rangle + \langle \hat{y}^*, \Lambda h\rangle = 0, \quad \forall h \in X,$$

то $\hat{\lambda}_0 = 0$ (ибо иначе $x^* \in \operatorname{Im}\Lambda^*$), а значит, $\Lambda^*\hat{y}^* = 0 \Rightarrow \hat{y}^* = 0$.

Упражнение[1]). Пусть $Z = Y = l_2$, $z = (z_1, \ldots, z_n, \ldots) \in l_2$, $\Lambda z = (z_1, z_2/2, \ldots, z_n/n, \ldots)$, $\hat{y} \notin \operatorname{Im}\Lambda$, $X = \mathbf{R} \times Z$, $f(x) = f(\alpha, z) = \alpha$, $F(x) = F(\alpha, z) = \Lambda z + \alpha^2 \hat{y}$. Покажите, что для задачи $f(x) \to \inf$; $F(x) = 0$ принцип Лагранжа неверен.

§ 3.3*. Принцип Лагранжа и двойственность в задачах выпуклого программирования

3.3.1. Теорема Куна — Таккера (субдифференциальная форма). Принцип Лагранжа для задач выпуклого программирования (теорема Куна — Таккера) был уже доказан в п. 1.3.3. В этом пункте дается «субдифференциальная форма» этой теоремы и проясняются связи с другими понятиями выпуклого анализа.

Пусть X и Y — банаховы пространства, $\Lambda: X \to Y$ — линейный непрерывный оператор, $f_i: X \to \overline{\mathbf{R}}$, $i = 0, 1, \ldots, m$ — выпуклые функции, $a = (a_1, \ldots, a_m) \in \mathbf{R}^m$, $b \in Y$, A — выпуклое множество в X. Рассмотрим следующую задачу выпуклого программирования:

$$f_0(x) \to \inf; \quad f_i(x) \leqslant a_i, \quad i = 1, \ldots, m, \quad \Lambda x = b, \quad x \in A. \qquad (\mathfrak{z})$$

Множество $\{x \mid \Lambda x = b\}$ обозначим через B. Функцией Лагранжа задачи ($\mathfrak{z}$) является функция

$$\mathscr{L}(x, y^*, \lambda, \lambda_0) = \lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - a_i) + \langle y^*, \Lambda x - b\rangle.$$

---

[1]) Предложено студентом 4 курса В. В. Успенским.

Предложение. *Пусть* $\hat{x}$ — *точка абсолютного минимума в задаче* ($\mathfrak{z}$). *Тогда* $\hat{x}$ — *точка абсолютного минимума в элементарной задаче*

$$f(x) = \max(f_0(x) - f_0(\hat{x}),\ f_1(x) - a_1,\ \ldots,\ f_m(x) - a_m) + \delta(A \cap B)(x) \to \inf, \qquad (\mathfrak{z}')$$

где $\delta(A \cap B)$ — *индикаторная функция множества* $A \cap B$.

Действительно, если существует элемент $\tilde{x}$, для которого $f(\tilde{x}) < 0$, то это означает, во-первых, что $\tilde{x} \in A$, $\tilde{x} \in B\ (\Leftrightarrow \Lambda\tilde{x} = b)$ во-вторых, что $f_i(\tilde{x}) < a_i$; $i = 1, \ldots, m$ (т. е. что $\tilde{x}$ — допустимый элемент в задаче), и, в-третьих, что $f_0(\tilde{x}) < f_0(\hat{x})$ вопреки условию. ■

Теорема (субдифференциальная форма теоремы Куна—Таккера). *Пусть в* ($\mathfrak{z}$) *функции* f_i, $i = 0, 1, \ldots, m$, *непрерывны в точке* $\hat{x} \in A \cap B$, *доставляющей абсолютный минимум в задаче. Тогда найдутся числа* $\hat{\lambda}_i \geqslant 0$ *такие, что* $\sum_{i=0}^{m} \hat{\lambda}_i = 1$, $\hat{\lambda}_i f_i(\hat{x}) = 0$, $i \geqslant 1$, *и элемент* $\hat{x}^* \in \partial\delta(A \cap B)(\hat{x})$, *для которых*

$$0 \in \sum_{i=0}^{m} \hat{\lambda}_i \partial f_i(\hat{x}) + \hat{x}^*. \qquad (1)$$

Доказательство. Согласно предложению $\hat{x}$ доставляет абсолютный минимум в элементарной задаче ($\mathfrak{z}'$). По теореме 1 п. 3.1.1 $0 \in \partial f(\hat{x})$. Функция $g(x) = \max(f_0(x) - f_0(\hat{x}),\ f_1(x) - a_1,\ \ldots,\ f_m(x) - a_m)$ выпукла и непрерывна в точке $\hat{x} \in A \cap B$ (т. е. принадлежащей $\operatorname{dom}\delta(A \cap B)$ и, значит, по теореме Моро—Рокафеллара (п. 2.6.4) $\partial f(\hat{x}) = \partial g(\hat{x}) + \partial\delta(A \cap B)(\hat{x})$. Наконец, по теореме Дубовицкого — Милютина (п. 2.6.4) $\partial g(\hat{x}) = \operatorname{conv}(\partial f_{i_1}(\hat{x}) \cup \ldots \cup \partial f_{i_s}(\hat{x}))$, где i_j — те и только те индексы, для которых $f_i(\hat{x}) = g(\hat{x})$. Таким образом, существуют два элемента $\hat{\xi}^* \in \partial g(\hat{x})$ и $\hat{x}^* \in \partial\delta(A \cap B)(\hat{x})$, для которых

$$\hat{\xi}^* + \hat{x}^* = 0,\ \hat{\xi}^* \in \operatorname{conv}(\partial f_{i_1}(\hat{x}) \cup \ldots \cup \partial f_{i_s}(\hat{x})) \Leftrightarrow \hat{\xi}^* = \sum_{j=1}^{s} \hat{\lambda}_{i_j} \hat{x}_{i_j}^*,\ \hat{x}_{i_j}^* \in \partial f_{i_j}(\hat{x}),\ \sum_{j=1}^{s} \hat{\lambda}_{i_j} = 1,\ \hat{\lambda}_{i_j} \geqslant 0.$$

Осталось положить $\hat{\lambda}_i = 0$, $i \notin \{i_1, \ldots, i_s\}$. ■

Упражнение. Пусть $B=\{x \mid \Lambda x=b\}$. Докажите, что если $x_0 \in B$, то $\partial\delta(B)(x_0)=(\operatorname{Ker}\Lambda)^{\perp}$.

Следствие 1 (принцип Лагранжа для задачи выпуклого программирования с ограничениями типа равенств и неравенств). *Пусть в условиях теоремы $A=X$ и образ X при отображении Λ замкнут в Y. Тогда найдутся такие множители Лагранжа $\hat{\lambda}_0 \in \mathbf{R}$, $\hat{\lambda} \in \mathbf{R}^{m*}$, $\hat{y}^* \in Y^*$, что*

$$\hat{\lambda}_i \geqslant 0, \quad i \geqslant 0, \quad \hat{\lambda}_i f_i(\hat{x})=0, \quad i \geqslant 1,$$
$$\min_x \mathscr{L}(x, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0)=\mathscr{L}(\hat{x}, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0). \qquad (2)$$

Действительно, если $\operatorname{Im}\Lambda$ есть собственное подпространство в Y, то можно положить $\hat{\lambda}_i=\hat{\lambda}_0=0$, $\hat{y}^* \in (\operatorname{Im}\Lambda)^{\perp}$. Если же Λ—сюръективный оператор, то $\partial\delta B(\hat{x})=(\operatorname{Ker}\Lambda)^{\perp}=\operatorname{Im}\Lambda^*$. Тогда, согласно (1), $0 \in \partial_x \mathscr{L}(\hat{x}, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0)=\sum_{i=0}^{m} \hat{\lambda}_i \partial f_i(\hat{x})+\operatorname{Im}\Lambda^*$, и, значит, по определению субдифференциала

$$\mathscr{L}(x, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0)-\mathscr{L}(\hat{x}, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0)-\langle 0, x-\hat{x}\rangle \geqslant 0 \Leftrightarrow$$
$$\Leftrightarrow \mathscr{L}(x, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0) \geqslant \mathscr{L}(\hat{x}, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0), \quad \forall x. \ \blacksquare$$

3.3.2. Метод возмущений и теорема двойственности. В предыдущем пункте задачу выпуклого программирования ($\mathfrak{z}$) мы рассматривали как одну индивидуальную задачу. Со многих точек зрения оказывается естественным и плодотворным рассмотрение целых семейств задач подобного рода. Фиксировав f_i, Λ, a_i, A и b, включим задачу ($\mathfrak{z}$) в семейство

$$f_0(x) \to \inf; \ f_i(x)+\alpha_i \leqslant a_i, \ i=1, \ldots, m,$$
$$\Lambda x+\eta=b, \quad x \in A. \qquad (\mathfrak{z}(\alpha, \eta)).$$

(Разумеется, можно было бы просто объявить a_i и b переменными параметрами семейства, но, введя параметры α_i, η так, как указано выше, мы получаем более красивые формулы.) Совокупность задач $\{(\mathfrak{z}(\alpha, \eta))\}$ назовем *возмущением* задачи $(\mathfrak{z})=(\mathfrak{z}(0, 0))$.

Мы уже видели в п. 3.3.1, что задача ($\mathfrak{z}$) сводится к элементарной. Сейчас мы проделаем то же самое для семейства $\mathfrak{z}(\alpha, \eta)$, однако несколько по-другому. А именно, обозначим

$$f(x;\ \alpha,\ \eta)=\begin{cases} f_0(x), \text{ если } f_i(x)+\alpha_i\leqslant a_i,\ \Lambda x+\eta=b,\ x\in A,\\ +\infty, \text{ в остальных случаях.}\end{cases} \tag{1}$$

Тогда семейство $(\mathfrak{z}(\alpha,\ \eta))$ можно записать в виде элементарной задачи

$$f(x;\ \alpha,\ \eta)\to\inf \qquad (\text{по } x\in X). \tag{2}$$

В дальнейшем, говоря о задаче $(\mathfrak{z}(\alpha,\ \eta))$, мы не будем различать ее исходную формулировку от (2). Значение этой задачи, т. е. $\inf f_0$, при указанных ограничениях является функцией от $\alpha=(\alpha_1,\ \ldots,\ \alpha_m)$ и η, которую мы обозначим через S, S: $\mathbf{R}^m\times Y\to\overline{\mathbf{R}}$ и иногда будем называть *S-функцией*:

$$S(\alpha,\ \eta)=\inf_x f(x;\ \alpha,\ \eta)=\inf_{x\in A,\ f_i(x)+\alpha_i\leqslant a_i,\ \Lambda x+\eta=b} f_0(x). \tag{3}$$

Лемма. *Пусть $F(x,\ z)$—выпуклая функция на произведении линейных пространств X и Z. Тогда функция*

$$S(z)=\inf_x F(x,\ z)$$

выпукла на Z.

Доказательство. Пусть $(z_i,\ t_i)\in\operatorname{epi}S$, $i=1,\ 2$ и $\lambda\in[0,\ 1]$. Тогда $S(z_i)\leqslant t_i$, $i=1,\ 2$, и для любого $\varepsilon>0$ существуют $(x_i,\ z_i)$ такие, что $F(x_i,\ z_i)<t_i+\varepsilon$, $i=1,\ 2$. Отсюда в силу выпуклости F

$$F(\lambda x_1+(1-\lambda)x_2,\ \lambda z_1+(1-\lambda)z_2)\leqslant$$
$$\leqslant\lambda F(x_1,\ z_1)+(1-\lambda)F(x_2,\ z_2)<\lambda(t_1+\varepsilon)+(1-\lambda)(t_2+\varepsilon)=$$
$$=\lambda t_1+(1-\lambda)t_2+\varepsilon.$$

Ввиду произвольности $\varepsilon>0$

$$F(\lambda x_1+(1-\lambda)x_2,\ \lambda z_1+(1-\lambda)z_2)\leqslant$$
$$\leqslant\lambda t_1+(1-\lambda)t_2\Rightarrow S(\lambda z_1+(1-\lambda)z_2)\leqslant$$
$$\leqslant\lambda t_1+(1-\lambda)t_2\Rightarrow(\lambda z_1+(1-\lambda)z_2,\ \lambda t_1+(1-\lambda)t_2)\in\operatorname{epi}S. \blacksquare$$

Предложение 1. *Пусть X и Y—линейные пространства, $A\subset X$—выпуклое множество, Λ: $X\to Y$—линейное отображение.*

Если функции f_i: $X\to\overline{\mathbf{R}}$, $i=0,\ 1,\ \ldots,\ m$, выпуклы, то функция $f(x;\ \alpha,\ \eta)$, определяемая равенством (1), выпукла на $X\times\mathbf{R}^m\times Y$.

Доказательство. Множества

$$M_0=\{(x,\ \alpha,\ \eta,\ t)\,|\,(x,\ t)\in \operatorname{epi} f_0\},$$
$$M_i=\{(x,\ \alpha,\ \eta,\ t)\,|\,f_i(x)+\alpha_i\leqslant a_i\},$$
$$M_A=\{(x,\ \alpha,\ \eta,\ t)\,|\,x\in A\},$$
$$M_\Lambda=\{(x,\ \alpha,\ \eta,\ t)\,|\,\Lambda x+\eta=b\}$$

выпуклы в $X\times \mathbf{R}^m\times Y\times \mathbf{R}$. Действительно (нарушая иногда для удобства порядок сомножителей): $M_0=\operatorname{epi} f_0\times \mathbf{R}^m\times Y$, $M_A=A\times \mathbf{R}^m\times Y\times \mathbf{R}$, $M_\Lambda=\tilde\Lambda^{-1}(b)\times \mathbf{R}^m\times \mathbf{R}$, где $\tilde\Lambda\colon (x,\ y)\mapsto \Lambda x+y$ — линейное отображение из $X\times Y$ в Y и $M_i=\sigma \operatorname{epi}\{f_i-a_i\}\times \mathbf{R}^{m-1}\times Y\times \mathbf{R}$, где $\sigma\colon\ (x,\ t)\mapsto (x,\ -t)$ симметрия в $X\times \mathbf{R}$. Отсюда видно, что все эти множества выпуклы, и остается заметить, что

$$\operatorname{epi} f=\bigcap_{i=0}^{m} M_i\cap M_A\cap M_\Lambda. \ \blacksquare$$

Следствие 1. *S-функция задачи* ($\mathfrak{z}(\alpha,\ \eta)$) *выпукла на* $\mathbf{R}^m\times Y$.

В § 2.6 мы уже видели, что выпуклость позволяет сопоставлять различным объектам (функциям, множествам) двойственные объекты в сопряженном пространстве. То же самое справедливо и для задач выпуклого программирования. Далее мы будем предполагать, что X и Y — локально выпуклые пространства, X^* и Y^* — их сопряженные, $\lambda=(\lambda_1,\ \ldots,\ \lambda_m)\in \mathbf{R}^{m*}$.

Определение 1. Семейство экстремальных задач

$$g(x^*;\lambda,\ \eta^*)\to \sup\,(\text{по } \lambda\in \mathbf{R}^{m*},\ \eta^*\in Y^*), \qquad (\mathfrak{z}^*(x^*))$$

где $(-1)\,g(x^*;\ \ \lambda,\ \ \eta^*)=f^*(x^*;\ \ \lambda,\ \ \eta^*)$ — преобразование Юнга—Фенхеля (п. 2.6.3) функции $(x,\ \alpha,\ \eta)\mapsto f(x,\ \alpha,\ \eta)$, определяемой равенством (1), называется *двойственным* к семейству $\mathfrak{z}(\alpha,\ \eta)\Leftrightarrow(2)$.

Задача $\mathfrak{z}^*=\mathfrak{z}^*(0)$ называется *двойственной* к задаче $\mathfrak{z}=\mathfrak{z}(0,\ 0)$ (относительно семейства возмущений $\mathfrak{z}(\alpha,\ \eta)$).

Значение задачи $\mathfrak{z}^*(x^*)$ обозначим

$$\Sigma(x^*)=\sup_{(\lambda,\ \eta^*)} g(x^*;\ \lambda,\ \ \eta^*). \qquad (4)$$

Поскольку функция f^* выпукла, противоположная ей функция g вогнута. Двойственные задачи можно было бы называть задачами вогнутого программирования, но мы предпочитаем обойтись без нового термина.

Определение 2. *Функцией Лагранжа* пары семейств $\mathfrak{z}(\alpha, \beta)$ и $\mathfrak{z}^*(x^*)$ или *расширенной функцией Лагранжа* задачи ($\mathfrak{z}$) называется функция $\tilde{\mathscr{L}}: X \times \mathbf{R}^{m*} \times Y^* \to \mathbf{R}$, определяемая соотношениями

$$\tilde{\mathscr{L}}(x;\ \lambda,\ \eta^*) = \begin{cases} f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - a_i) + \langle \eta^*,\ \Lambda x - b\rangle, \\ x \in A,\ \lambda \in \mathbf{R}_+^{m*}, \\ -\infty,\ x \in A,\ \lambda \notin \mathbf{R}_+^{m*}, \\ +\infty,\ x \notin A. \end{cases} \tag{5}$$

При фиксированных (λ, η^*) эта функция выпукла по x, а при фиксированном x выпукла по (λ, η^*) противоположная функция $-\tilde{\mathscr{L}}$.

Заметим, что

$$\tilde{\mathscr{L}}(x;\ \lambda,\ \eta^*) = \mathscr{L}(x,\ \eta^*,\ \lambda,\ 1), \quad x \in A, \quad \lambda \in \mathbf{R}_+^{m*}, \tag{6}$$

где $\mathscr{L}$ — функция Лагранжа задачи ($\mathfrak{z}$), определенная в п. 3.3.1.

Предложение 2. *Пара двойственных семейств $\mathfrak{z}(\alpha, \eta)$ и $\mathfrak{z}^*(x^*)$ однозначно определяется их функцией Лагранжа, поскольку*

$$f(x;\ \alpha,\ \eta) = (-\tilde{\mathscr{L}})^{*(2)}, \quad g(x^*;\ \lambda,\ \eta^*) = -\tilde{\mathscr{L}}^{*(1)}, \tag{7}$$

где $*(1)$ и $*(2)$ обозначают преобразования Юнга—Фенхеля по аргументам x и (λ, η^*) соответственно.

Доказательство. По определению преобразования Юнга—Фенхеля

$$\begin{aligned} -g(x^*;\ \lambda,\ \eta^*) &= f^*(x^*;\ \lambda,\ \eta^*) = \\ &= \sup_{(x,\ \alpha,\ \eta)} \{\langle x^*,\ x\rangle + \lambda\alpha + \langle \eta^*,\ \eta\rangle - f(x;\ \lambda,\ \eta)\} = \\ &= \sup_{\substack{x \in A,\ \Lambda x + \eta = b \\ f_i(x) + \alpha_i \leqslant a_i}} \{\langle x^*,\ x\rangle + \lambda\alpha + \langle \eta^*,\ \eta\rangle - f_0(x)\} = \end{aligned}$$

$$= \begin{cases} \sup\limits_{x \in A} \left\{ \langle x^*, x\rangle - f_0(x) - \sum_{i=1}^{m} \lambda_i (f_i(x) - a_i) - \langle \eta^*, \Lambda x - b\rangle \right\}, \\ \lambda \in \mathbf{R}_+^{m*} \\ +\infty,\ \lambda \notin \mathbf{R}_+^{m*} \end{cases} =$$

$$= \tilde{\mathscr{L}}^{*(1)}(x^*;\ \lambda,\ \eta^*),$$

чем доказано второе из равенств (7). Попутно мы получили полезное соотношение

$$g(x^*;\lambda,\eta^*)=\begin{cases}\inf\limits_{x\in A}(\tilde{\mathscr{L}}(x;\ \lambda,\ \eta^*)-\langle x^*,\ x\rangle),\ \lambda\in\mathbf{R}_+^{m*},\\ -\infty,\quad \lambda\notin\mathbf{R}_+^{m*}.\end{cases}\tag{8}$$

С другой стороны[1]),

$$(-\tilde{\mathscr{L}})^{*(2)}=\sup_{(\lambda,\ \eta^*)}(\lambda\alpha+\langle\eta^*,\ \eta\rangle+\tilde{\mathscr{L}}(x;\ \lambda,\ \eta^*))=$$

$$=\begin{cases}+\infty,\ x\notin A,\\ \sup\limits_{\lambda\geqslant 0,\ \eta^*}(\lambda\alpha+\langle\eta^*,\ \eta\rangle+f_0(x)+\\ \quad+\sum\limits_{i=1}^{m}\lambda_i(f_i(x)-a_i)+\langle\eta^*,\Lambda x-b\rangle)\end{cases}=$$

$$=\begin{cases}+\infty,\text{ если } x\notin A \text{ или } \Lambda x-b\neq-\eta \text{ или } f_i(x)-a_i>-\alpha_i\\ \qquad\text{для некоторого } i,\\ f_0(x)\quad\text{в остальных случаях}\end{cases}=$$

$$=f(x;\ \alpha,\ \eta).\ \blacksquare$$

Следствие 2. *Сопряженная функция к S-функции задачи* $\mathfrak{z}(\alpha,\ \eta)$ *имеет вид*

$$S^*(\lambda,\ \eta^*)=-g(0;\ \lambda,\ \eta^*)=\begin{cases}-\inf\limits_{x\in A}\mathscr{L}(x,\eta^*,\lambda,1),\ \lambda\in\mathbf{R}_+^{m*},\\ +\infty,\ \lambda\notin\mathbf{R}_+^{m*}.\end{cases}\tag{9}$$

Доказательство. По определению

$$S^*(\lambda,\ \eta^*)=\sup_{(\alpha,\ \eta)}(\lambda\alpha+\langle\eta^*,\ \eta\rangle-\inf_x f(x;\ \alpha,\ \eta))=$$
$$=\sup_{(\alpha,\ \eta,\ x)}(\lambda\alpha+\langle\eta^*,\ \eta\rangle-f(x;\ \alpha,\ \eta))=-g(0;\ \lambda,\ \eta^*),$$

после чего (9) следует из (6) и (8).

Таким образом, двойственная к задаче $\mathfrak{z}=\mathfrak{z}(0,\ 0)$ задача $\mathfrak{z}^*=\mathfrak{z}^*(0)$ может быть сформулирована также и так:

$$-S^*(\lambda,\ \eta^*)=\inf\mathscr{L}(x,\ \eta^*,\ \lambda,\ 1)\to\sup;\quad \lambda\in\mathbf{R}_+^{m*},\ \ \eta^*\in Y^*.\tag{10}$$

Замечание. Определение двойственной задачи зависит, вообще говоря, от того, в какое семейство

[1]) Как и в п. 2.6.3, вычисляя сопряженную к функции, определенной на сопряженном пространстве (в нашем случае на $\mathbf{R}^{m*}\times Y^*$), мы считаем результат функцией на исходном пространстве, а не на втором сопряженном.

возмущений мы включим исходную задачу ($\mathfrak{z}$). Равенства (5) и (6), определяющие расширенную функцию Лагранжа и равенства (7), показывают тем не менее, что семейства $\{(\mathfrak{z}(\alpha, \eta))\}$ и $\{(\mathfrak{z}^*(x^*))\}$ в некотором смысле естественно соответствуют задаче ($\mathfrak{z}$) и, таким образом, задача ($\mathfrak{z}^*$), эквивалентная (10), в том же смысле является ее естественной двойственной.

Упражнение. Покажите, что если функции f_i и множество A выпуклы и замкнуты и если $f_0(x)$ не обращается в $-\infty$ на множестве допустимых x задачи ($\mathfrak{z}$), то двойственное (с учетом сноски к доказательству предложения 2) к семейству $\{(\mathfrak{z}^*(x^*))\}$ семейство совпадает с $\{(\mathfrak{z}(\alpha, \eta))\}$, и, таким образом, задачи ($\mathfrak{z}$) и ($\mathfrak{z}^*$) образуют пару двойственных друг другу задач.

Теорема двойственности для задач выпуклого программирования. *Предположим, что S-функция семейства $\{(\mathfrak{z}(\alpha, \eta))\}$ непрерывна в точке* $(0, 0)$. *Тогда: для любых* $(\alpha, \eta) \in \operatorname{int}(\operatorname{dom} S)$

$$S(\alpha, \eta) = \sup_{\lambda \geqslant 0,\ \eta^* \in Y^*} \left(\lambda\alpha + \langle\eta^*, \eta\rangle + \inf_{x \in A} \mathscr{L}(x, \eta^*, \lambda, 1)\right). \quad (11)$$

Значения задачи ($\mathfrak{z}$) *и двойственной ей задачи* ($\mathfrak{z}^*$) *равны между собой*

$$\inf_{\substack{x \in A \\ \Lambda x = b \\ f_i(x) \leqslant a_i,\ i = 1, \ldots, m}} f_0(x) = S(0, 0) = \Sigma(0) =$$

$$= \sup_{\substack{\lambda \geqslant 0 \\ \eta^* \in Y^*}} \left\{\inf_{x \in A}\left\{f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - a_i) + \langle\eta^*, \Lambda x - b\rangle\right\}\right\}. \quad (12)$$

Доказательство. По следствию 1 функция S выпукла, а из непрерывности ее в одной точке вытекает (предложение 3 п. 2.6.2) непрерывность и равенство $S(\alpha, \eta) = (\overline{\operatorname{conv} S})(\alpha, \eta)$ для всех $(\alpha, \eta) \in \operatorname{int}(\operatorname{dom} S)$. По теореме Фенхеля—Моро (п. 2.6.3) и следствию 2

$$S(\alpha, \eta) = S^{**}(\alpha, \eta) = \sup_{(\lambda, \eta^*)} (\lambda\alpha + \langle\eta^*, \eta\rangle - S^*(\lambda, \eta^*)) =$$

$$= \sup_{\substack{\lambda \geqslant 0 \\ \eta^* \in Y^*}} \left(\lambda\alpha + \langle\eta^*, \eta\rangle + \inf_{x \in A} \mathscr{L}(x, \eta^*, \lambda, 1)\right),$$

чем доказано (11). Подставляя $\alpha = 0$, $\eta = 0$, получаем (12). ■

С л е д с т в и е 3 (т е о р е м а о м и н и м а к с е). *В условиях теоремы двойственности справедливо соотношение*

$$\sup_{\substack{\lambda \geqslant 0 \\ \eta^* \in Y^*}} \inf_{x \in A} \mathscr{L}(x, \eta^*, \lambda, 1) = \inf_{x \in A} \sup_{\substack{\lambda \geqslant 0 \\ \eta^* \in Y^*}} \mathscr{L}(x, \eta^*, \lambda, 1). \quad (13)$$

Д о к а з а т е л ь с т в о. Левая часть (13) равна $\Sigma(0)$ и в силу (12) совпадает с $S(0, 0)$. Что же касается правой части, то

$$\inf_{x \in A} \sup_{\substack{\lambda \geqslant 0 \\ \eta^* \in Y^*}} \mathscr{L}(x, \eta^*, \lambda, 1) = \inf_{x \in A} \sup_{(\lambda, \eta^*)} \tilde{\mathscr{L}}(x; \lambda, \eta^*) =$$

$$= \inf_{x \in A} ((-\tilde{\mathscr{L}})^{*(2)}(x; 0, 0)) = \inf_{x \in A} f(x; 0, 0)$$

согласно (7) и также совпадает с $S(0, 0)$.

3.3.3. Линейное программирование: теорема существования и теорема двойственности. Пусть X — линейное пространство, X' — алгебраически сопряженное с ним (т. е. совокупность *всех* линейных функционалов на X), K — некоторый полиэдральный конус (т. е. пересечение конечного числа полупространств $H_j = \{\langle x'_j, x\rangle \leqslant 0, x'_j \in X', j = 1, \ldots, s\}$ — иногда, впрочем, от условия полиэдральности отказываются).

Экстремальные задачи вида

$$\langle x'_0, x\rangle \to \inf; \ f_i(x) = \langle x'_i, x\rangle \geqslant b_i, \ i = 1, \ldots, m, \ x \in K, \quad (1)$$

составляют отдельный класс задач *линейного программирования*.

В этом пункте будет рассмотрен *конечномерный случай*, когда $X = \mathbf{R}^n$, $X' = \mathbf{R}^{n*}$, $K = \mathbf{R}^n_+$. В этом случае задачу (1) можно переписать так[1]):

$$cx \to \inf; \quad Ax \geqslant b, \quad x \geqslant 0, \quad (2)$$

где $c \in \mathbf{R}^{n*}$, $A = (a_{ij})$, $j = 1, \ldots, n$, $i = 1, \ldots, m$, — матрица, задающая линейный оператор из $\mathbf{R}^n$ в $\mathbf{R}^m$, $b \in \mathbf{R}^m$. Символ $z' \geqslant z$ для конечномерных векторов z' и z означает, что все координаты вектора не превосходят соответствующих координат вектора z'. Задачу (2) назовем *конечномерной задачей линейного программирования*. Эта

1) Напомним, что для конечномерных пространств мы имеем два обозначения:

$$\langle c, x\rangle = cx = \sum_{i=1}^{n} c_i x_i, \ c \in \mathbf{R}^{n*}, \ x \in \mathbf{R}^n.$$

задача — частный случай общей задачи выпуклого программирования. Однако в ней имеется еще более специальная структура: *все функции* здесь *линейны и конус полиэдрален*. Как и обычно, условимся, что если задача (2) несовместна, т. е. не имеет допустимых элементов, то ее значение считается равным $+\infty$.

Теорема существования. *Если множество допустимых значений конечномерной задачи линейного программирования* (2) *непусто и ее значение конечно, то задача имеет решение.*

Доказательство теоремы 1 опирается на один простой факт конечномерной геометрии. Напомним, что коническая оболочка cone C некоторого множества $C \subset X$ определяется равенством

$$\operatorname{cone} C = \left\{ x \in X \;\middle|\; x = \sum_{i=1}^{s} \lambda_i x_i,\ \lambda_i \geqslant 0,\ x_i \in C \right\}$$

и является выпуклым множеством (п. 2.6.1).

Мы будем говорить, что конус cone C порожден множеством C.

А) Лемма о замкнутости конечнопорожденного конуса в конечномерном пространстве. *Всякий конус в конечномерном пространстве, порожденный конечным множеством точек, замкнут.*

Доказательство. Пусть $K = \operatorname{cone} C \subset \mathbf{R}^N$, $C = \{z_1, \ldots, z_k\}$, $z_i \in \mathbf{R}^N$. Доказательство проведем по индукции.

1) $k = 1$. Тогда $K = \operatorname{cone}\{z_1\} = \{x \mid x = \lambda_1 z_1,\ \lambda_1 \geqslant 0\}$ — полупрямая, т. е. замкнутое множество.

2) Допустим теперь, что лемма верна для $k = s-1$. Докажем ее для $k = s$. Возможны два случая:

а) конус K содержит векторы $-z_1, \ldots, -z_s$. Тогда K есть подпространство конечномерного пространства и, следовательно, замкнутое подмножество;

б) хотя бы один из векторов $(-z_i)$, $i = 1, \ldots, s$, например $(-z_s)$ не принадлежит конусу K. Обозначим через $K_1 = \operatorname{cone}\{z_1, \ldots, z_{s-1}\}$. По индуктивному предположению конус K_1 замкнут. Пусть вектор z принадлежит замыканию конуса K, т. е. $z = \lim\limits_n \xi_n$; $\xi_n \in K$, $n \geqslant 1$. По определению

$$\xi_n \in K = \operatorname{cone}\{z_1, \ldots, z_s\} \Leftrightarrow \xi_n = \sum_{i=1}^{s} \lambda_{in} z_i = \zeta_n + \lambda_{sn} z_s,$$

где $\zeta_n \in K_1$ и $\lambda_{sn} \geqslant 0$.

Если допустить, что $\lambda_{sn} \to \infty$ при $n \to \infty$, то

$$-z_s = \lim_{n\to\infty} \frac{\zeta_n - \xi_n}{\lambda_{sn}} = \lim_{n\to\infty} \frac{\zeta_n}{\lambda_{sn}} \in K_1$$

(напомним, что $\xi_n \to z$, и потому $\xi_n/\lambda_{sn} \to 0$, а $\zeta_n/\lambda_{sn} \in K_1$ вместе с ζ_n), т. е. $-z_s \in K_1$ вследствие замкнутости K_1, что противоречит допущению. Значит, λ_{sn} не стремится к бесконечности и можно выбрать подпоследовательность λ_{sn_k}, сходящуюся к некоторому числу $\lambda_{s0} \geqslant 0$. В этом случае

$$\zeta_{n_k} = \xi_{n_k} - \lambda_{sn_k} z_s \to z - \lambda_{s0} z_s = \tilde{z}.$$

В силу замкнутости K_1 вектор $\tilde{z} \in K_1$ и, следовательно,

$$z = \tilde{z} + \lambda_{s0} z_s \in K. \blacksquare$$

Б) Докажем теперь теорему существования.

Рассмотрим в пространстве $\mathbf{R}^{m+1}$ множество K, образованное такими векторами (α, z), $\alpha \in \mathbf{R}$, $z \in \mathbf{R}^m$, для каждого из которых найдется хотя бы один вектор $\tilde{x} \in \mathbf{R}^n_+$ такой, что $c\tilde{x} \leqslant \alpha$, $A\tilde{x} \geqslant z$.

Множество K — это конус, ибо если $(\alpha, z) \in K$, то для некоторого $\tilde{x} \in \mathbf{R}^n_+$, $c\tilde{x} \leqslant \alpha$, $A\tilde{x} \geqslant z$. Но тогда для любого $t \geqslant 0$, $t\tilde{x} \in \mathbf{R}^n_+$, $c(t\tilde{x}) \leqslant t\alpha$, $A(t\tilde{x}) \geqslant tz$, т. е. $t(\alpha, z) \in K$.

Покажем, что конус K порожден конечным множеством векторов в $\mathbf{R}^{m+1}$:

$$\zeta_1 = (c_1, a_{11}, \ldots, a_{m1}),\ \zeta_2 = (c_2, a_{12}, \ldots, a_{m2}), \ldots, \zeta_n = (c_n, a_{1n}, \ldots, a_{mn}),$$
$$\zeta_{n+1} = (1, 0, \ldots, 0),\ \zeta_{n+2} = (0, -1, 0, \ldots, 0), \ldots, \zeta_{n+m+1} = (0, \ldots, 0, -1).$$

Прежде всего $\zeta_i \in K$, так что cone $\{\zeta_1, \ldots, \zeta_{n+m+1}\} \subset K$. Действительно, если $1 \leqslant i \leqslant n$, надо взять в качестве $\tilde{x}$ стандартный базисный вектор e_i; для остальных i надо взять $\tilde{x} = 0$. Пусть теперь вектор $\zeta = (\alpha, z) = (\alpha, z_1, \ldots, z_m) \in K$. Тогда для некоторого $x \in \mathbf{R}^n$, $x = (x_1, \ldots, x_n)$ такого, что $x_1 \geqslant 0, \ldots, x_n \geqslant 0$ и некоторых $\beta_0 \geqslant 0$, $\beta_j \geqslant 0$, выполняются соотношения

$$\sum_{j=1}^{n} c_j x_j + \beta_0 = \alpha, \quad \sum_{j=1}^{n} a_{ij} x_j - \beta_j = z_i, \quad i = 1, \ldots, m.$$

Но это как раз и означает, что

$$\zeta=(\alpha,\ z)=\sum_{j=1}^{n} x_j\zeta_j+\beta_0\zeta_{n+1}+\sum_{i=1}^{m}\beta_i\zeta_{n+i+1},$$
$$x_j\geqslant 0,\quad \beta_0\geqslant 0,\quad \beta_i\geqslant 0,$$

т. е.

$$\zeta\in\operatorname{cone}\{\zeta_1,\ \ldots,\ \zeta_{n+m+1}\}.$$

По лемме из предыдущего пункта конус K замкнут. По условию теоремы задача (2) имеет допустимый элемент $\tilde{x}$ и значение задачи $\hat{\alpha}>-\infty$. Из определения K следует, что точка $(\tilde{\alpha},\ b)$, где $\tilde{\alpha}=c\tilde{x}$, принадлежит K. При этом, очевидно, что $\tilde{\alpha}\geqslant\hat{\alpha}$. Следовательно, множество $\mathfrak{A}=\{\alpha\in\mathbf{R}\,|\,(\alpha,\ b)\in K\}$ непусто и $\hat{\alpha}=\inf\{\alpha\,|\,\alpha\in\mathfrak{A}\}$, т. е. $(\hat{\alpha}, b)$ принадлежит замыканию конуса K, а значит, и самому K. Следовательно, существует элемент $\hat{x}\geqslant 0$, для которого $A\hat{x}\geqslant b$ и $c\hat{x}\leqslant\hat{\alpha}$, т. е. $\hat{x}$ — решение задачи. ■

Переходим к рассмотрению двойственной задачи. Теорема двойственности приобретает здесь более законченный вид, чем в предыдущем пункте, поскольку ввиду специальной структуры S-функция задачи оказывается замкнутой.

В соответствии с формулами (5) и (6) п. 3.3.2 расширенная функция Лагранжа имеет вид

$$\bar{\mathscr{L}}(x;\lambda)=\begin{cases} cx+\lambda(b-Ax), & x\in\mathbf{R}_+^n,\quad \lambda\in\mathbf{R}_+^{m*},\\ -\infty, & x\in\mathbf{R}_+^n,\quad \lambda\notin\mathbf{R}_+^{m*},\\ +\infty, & x\notin\mathbf{R}_+^n.\end{cases}$$

Отсюда мы находим семейство возмущений задачи (2) и двойственное семейство. Согласно (7) п. 3.3.2

$$f(x;\alpha)=(-\bar{\mathscr{L}})^{*(2)}=\sup_{\lambda}(\lambda\alpha+\mathscr{L}(x,\lambda))=$$
$$=\begin{cases} +\infty, & x\notin\mathbf{R}_+^n,\\ \sup\limits_{\lambda\geqslant 0}(\lambda\alpha+cx+\lambda(b-Ax)), & x\in\mathbf{R}_+^n=\end{cases}$$
$$=\begin{cases} cx, & x\geqslant 0,\quad Ax\geqslant b+\alpha,\\ +\infty & \text{в остальных случаях.}\end{cases}$$

Обозначая $z=b+\alpha$, получаем возмущение задачи (2)

$$cx\to\inf,\quad x\geqslant 0,\quad Ax\geqslant z. \tag{3}$$

Аналогично для определения двойственной задачи мы должны вычислить

$$g(0, \lambda) = (-\tilde{\mathcal{L}}^{*(1)})(0, \lambda) = -\sup_x(-\tilde{\mathcal{L}}(x, \lambda)) =$$

$$= \begin{cases} -\sup\limits_{x \geqslant 0}(-cx - \lambda(b - Ax)), & \lambda \geqslant 0, \\ -\infty, & \lambda \notin \mathbf{R}_+^{m*}, \end{cases} =$$

$$= \begin{cases} \lambda b, & \lambda \geqslant 0, \quad \lambda A \leqslant c, \\ -\infty & \text{в остальных случаях.} \end{cases}$$

Это дает задачу

$$\lambda b \to \sup; \quad \lambda A \leqslant c, \quad \lambda \geqslant 0, \tag{4}$$

двойственную задаче (2). Она имеет такое же строение, как задача (2), и легко понять, что если, отправляясь от задачи (4), построить двойственную задачу по тому же правилу, по которому из задачи (2) получилась задача (4), то мы вернемся к задаче (2). Поэтому имеет смысл говорить о *паре двойственных задач линейного программирования.*

Теорема двойственности. *Для пары двойственных задач линейного программирования справедлива следующая альтернатива: либо значения задач конечны и равны и в обеих задачах существует решение, либо в одной из задач множество допустимых элементов пусто.*

В первом случае $\hat{x} \in \mathbf{R}^n$ и $\hat{\lambda} \in \mathbf{R}^{m*}$, в том и только в том случае, будут решениями задач (2) *и* (4) *соответственно, когда они допустимы в этих задачах и удовлетворяют одному из двух соотношений*

$$c\hat{x} = \hat{\lambda} b, \tag{5}$$

$$\hat{\lambda}(A\hat{x} - b) = (\hat{\lambda} A - c)\hat{x}. \tag{6}$$

Во втором случае одна из задач несовместна, а другая либо несовместна (т. е. ее множество допустимых элементов пусто), либо имеет бесконечное значение.

Доказательство. А) Построение S-функции. Снова рассмотрим тот же конус K, что и в теореме существования. Если $(\alpha, z) \in K$ и $\beta \geqslant \alpha$, то $(\beta, z) \in K$, а потому K является надграфиком функции

$$S(z) = \inf\{\alpha \mid (\alpha, z) \in K\}. \tag{7}$$

Из доказательства теоремы существования следует, что inf достигается и $S(z)$ есть значение задачи (3) и, следовательно, формула (7) определяет S-функцию задачи (2).

Поскольку K — выпуклый и замкнутый конус, S-функция задачи (2) замкнута и выпукла.

Б) Вычислим функцию S^*, сопряженную с функцией S. По определению,

$$S^*(\lambda) = \\ = \sup_z (\lambda z - S(z)) = \sup_z \{\lambda z - \inf_x \{cx \mid x \in \mathbf{R}_+^n,\ Ax \geqslant z\}\} = \\ = \sup_{(z,\,x)} \{\lambda z - cx \mid x \in \mathbf{R}_+^n,\ z \in \mathbf{R}^m,\ z \leqslant Ax\}.$$

Очевидно, что $\sup\{\lambda z \mid z \in \mathbf{R}^m,\ z \leqslant Ax\} = \lambda Ax < \infty$ в том и только в том случае, когда $\lambda \geqslant 0$.

Таким образом,

$$S^*(\lambda) = \begin{cases} \sup\limits_{x \geqslant 0} (\lambda A - c)\,x, & \text{если } \lambda \in \mathbf{R}_+^{m*}, \\ +\infty, & \text{если } \lambda \notin \mathbf{R}_+^{m*} \end{cases} = \\ = \begin{cases} 0, & \text{если } \lambda A \leqslant c,\ \lambda \geqslant 0, \\ +\infty & \text{в остальных случаях.} \end{cases}$$

Поэтому

$$S^{**}(z) = \sup\{\lambda z \mid \lambda A \leqslant c,\ \lambda \geqslant 0\}. \tag{8}$$

В частности, отсюда видно, что $S^{**}(b)$ — значение двойственной задачи (4).

В) З а в е р ш е н и е   д о к а з а т е л ь с т в а. Функция S не может равняться тождественно $+\infty$, ибо $S(0) \leqslant 0$, так как нуль — допустимый элемент в задаче $cx \to \inf$, $Ax \leqslant 0$, $x \geqslant 0$; следовательно, $\operatorname{dom} S \neq \varnothing$.

Возможно одно из двух:

1) $S(z) > -\infty$, $\forall z$ или

2) существует $\tilde{z}$, для которого $S(\tilde{z}) = -\infty$.

Случай 1) в свою очередь распадается на два: 1а) $b \in \operatorname{dom} S$ и 1б) $b \notin \operatorname{dom} S$. В случае 1а) S — собственная функция и значение задачи конечно. Вследствие замкнутости S по теореме Фенхеля—Моро (п. 2.6.3) $S^{**}(b) = S(b)$ и теперь из (8) видно, что двойственная задача (4) имеет то же значение, что и прямая (2) и, в частности, совместна. В силу теоремы существования решение существует в обеих задачах. В случае 1б) $S(b) = +\infty$, т. е. задача (2) не совместна. Снова из теоремы Фенхеля—Моро получаем, что $S^{**}(b) = S(b)$ и, значит, задача (4) совместна, но ее значение бесконечно.

В случае 2) $S(z) = -\infty$ для всех $z \in \operatorname{dom} S$ (см. упражнение 8 п. 2.6.2). По определению $S^*(\lambda) \equiv +\infty$ и

$S^{**}(z)\equiv-\infty$. В частности, $S^{**}(b)=-\infty$, т. е. задача (4) несовместна.

Если $b\in\operatorname{dom}S$, то задача (2) совместна и ее значение бесконечно: $S(b)=-\infty$. Если же $b\notin\operatorname{dom}S$, то $S(b)=+\infty$ задача (2) несовместна. Альтернатива полностью обоснована.

Вернемся теперь к случаю 1а). Там, как было доказано, существуют решения задач (2) и (4). Обозначим их $\hat{x}$ и $\hat{\lambda}$ соответственно. Доказано, что значения задач равны: $c\hat{x}=\hat{\lambda}b$, т. е. выполнено (5) и, значит,

$$\hat{\lambda}(A\hat{x}-b)=\hat{\lambda}A\hat{x}-\hat{\lambda}b=\hat{\lambda}A\hat{x}-c\hat{x}=(\hat{\lambda}A-c)\hat{x},$$

т. е. выполнено (6). Далее, если x и λ — допустимые элементы (т. е. $x\geqslant 0$, $\lambda\geqslant 0$, $\lambda A\leqslant c$, $Ax\geqslant b$), то

$$cx\geqslant\lambda Ax\geqslant\lambda b.$$

Поэтому, если $c\hat{x}=\hat{\lambda}b$, то $\hat{x}$ и $\hat{\lambda}$ — решения задач. Далее, если $\hat{\lambda}(A\hat{x}-b)=(\hat{\lambda}A-c)\hat{x}$, то $c\hat{x}=\hat{\lambda}b$, т. е. $\hat{x}$ и $\hat{\lambda}$ — решения задачи.

Таким образом, если $\hat{x}$ и $\hat{\lambda}$ — решения, то выполнены и (5) и (6); если $\hat{x}$ и $\hat{\lambda}$ допустимы и выполнено либо (5), либо (6), то $\hat{x}$ и $\hat{\lambda}$ — решения задач. Теорема полностью доказана. ■

Упражнение. Во что превращается в рассматриваемой ситуации теорема о минимаксе (следствие 3 п. 3.3.2)?

3.3.4. Теорема двойственности для задачи о кратчайшем расстоянии. Лемма Хоффмана и лемма о минимаксе. Пусть Y — нормированное пространство, $B\subset Y$ — непустое подмножество Y. Величина

$$S_B(\eta)=\rho(\eta, B)=\inf_{y\in B}\|y-\eta\| \tag{1}$$

называется *расстоянием от точки* η до множества B. Исследование функции $S_B(\eta)$ является одной из основных в теории приближений (см. [84]). Если B — выпуклое множество, мы получаем задачу выпуклого программирования. Доказываемая ниже теорема двойственности имеет в анализе многочисленные приложения.

Итак, пусть B — выпуклое множество. (Далее оно фиксировано, и мы опускаем индекс B в обозначении функции S.) Тогда функция $\eta\mapsto S(\eta)$ является

S-функцией такой задачи:

$$\|z\| \to \inf; \quad y - z = \eta, \quad y \in B. \tag{2}$$

Приведем (2) к стандартному виду задачи выпуклого программирования (см. ($\mathfrak{z}$) в п. 3.3.1). Для этого положим $X = Y \times Y$, $x = (y, z)$, $f_0(x) = \|z\|$, $\Lambda x = z - y$, $\Lambda: X \to Y$, $A = \{x = (y, z) \mid y \in B\}$ и тогда задача (2) примет вид

$$f_0(x) \to \inf; \quad \Lambda x + \eta = 0, \quad x \in A. \tag{3}$$

Из сказанного вытекает выпуклость функции S (следствие 1 п. 3.3.2).

В дальнейшем нам понадобится следующее важное геометрическое понятие.

Определение. *Опорной функцией множества $B \subset Y$ называется функция sB: $Y^* \to \overline{\mathbb{R}}$, определяемая равенством*

$$sB(y^*) = \sup_{y \in B} \langle y^*, y \rangle.$$

Теорема двойственности для задачи о кратчайшем расстоянии. *Величина $S(\eta)$ допускает следующее двойственное представление*:

$$S(\eta) = \rho(\eta, B) = \sup \{\langle y^*, y \rangle - sB(y^*) \mid \|y^*\| \leqslant 1\}. \tag{4}$$

Доказательство. Очевидно, что $S(\eta) \geqslant 0$, $\forall \eta$. С другой стороны, $S(\eta) \leqslant \|y_0 - \eta\|$, где y_0 — любая точка из B. Значит, S — ограничена сверху и снизу и, следовательно, непрерывна всюду на Y (предложение 3 п. 2.6.2). Теперь можно применить теорему двойственности из п. 3.3.2. Поскольку в (2) неравенства отсутствуют, функция Лагранжа имеет вид

$$\mathscr{L}(x, y^*, 1) = \|z\| + \langle y^*, z - y \rangle.$$

Из формулы (11) п. 3.3.2 получаем

$$S(\eta) = \sup_{y^*} \left(\langle y^*, \eta \rangle + \inf_{\substack{y \in B \\ z \in Y}} (\|z\| + \langle y^*, z - y \rangle) \right) =$$

$$= \sup_{y^*} \left\{ \left(\langle y^*, \eta \rangle - \sup_{y \in B} \langle y^*, y \rangle \right) - \sup_{z \in Y} (\langle -y^*, z \rangle - \|z\|) \right\} =$$

$$= \sup_{y^*} \{\langle y^*, \eta \rangle - sB(y^*) - N^*(-y^*)\},$$

где $N(z) = \|z\|$, а $N^*(z) = 0$ при $\|z^*\| \leqslant 1$ и $+\infty$ при $\|z^*\| > 1$. (Предложение 3 п. 2.6.3.) Отсюда следует (4). ■

В дальнейшем понадобится следующее обобщение следствия 2 п. 2.6.4.

Лемма о сопряженном конусе. *Пусть X и Y — банаховы пространства, $\Lambda\colon X\to Y$ — линейный сюръективный оператор из X в Y, $x_1^*, \ldots, x_s^*$ — элементы сопряженного пространства X^*, и пусть*

$$K=\{x\mid \langle x_i^*, x\rangle\leqslant 0,\ i=1, \ldots, s,\ \Lambda x=0\}$$

— конус в X.

Тогда всякий элемент x_0^* из сопряженного конуса

$$K^*=\{x^*\mid \langle x^*, x\rangle\geqslant 0,\ x\in K\}$$

допускает представление в виде

$$-x_0^*=\sum_{i=1}^{s}\lambda_i x_i^*+\Lambda^* y^*$$

для некоторых $\lambda_i\geqslant 0$ и $y^*\in Y^*$.

Доказательство. А) Обозначим $L=\bigcap\limits_{i=0}^{s}\operatorname{Ker}x_i^*$, $Z=\operatorname{Ker}\Lambda/L$, и пусть $\pi\colon \operatorname{Ker}\Lambda\to Z$ — естественная проекция, отображающая $x\in\operatorname{Ker}\Lambda$ в класс $\pi(x)\in Z$, его содержащий. Покажем, что $\dim Z\leqslant s+1$. Действительно, пусть $z_0, \ldots, z_{s+1}$ — произвольные элементы из Z, $z_i=\pi(x_i)$. Однородная линейная система $s+1$ уравнений

$$\sum_{j=0}^{s+1}\langle x_i^*, x_j\rangle\lambda_j=\left\langle x_i^*, \sum_{j=0}^{s+1}\lambda_j x_j\right\rangle=0,\quad i=0, 1, \ldots, s,$$

с $(s+2)$ неизвестным имеет ненулевое решение $\hat\lambda_0, \ldots, \hat\lambda_{s+1}$. Тогда

$$\hat x=\sum_{j=0}^{s+1}\hat\lambda_j x_j\in\operatorname{Ker}x_i^*,\quad i=0, \ldots, s, \Rightarrow \hat x\in L\Rightarrow 0=\pi(\hat x)=$$
$$=\sum_{j=0}^{s+1}\hat\lambda_j\pi(x_j)=\sum_{j=0}^{s+1}\hat\lambda_j z_j.$$

Следовательно, любые $s+2$ элементов из Z линейно зависимы и $d=\dim Z\leqslant s+1$.

Б) Выбрав в Z некоторый базис $f_1, \ldots, f_d$, установим стандартный изоморфизм между Z и $\mathbf{R}^d$ и между Z^* и $\mathbf{R}^{d*}$

$$z=\sum_{i=1}^{d}\zeta_i f_i\mapsto(\zeta_1, \ldots, \zeta_d),$$

$$\langle z^*, z\rangle=\sum_{i=1}^{d}\langle z^*, f_i\rangle\zeta_i\Rightarrow$$
$$\Rightarrow z^*\mapsto(\zeta_1^*, \ldots, \zeta_d^*)=(\langle z^*, f_1\rangle, \ldots, \langle z^*, f_d\rangle).$$

Далее, определим функционалы $z_i^* \in Z^*$, $i = 0, 1, \ldots, s$, равенствами

$$\langle z_i^*, \pi(x)\rangle = \langle x_i^*, x\rangle$$

и рассмотрим в Z^* выпуклый конус

$$\tilde{K} = \operatorname{cone}\{z_1^*, \ldots, z_s^*\} = \left\{ z^* \,\middle|\, \sum_{i=1}^{s} \lambda_i z_i^*,\ \lambda_i \geqslant 0 \right\}.$$

Поскольку Z^* конечномерно, а конус $\tilde{K}$ конечно порожден, то по лемме из п. 3.3.3 он и замкнут.

Предположим, что $-z_0^* \notin \tilde{K}$. Тогда по второй теореме отделимости (п. 2.1.4) существует линейный функционал $l \in (Z^*)^*$, строго разделяющий $\{-z_0^*\}$ и $\tilde{K}$:

$$\langle l, -z_0^*\rangle > \sup\{\langle l, z^*\rangle \mid z^* \in \tilde{K}\} = \sup\left\{ \sum_{i=1}^{s} \lambda_i \langle l, z_i^*\rangle \,\middle|\, \lambda_i \geqslant 0 \right\}.$$

Но тогда обязательно $\langle l, z_i^*\rangle \leqslant 0$ (иначе $\sup = +\infty$), $0 = \sup\{\langle l, z^*\rangle \mid z^* \in \tilde{K}\}$ и $\langle l, z_0^*\rangle < 0$.

Теперь заметим, что l, как и всякий линейный функционал на конечномерном пространстве Z^*, задается линейной формой

$$\langle l, z^*\rangle = \sum_{i=1}^{d} a_i \zeta_i^* = \sum_{i=1}^{d} a_i \langle z^*, f_i\rangle = \left\langle z^*, \sum_{i=1}^{d} a_i f_i \right\rangle = \langle z^*, a\rangle.$$

Выбрав в классе $a = \sum_{i=1}^{d} a_i f_i \in Z = \operatorname{Ker}\Lambda/L$ представителя $x_0 \in \operatorname{Ker}\Lambda$, $a = \pi(x_0)$ имеем, с одной стороны,

$$\langle x_i^*, x_0\rangle = \langle z_i^*, \pi(x_0)\rangle = \langle z_i^*, a\rangle = \langle l, z_i^*\rangle \leqslant 0$$

и, следовательно,

$$x_0 \in \operatorname{Ker}\Lambda \cap \bigcap_{i=1}^{s} \{x \mid \langle x_i^*, x\rangle \leqslant 0\} = K.$$

С другой стороны,

$$\langle x_0^*, x_0\rangle = \langle z_0^*, \pi(x_0)\rangle = \langle z_0^*, a\rangle = \langle l, z_0\rangle < 0$$

и, следовательно, $x_0^* \notin K^*$ вопреки условию леммы.

Таким образом, предположение $-z_0^* \notin \tilde{K}$ приводит к противоречию, и потому $-z_0^* = \sum_{i=1}^{s} \lambda_i z_i^*$ для некоторых $\lambda_i \geqslant 0$.

В) Для любого $x \in \operatorname{Ker} \Lambda$

$$\left\langle x_0^* + \sum_{i=1}^{s} \lambda_i x_i^*, x \right\rangle = \left\langle z_0^* + \sum_{i=1}^{s} \lambda_i z_i^*, \pi(x) \right\rangle = 0.$$

Поэтому $x_0^* + \sum_{i=1}^{s} \lambda_i x_i^* \in (\operatorname{Ker} \Lambda)^{\perp}$, и по лемме об ядре регулярного оператора (п. 2.1.7) $x_0^* + \sum_{i=1}^{s} \lambda_i x_i^* = \Lambda^*(-y^*)$ для некоторого $y^* \in Y^*$. ■

Лемма Хоффмана. *Пусть выполнены те же условия, что и в лемме о сопряженном конусе.*

Тогда для функции расстояния от точки x до K справедливо неравенство

$$\rho(x, K) \leqslant C \left\{ \sum_{i=1}^{s} \langle x_i^*, x \rangle_+ + \|\Lambda x\| \right\}, \tag{5}$$

где $\langle x_i^*, x \rangle_+$ равно $\langle x_i^*, x \rangle$, если $\langle x_i^*, x \rangle \geqslant 0$, и нулю в остальных случаях, а константа C не зависит от x.

Доказательство. Множество K, являющееся пересечением конечного числа полупространств и подпространства,— выпуклый конус в X. Вычислим его опорную функцию sK. Пусть $x^* \in X^*$. Возможно одно из двух: либо найдется такой элемент $x_0 \in K$, что $\langle x^*, x_0 \rangle > 0$, либо $\langle x^*, x \rangle \leqslant 0$, $\forall x \in K$. В первом случае $sK(x^*) \geqslant t \langle x^*, x_0 \rangle$ $\forall t \in \mathbf{R}_+$ и, значит, $sK(x^*) = +\infty$.

Во втором случае $(-1)x^*$ принадлежит сопряженному конусу конуса K и, значит, по предыдущей лемме $x^* = \sum_{i=1}^{s} \lambda_i x_i^* + \Lambda^* y^*$, $\lambda_i \geqslant 0$, $y^* \in Y^*$. Итак,

$$sK(x^*) = \begin{cases} 0, & \text{если } \exists \lambda_i \in 0,\ y^* \in Y^*\colon\ x^* = \sum_{i=1}^{s} \lambda_i x_i^* + \Lambda^* y, \\ +\infty & \text{в остальных случаях.} \end{cases}$$

Применив теперь формулу (4) (теорему двойственности) к нашей задаче, получаем

$$\rho(x, K) =$$
$$= \sup \left\{ \langle x^*, x \rangle \mid x^* = \sum_{i=1}^{s} \lambda_i x_i^* + \Lambda^* y^*,\ \lambda_i \geqslant 0,\ \|x^*\| \leqslant 1 \right\}. \tag{6}$$

Подпространство $L = \operatorname{lin}\{x_1^*, \ldots, x_s^*\} + \operatorname{Im} \Lambda^*$ является суммой замкнутого подпространства $L_1 = \operatorname{Im} \Lambda^*$ (ибо

$\operatorname{Im}\Lambda^* = (\operatorname{Ker}\Lambda)^\perp$, а аннулятор всегда замкнут) и конечномерного подпространства $L_2 = \{x_1^*, \ldots, x_s^*\}$. Поэтому L замкнуто в X (докажите!) и, следовательно, банахово. Оператор Λ_1: $\mathbb{R}^s \times Y^* \to L$, $\Lambda_1(\lambda, y) = \sum_{i=1}^{s} \lambda_i x_i^* + \Lambda^* y^*$ линеен, непрерывен и отображает $\mathbb{R}^s \times Y^*$ на банахово пространство L. По лемме о правом обратном отображении (п. 2.1.5) существует отображение M_1: $L \to \mathbb{R}^s \times Y^*$ такое, что $\Lambda_1 \circ M_1 = I_L$, $\|M_1 x^*\| \leqslant C \|x^*\|$. Тогда, если $\|x^*\| \leqslant 1$, то $\|M_1 x^*\|_{\mathbb{R}^s \times Y^*} = \sum_{i=1}^{s} |\lambda_i| + \|y^*\| \leqslant C$. Таким образом, в выражении (6) можно считать, что $0 \leqslant \lambda_i \leqslant C$, $\|y^*\| \leqslant C$, откуда и получаем

$$S(x) = \rho(x, K) \leqslant$$
$$\leqslant \sup \left\{ \left\langle \sum_{i=1}^{s} \lambda_i x_i^* + \Lambda^* y^*, x \right\rangle \Big| 0 \leqslant \lambda_i \leqslant C, \|y^*\| \leqslant C \right\} \leqslant$$
$$\leqslant C \left\{ \sum_{i=1}^{s} \langle x_i^*, x \rangle_+ + \|\Lambda x\| \right\}. \blacksquare$$

Заметим, что если конус K задается лишь равенствами

$$K = \{x \mid \langle x_i^*, x \rangle = 0, \ \Lambda x = 0\}$$

(т. е. является подпространством), то он допускает определение и через неравенства

$$K = \{x \mid \langle x_i^*, x \rangle \leqslant 0, \ \langle (-1) x_i^*, x \rangle \leqslant 0, \ \Lambda x = 0\}.$$

Применив лемму Хоффмана, получаем и в этом случае, что

$$\rho(x, K) \leqslant C \left\{ \sum_{i=1}^{s} |\langle x_i^*, x \rangle| + \|\Lambda x\| \right\}. \qquad (5')$$

Лемма о минимаксе. *Пусть X и Y — банаховы пространства, Λ: $X \to Y$ — сюръективный оператор, $x_i^* \in X^*$, $i = 1, \ldots, s$, $a = (a_1, \ldots, a_s)$. Определим функцию S: $\mathbb{R}^s \times Y \to \overline{\mathbb{R}}$ равенством*

$$S(a, y) = \inf_{\Lambda x + y = 0} \max_{1 \leqslant i \leqslant s} (a_i + \langle x_i^*, x \rangle). \qquad (7)$$

Если $\max\limits_{1\leqslant i\leqslant s}\langle x_i^*, x\rangle\geqslant 0$ *для любого* $x\in\operatorname{Ker}\Lambda$, *то имеет место следующая формула двойственности:*

$$S(a, y)=\sup\left\{\sum_{i=1}^{s}\alpha_i a_i+\langle y^*, y\rangle \mid \alpha_i\geqslant 0,\right.$$
$$\left.\sum_{i=1}^{s}\alpha_i=1,\ \Lambda^* y^*+\sum_{i=1}^{s}\alpha_i x_i^*=0\right\}. \quad (8)$$

При этом inf *в* (7) *достигается на некотором* (*быть может, не единственном*) $\hat{x}=\hat{x}(a, y)$, *и существуют такие* $C>0, \tilde{C}>0$, *не зависящие от* a, y, *что при подходящем выборе* $\hat{x}(a, y)$

$$\|\hat{x}(a, y)\|\leqslant C\{|S(a, y)|+|a|+\|y\|\}\leqslant\tilde{C}(|a|+\|y\|). \quad (9)$$

Доказательство. Существование минимума в (7), выпуклость S и формулу (8) мы докажем редукцией к двум стандартным задачам, рассмотренным выше.

А) Вспомогательные отображения. Нам придется дважды воспользоваться леммой о правом обратном отображении (п. 2.1.5). Во-первых, существует $\mathrm{M}\colon Y\to X$ такое, что $\Lambda\circ\mathrm{M}=I$, $\|\mathrm{M}(y)\|\leqslant C_1\|y\|$. Во-вторых, пусть $\varphi\colon X\to\mathbf{R}^s$ определяется равенством $\varphi(x)=(\langle x_1^*, x\rangle, \ldots, \langle x_s^*, x\rangle)$. Обозначим $L=\varphi(\operatorname{Ker}\Lambda)$. По той же лемме существует отображение $\mu\colon L\to\operatorname{Ker}\Lambda$ такое, что

$$\varphi\circ\mu=I,\ \|\mu(\xi)\|\leqslant C_2|\xi|.$$

Далее, как и всякое подпространство в $\mathbf{R}^s$, L может быть задано некоторой системой линейных уравнений $\sum\limits_{j=1}^{s}b_{lj}\xi_j=0$, $l=1, \ldots, p$, или при помощи отображения $\beta\colon\mathbf{R}^s\to\mathbf{R}^p$, $\beta\xi=\left(\sum\limits_{j=1}^{s}b_{lj}\xi_j\right)$ условием $\beta\xi=0$.

Б) Ограниченность $S(a, y)$. Поскольку

$$\Lambda x+y=0\Leftrightarrow x=\mathrm{M}(-y)+x_0,\quad x_0\in\operatorname{Ker}\Lambda\Rightarrow a_i+\langle x_i^*, x\rangle=$$
$$=a_i+\langle x_i^*, \mathrm{M}(-y)\rangle+\langle x_i^*, x_0\rangle, \quad (10)$$

то, полагая, $x_0=0$, мы получаем, с одной стороны, неравенство

$$\inf_{\Lambda x+y=0}\max_{1\leqslant i\leqslant s}(a_i+\langle x_i^*, x\rangle)\leqslant$$
$$\leqslant\max_{1\leqslant i\leqslant s}\{|a_i|+C_1\|x_i^*\|\|y\|\}\overset{\text{def}}{=}K, \quad (11)$$

а, с другой стороны, при $\Lambda x+y=0$, согласно (10),

$$\max_{1\leqslant i\leqslant s}(a_i+\langle x_i^*, x\rangle)\geqslant\min_{1\leqslant i\leqslant s}(\tilde{a}_i-C_1\|x_i^*\|\|y\|)\geqslant -K, \quad (12)$$

так как по условию $\max\limits_{1\leqslant i\leqslant s}(\langle x_i^*, x_0\rangle)\geqslant 0$ $(x_0\in\operatorname{Ker}\Lambda)$. Из (11) и (12) следует оценка

$$|S(a, y)|\leqslant K=\max_{1\leqslant i\leqslant s}\{|a_i|+C_1\|x_i^*\|\|y\|\}. \quad (13)$$

В) Существование минимума. Обозначив $\tilde{a}_i=a_i+\langle x_i^*, \mathrm{M}(-y)\rangle+K$ и вспоминая, что

$$x_0\in\operatorname{Ker}\Lambda\Leftrightarrow(\langle x_1^*, x_0\rangle, \ldots, \langle x_s^*, x_0\rangle)=$$
$$=\varphi(x_0)\in L=\varphi(\operatorname{Ker}\Lambda)=\operatorname{Ker}\beta,$$

мы видим из (10) и (7), что $S(a, y)+K$ является значением задачи

$$\max(\tilde{a}_i+\xi_i)\to\inf;\ \xi\in L\Leftrightarrow c\to\inf;\ \tilde{a}_i+\xi_i\leqslant c, \beta\xi=0. \quad (14)$$

В силу (13) $S(a, y)+K\geqslant 0$, и потому к (14) можно безболезненно присоединить условие $c\geqslant 0$.

Задачу (14) можно привести к стандартному виду задачи линейного программирования (2) п. 3.3.3) в $\mathbb{R}^{s+1}$. Для этого обозначим $z_0=c$, $z_i=c-\xi_i-\tilde{a}_i$, $z=(z_1, \ldots, z_s)$, $\theta=(1, \ldots, 1)$, $\tilde{a}=(\tilde{a}_1, \ldots, \tilde{a}_s)$, $\tilde{z}=(z_0, z)$. Условие $\beta\xi=0$ записывается в виде

$$\beta\xi=0\Leftrightarrow\beta(c\theta-z-\tilde{a})=c\beta\theta-\beta z-\beta\tilde{a}=0\Leftrightarrow\tilde{B}\tilde{z}\geqslant\tilde{b},$$

где

$$\tilde{B}=\begin{pmatrix}\beta\theta & -\beta\\ -\beta\theta & \beta\end{pmatrix}, \quad \tilde{b}=\begin{pmatrix}\beta\tilde{a}\\ -\beta\tilde{a}\end{pmatrix}$$

— матрица порядка $2p\times(s+1)$ и $2p$-мерный вектор (фактически мы заменили равенство $\beta\xi=0$ двумя неравенствами $\beta\xi\geqslant 0$ и $\beta\xi\leqslant 0$).

Поэтому задача (14) эквивалентна задаче

$$z_0\to\inf,\quad \tilde{B}\tilde{z}\geqslant\tilde{b},\quad \tilde{z}\geqslant 0, \quad (15)$$

имеющей стандартный вид. Эта задача совместна (например, вектор $(2K, 2K-\tilde{a}_1, \ldots, 2K-\tilde{a}_s)$ допустимый; соответствующее $\xi=0\in L$ и ее значение $S(a, y)+K$ конечно (и даже неотрицательно). По теореме существования из п. 3.3.3 она имеет решение $\hat{z}$. Следовательно, задача (14) имеет решение $\hat{\xi}=(S(a, y)+K)\theta-\hat{z}-\tilde{a}$, и тогда

$$S(a, y)+K=\max_{1\leqslant i\leqslant s}\{\tilde{a}_i+\hat{\xi}_i\}=$$
$$=\max_{1\leqslant i\leqslant s}\{a_i+\langle x_i^*, \mathrm{M}(-y)\rangle+\hat{\xi}_i\}+K$$

и

$$S(a, y)=\max_{1\leqslant i\leqslant s}\{a_i+\langle x_i^*, \hat{x}\rangle\},$$

где, согласно (10) и определению отображения $\mu: L\to\operatorname{Ker}\Lambda$,

$$\hat{x}=\mathrm{M}(-y)+\mu\hat{\xi}. \tag{16}$$

Г) Выпуклость S и теорема двойственности. Функция $(a, y)\mapsto S(a, y)$ является S-функцией следующей задачи выпуклого программирования:

$$\max(\eta_1, \ldots, \eta_s)\to\inf, \quad -\eta_i+\langle x_i^*, x\rangle+a_i=0,$$
$$\Lambda x+y=0. \tag{17}$$

Полагая $Z=\mathbf{R}^s\times X$, $\tilde{Y}=\mathbf{R}^s\times Y$, $z=(\eta, x)$,

$$f_0(z)=\max(\eta_1, \ldots, \eta_s),$$
$$\tilde{\Lambda}z=(\langle x_1^*, x\rangle-\eta_1, \ldots, \langle x_s^*, x\rangle-\eta_s, \Lambda x),$$

приведем (17) к стандартному виду (($\mathfrak{z}$) в п. 3.3.2):

$$f_0(z)\to\inf, \quad \tilde{\Lambda}z+(a, y)=0. \tag{18}$$

Следовательно, функция $(a, y)\mapsto S(a, y)$ выпукла (следствие 1 п. 3.3.2). Согласно (13) она ограничена, а потому и непрерывна на всем пространстве $\tilde{Y}$ (предложение 3 п. 2.6.2). По теореме двойственности из п. 3.3.2 (напомним, что неравенства в задаче (18) отсутствуют)

$$S(a, y)=$$
$$=\sup_{z^*}\{\langle z^*, (a, y)\rangle+\inf_z[\langle z^*, \tilde{\Lambda}z\rangle+\max(\eta_1, \ldots, \eta_s)]\}. \tag{19}$$

Но $z^*\in(\mathbf{R}^s\times Y)^*=\mathbf{R}^{s*}\times Y^*$, т. е. z^* представим в виде (α, y^*), $\alpha=(\alpha_1, \ldots, \alpha_s)\in\mathbf{R}^{s*}$, $y^*\in Y^*$, так что

$$\begin{aligned}&\langle z^*, (a, y)\rangle=\alpha a+\langle y^*, y\rangle,\\ &\langle z^*, \tilde{\Lambda}z\rangle=\sum_{i=1}^s\alpha_i(\langle x_i^*, x\rangle-\eta_i)+\langle y^*, \Lambda x\rangle.\end{aligned} \tag{20}$$

Согласно предложению 2 п. 2.6.3 преобразование Юнга — Фенхеля функции $f(\eta)=\max(\eta_1, \ldots, \eta_s)$ имеет вид

$$f^*(\alpha)=\sup_\eta(\alpha\eta-f(\eta))=\begin{cases} 0, & \text{если } \alpha_i\geqslant 0,\ \sum_{i=1}^s \alpha_i=1, \\ +\infty & \text{в остальных случаях.}\end{cases} \tag{21}$$

Теперь из (20) и (21) имеем

$$\inf_z(\max(\eta_1, \ldots, \eta_s)+\langle z^*, \tilde{\Lambda}z\rangle)=$$

$$=-\sup_{(\eta, x)}\left\{\alpha\eta-\max(\eta_1, \ldots, \eta_s)-\left\langle\sum_{i=1}^s \alpha_i x_i^*+\Lambda^*y^*,\ x\right\rangle\right\}=$$

$$=\begin{cases} 0, \text{ если } \alpha_i\geqslant 0,\ \sum_{i=1}^s \alpha_i=1,\ \Lambda^*y^*+\sum_{i=1}^s \alpha_i x_i^*=0, \\ -\infty \text{ в остальных случаях.}\end{cases} \tag{22}$$

Подставляя это в (19), получаем (8).

Д) Геометрическая лемма. *Пусть L — подпространство, а K — конечно порожденный конус в $\mathbf{R}^s$. Тогда существует такое $N>0$, что для любого $a\in\mathbf{R}^s$*

$$\rho(0,\ (L+a)\cap K)\leqslant N\rho(0,\ L+a)\leqslant N|a| \tag{23}$$

(формула остается справедливой и в случае $(L+a)\cap K=\varnothing$, если мы условимся считать, что $\rho(0, \varnothing)=-\infty$).

Доказательство. а) Пусть $K=\operatorname{cone}\{x_1, \ldots, x_m\}$. Разложим каждый вектор на две компоненты — содержащуюся в подпространстве L и ортогональную к нему:

$$a=b+c,\ x_i=y_i+z_i,\quad i=1, \ldots, m,\ b,\ y_i\in L,\ c,\ z_i\in L^\perp.$$

Тогда

$$\rho(0,\ (L+a)\cap K)=\inf\{|\xi|\mid \xi\in K,\ (\xi-a)\in L\}=$$

$$=\inf\left\{\left|\sum_{i=1}^m \lambda_i(y_i+z_i)\right|\,\middle|\,\lambda_i\geqslant 0,\ \sum_{i=1}^m \lambda_i(y_i+z_i)-b-c\in L\right\}=$$

$$=\inf\left\{\left|\sum_{i=1}^m \lambda_i(y_i+z_i)\right|\,\middle|\,\lambda_i\geqslant 0,\ \sum_{i=1}^m \lambda_i z_i=c\right\}\leqslant$$

$$\leqslant\inf\left\{\sum_{i=1}^m \lambda_i \max_{1\leqslant i\leqslant m}\{|y_i|\}+|c|\,\middle|\,\lambda_i\geqslant 0,\ \sum_{i=1}^m \lambda_i z_i=c\right\}. \tag{24}$$

С другой стороны,

$$\rho(0,\ L+a)=\inf\{|\eta+a|\mid\eta\in L\}=$$

$$=\inf\{|\eta'+c|\mid\eta'\in L\}=|c|\leqslant\sqrt{|b|^2+|c|^2}=|a|.$$

Поскольку

$$\lambda_i \geqslant 0, \quad \sum_{i=1}^{m} \lambda_i z_i = c \Leftrightarrow c \in \operatorname{cone}\{z_1, \ldots, z_m\},$$

мы видим, что (23) будет следовать из (24), если для некоторого $N_1 > 0$

$$c \in \operatorname{cone}\{z_1, \ldots, z_m\} \Rightarrow \exists \lambda_i \geqslant 0,$$

$$\sum_{i=1}^{m} \lambda_i z_i = c, \quad \sum_{i=1}^{m} \lambda_i \leqslant N_1 |c| \tag{25}$$

(при этом в (23) мы будем иметь $N = N_1 \max\limits_{1 \leqslant i \leqslant m} \{|y_i|\} + 1$).

б) Пусть $L_1 = \operatorname{lin}\{z_1, \ldots, z_m\}$, $\dim L_1 = n$. Выделим всевозможные наборы индексов $I = \{i_1, \ldots, i_n\}$, для которых $\{z_{i_1}, \ldots, z_{i_n}\}$ — базис в L_1. Каждому такому набору отвечает линейное отображение $\Lambda_I: \mathbf{R}^n \to L_1$, определяемое формулой $\Lambda_I x = \sum\limits_{k=1}^{n} x_k z_{i_k}$. Поскольку $\{z_{i_1}, \ldots, z_{i_n}\}$ — базис в L_1, существует обратное отображение.

По теореме Каратеодори (п. 2.6.1) каждый вектор $c \in \operatorname{cone}\{z_1, \ldots, z_m\}$ является конической комбинацией не более чем n линейно независимых векторов z_i

$$c = \lambda_{i_1} z_{i_1} + \ldots + \lambda_{i_s} z_{i_s}, \quad \lambda_{i_k} > 0, \quad s \leqslant n.$$

Дополнив, если нужно, набор $\{z_{i_1}, \ldots, z_{i_s}\}$ до базиса в L_1 некоторыми векторами $z_{i_{s+1}}, \ldots, z_{i_n}$ и положив $\lambda_{i_{s+1}} = \ldots = \lambda_{i_n} = 0$, мы видим, что

$$c = \Lambda_I \lambda, \quad \lambda = (\lambda_{i_1}, \ldots, \lambda_{i_n}), \quad \lambda_{i_k} \geqslant 0, \quad I = (i_1, \ldots, i_n).$$

Поэтому

$$\sum_{k=1}^{n} \lambda_{i_k} \leqslant n |\lambda| \leqslant n \|\Lambda_I^{-1}\| |c| \leqslant n \max_I \{\|\Lambda_I^{-1}\|\} |c|.$$

Этим доказано (25) для $N_1 = n \max\limits_I \{\|\Lambda_I^{-1}\|\}$.

Е) Завершение доказательства леммы о минимаксе. Последнее, что осталось нам сделать, — получить оценку (9). Для этого вернемся к левой из задач (14). Элемент $\bar{\xi} = (\bar{\xi}_1, \ldots, \bar{\xi}_s)$ является ее решением тогда и только тогда, когда

$$\bar{\xi}_i + \tilde{a}_i \leqslant S(a, y) + K, \quad i = 1, \ldots, s, \quad \bar{\xi} \in L \tag{26}$$

(поскольку $S(a, y)+K$ — значение задачи, по крайней мере для одного индекса здесь будет иметь место равенство). Вспоминая определение $\tilde{a}_i$ в п. В) и обозначая $\hat{a}=(\hat{a}_1, \ldots, \hat{a}_s)$, где

$$\hat{a}_i=\tilde{a}_i-S(a, y)-K=a_i+\langle x_i^*, \mathrm{M}(-y)\rangle-S(a, y), \tag{27}$$

переписываем (26) в виде $\tilde{\xi}_i+\hat{a}_i\leqslant 0$.

Таким образом $\tilde{\xi}$ — решение (14)$\Leftrightarrow\tilde{\xi}+\hat{a}\in(L+\hat{a})\cap(-\mathrm{R}_+^s)$. Поэтому

$$\inf\{|\tilde{\xi}+\hat{a}|\,|\,\tilde{\xi} \text{ решение } (14)\}=\rho(0, (L+\hat{a})\cap(-\mathrm{R}_+^s)).$$

Поскольку $-\mathrm{R}_+^s=\mathrm{cone}\{-e_1, \ldots, -e_s\}$ — конечно порожденный конус, правая часть по геометрической лемме п. Д) не превосходит $N\rho(0, L+\hat{a})\leqslant N|a|$. Следовательно,

$$\inf\{|\tilde{\xi}+\hat{a}|\,|\,\tilde{\xi} \text{ решение } (14)\}\leqslant N|\hat{a}|,$$

и так как $|\tilde{\xi}|\leqslant|\tilde{\xi}+\hat{a}|+|\hat{a}|$,

$$\inf\{|\tilde{\xi}|\,|\,\tilde{\xi} \text{ решение } (14)\}\leqslant(N+1)|\hat{a}|.$$

Если $\hat{a}\neq 0$, то $(N+1)|\hat{a}|<(N+2)|\hat{a}|$, и среди решений задачи (14) мы можем выбрать такое, обозначим его $\hat{\xi}$, у которого

$$|\hat{\xi}|\leqslant(N+2)|\hat{a}|. \tag{28}$$

Если $\hat{a}=0$, то $\hat{\xi}=0$ решение и (28) снова верно.

Воспользовавшись формулой (16), находим по $\hat{\xi}$ элемент $\hat{x}=\hat{x}(a, y)$. При его оценке учитываем неравенства

$$\|\mathrm{M}(-y)\|\leqslant C_1\|y\|, \quad \|\mu(\hat{\xi})\|\leqslant C_2|\hat{\xi}| \tag{29}$$

(ср. определение правых обратных отображений M и μ в пункте А)) и

$$|\hat{a}|=\left(\sum_{i=1}^{s}\hat{a}_i^2\right)^{1/2}\leqslant\sqrt{s}\max\{|\hat{a}_1|, \ldots, |\hat{a}_s|\}\overset{(27)}{\leqslant}$$
$$\leqslant\sqrt{s}\left[\max_{1\leqslant i\leqslant s}\{|a_i|+\|x_i^*\|C_1\|y\|\}+|S(a, y)|\right]\leqslant$$
$$\leqslant\sqrt{s}\left[|a|+C_1\max_{1\leqslant i\leqslant s}\|x_i^*\|\|y\|+|S(a, y)|\right], \tag{30}$$

а также оценку (13):

$$\|\hat{x}(a, y)\| \overset{(16),(29)}{\leqslant} C_1\|y\|+C_2|\hat{\xi}| \overset{(28)}{\leqslant} C_1\|y\|+C_2(N+2)|\hat{a}| \overset{(30)}{\leqslant}$$

$$\leqslant C_1(1+C_2(N+2)\sqrt{s}\max_{1\leqslant i\leqslant s}\|x_i^*\|)\|y\|+$$

$$+C_2(N+2)\sqrt{s}(|a|+|S(a, y)|) \overset{(13)}{\leqslant}$$

$$\leqslant C_1(1+\sqrt{s}\max_{1\leqslant i\leqslant s}\|x_i^*\|2C_2(N+2))\|y\|+$$

$$+2C_2(N+2)\sqrt{s}|a|.$$

Поэтому верно (9) с константами

$$C=\max\{C_1(1+C_2(N+2)\sqrt{s}\max_{1\leqslant i\leqslant s}\|x_i^*\|, C_2(N+2)\sqrt{s}\},$$

$$\tilde{C}=\max\{C_1(1+\sqrt{s}\max_{1\leqslant i\leqslant s}\|x_i^*\|2C_2(N+2)),$$

$$2C_2(N+2)\sqrt{s}\}.$$

§ 3.4*. Необходимые условия второго порядка и достаточные условия экстремума в гладких задачах

Снова разберем сначала случай, когда неравенства отсутствуют.

3.4.1. Гладкие задачи с равенствами. Рассмотрим задачу

$$f(x)\to \mathrm{extr};\quad F(x)=0. \tag{1}$$

Теорема 1 (необходимые условия второго порядка). *Пусть X и Y — банаховы пространства, U — открытое множество в X, функция $f\colon U\to\mathbf{R}$ и отображение $F\colon U\to Y$ строго дифференцируемы в точке $\hat{x}\in U$ и имеют в этой точке $\hat{x}$ вторую производную Фреше.*

Если $\hat{x}$ доставляет локальный минимум (максимум) в задаче (1) *и если F регулярно в точке $\hat{x}$* (т. е. $\mathrm{Im}\,F'(\hat{x})=Y$), *то существует множитель Лагранжа $\hat{y}^*\in Y^*$ такой, что*

$$\mathscr{L}_x(\hat{x}, \hat{y}^*, 1)=0, \tag{2}$$

$$\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, 1)[h, h]\geqslant 0\ (\leqslant 0),\quad \forall h\in\mathrm{Ker}\,F'(\hat{x}), \tag{3}$$

где $\mathscr{L}(x, \hat{y}^*, 1)=f(x)+\langle\hat{y}^*, F(x)\rangle$.

Доказательство. Равенство (2) было доказано в п. 3.2.2. Далее разбираем случай $\hat{x}\in\mathrm{locmin}$ (1). Пусть $h\in\mathrm{Ker}\,F'(\hat{x})$. Согласно теореме Люстерника (п. 2.3.5)

$h \in T_{\hat{x}}\mathscr{M}$, где $\mathscr{M} = \{x \mid F(x) = 0\}$, т. е. существует отображение $r(\cdot)\colon [-\varepsilon, \varepsilon] \to \mathscr{M}$ такое, что

$$F(\hat{x} + th + r(t)) = 0, \quad r(t) = o(t), \quad t \in [-\varepsilon, \varepsilon]. \qquad (4)$$

В силу (4) $\hat{x} + th + r(t)$ — допустимый элемент в задаче при $t \in [-\varepsilon, \varepsilon]$ и, следовательно, $f(\hat{x}) \leqslant f(\hat{x} + th + r(t))$, ибо $\hat{x} \in \operatorname{locmin}(1)$. Поэтому

$$\begin{aligned} f(\hat{x}) \leqslant f(\hat{x} + th + r(t)) = \mathscr{L}(\hat{x} + th + r(t), \hat{y}^*, 1) = \\ = \mathscr{L}(\hat{x}, \hat{y}^*, 1) + \mathscr{L}_x(\hat{x}, \hat{y}^*, 1)[th + r(t)] + \\ + \frac{1}{2}\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, 1)[th + r(t), th + r(t)] + o(t^2) = \\ = f(\hat{x}) + \frac{t^2}{2}\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, 1)[h, h] + o(t^2). \end{aligned}$$

Отсюда немедленно следует (3). ■

Теорема 2 (достаточное условие минимума). *Пусть выполнены условия предыдущей теоремы и, кроме того, для некоторого $\alpha > 0$ выполнено неравенство*

$$\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, 1)[h, h] \geqslant 2\alpha \|h\|^2, \quad \forall h \in \operatorname{Ker} F'(\hat{x}). \qquad (5)$$

Тогда $\hat{x}$ есть точка локального минимума в задаче (1).

Доказательство. Без ограничения общности можно считать, что $f(\hat{x}) = 0$. Обозначим через $B(h_1, h_2)$ билинейную форму $\frac{1}{2}\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, 1)[h_1, h_2]$. По теореме о смешанных производных (п. 2.2.5) B — непрерывная симметричная билинейная форма. Выберем $\varepsilon > 0$ так, чтобы

$$\varphi(\varepsilon) = \alpha(1-\varepsilon)^2 - 2\|B\|(1+\varepsilon)\varepsilon - \|B\|\varepsilon^2 - \frac{\alpha}{2} > 0 \qquad (6)$$

(поскольку $\varphi(0) = \alpha/2 > 0$, это сделать можно).

Функции F и $\mathscr{L}(\cdot, \hat{y}^*, 1)$ по условию имеют в точке $\hat{x}$ вторые производные Фреше по x. Воспользовавшись формулой Тейлора (теорема 2 п. 2.2.5) и учтя, что $F(\hat{x}) = 0$, $\mathscr{L}(\hat{x}, \hat{y}^*, 1) = 0$, $\mathscr{L}_x(\hat{x}, \hat{y}^*, 1) = 0$, найдем такое $\delta_1 > 0$, чтобы при $\|h\| < \delta_1$ выполнялись неравенства

$$\begin{aligned} \|F(\hat{x}+h) - F'(\hat{x})h\| = \|F(\hat{x}+h) - F(\hat{x}) - F'(\hat{x})h\| \leqslant C_1\|h\|^2, \\ |\mathscr{L}(\hat{x}+h, \hat{y}^*, 1) - B(h, h)| \leqslant \frac{\alpha}{2}\|h\|^2. \end{aligned} \qquad (7)$$

Применяя лемму из п. 2.1.5, построим правое обратное к $F'(\hat{x})$ отображение М: $Y \to X$, $F'(\hat{x}) \circ \mathrm{M} = I$, $\|\mathrm{M}(y)\| \leqslant C\|y\|$. По выбранному ранее $\varepsilon > 0$ подберем δ, $0 < \delta < \delta_1$ так, чтобы

$$\delta CC_1 < \varepsilon. \tag{8}$$

Пусть теперь $\|h\| < \delta$ и $\hat{x}+h$ — допустимый элемент в задаче, т. е. $F(\hat{x}+h) = 0$ Положим $h_2 = \mathrm{M}(F'(\hat{x})h)$ и обозначим через h_1 разность $h - h_2$. Тогда из оценки для $\mathrm{M}(y)$, (7) и (8) имеем

$$\|h_2\| \leqslant C\|F'(\hat{x})h\| \leqslant CC_1\|h\|^2 < \varepsilon\|h\|, \tag{9}$$

$$F'(\hat{x})h_1 = F'(\hat{x})(h - h_2) = F'(\hat{x})h - F'(\hat{x})\mathrm{M}(F'(\hat{x})h)) = 0.$$

Таким образом, $h_1 \in \operatorname{Ker} F'(\hat{x})$. Из (9) вытекает, что $(1-\varepsilon)\|h\| \leqslant \|h_1\| \leqslant (1+\varepsilon)\|h\|$. В итоге получаем, учитывая (6), (5) и (7):

$$f(\hat{x}+h) = \mathscr{L}(\hat{x}+h,\ \hat{y}^*,\ 1) \geqslant B(h,\ h) - \frac{\alpha}{2}\|h\|^2 =$$
$$= B(h_1+h_2,\ h_1+h_2) - \frac{\alpha}{2}\|h\|^2 \geqslant$$
$$\geqslant B(h_1,\ h_1) - 2\|B\|\|h_1\|\|h_2\| - \|B\|\|h_2\|^2 - \frac{\alpha}{2}\|h\|^2 \geqslant$$
$$\geqslant \left(\alpha(1-\varepsilon)^2 - 2\|B\|(1+\varepsilon)\varepsilon - \|B\|\varepsilon^2 - \frac{\alpha}{2}\right)\|h\|^2 > 0,$$

т. е. $\hat{x} \in \operatorname{locmin}(1)$. ■

3.4.2. Гладкие задачи с равенствами и неравенствами — необходимые условия второго порядка. Рассмотрим задачу:

$$f_0(x) \to \inf, \quad F(x) = 0, \quad f_i(x) \leqslant 0, \quad i = 1, \ldots, m. \tag{$\mathfrak{z}$}$$

Меняя если нужно, знаки функций мы можем свести к ($\mathfrak{z}$) любую задачу из § 3.2. Функция Лагранжа задачи ($\mathfrak{z}$) имеет вид

$$\mathscr{L}(x,\ y^*,\ \lambda,\ \lambda_0) = \sum_{i=0}^{m} \lambda_i f_i(x) + \langle y^*,\ F(x)\rangle.$$

Теорема. *Пусть X, Y — банаховы пространства, U — открытое множество в X, $\hat{x} \in U$, функции $f_i\colon U \to \mathbf{R}$, $i = 0, 1, \ldots, m$ и отображение $F\colon U \to Y$ строго дифференцируемы в точке $\hat{x}$, имеют в этой точке вторую производную Фреше и, кроме того, $f_i(\hat{x}) = 0$, $i = 1, \ldots, m$.*

Если $\hat{x}$ *доставляет локальный минимум в задаче* ($\mathfrak{z}$), *и если* F *регулярно в точке* $\hat{x}$ (*т. е.* $\operatorname{Im} F'(\hat{x}) = Y$), *то:*

а) *множество* D *множителей Лагранжа* $(y^*, \lambda, \lambda_0)$, $y^* \in Y^*$, $\lambda \in \mathbf{R}^{m*}$, λ_0 *таких, что*

$$\begin{gathered}\lambda_0 \geqslant 0, \quad \lambda_i \geqslant 0, \quad \sum_{i=0}^{m} \lambda_i = 1, \\ \mathscr{L}_x(\hat{x}, y^*, \lambda, \lambda_0) = 0\end{gathered} \tag{1}$$

является непустым выпуклым компактом в $Y^* \times \mathbf{R}^{m*} \times \mathbf{R}$;

б) *для любого* h_0, *лежащего в подпространстве*

$$L = \{h \mid \langle f_i'(\hat{x}), h\rangle = 0, \quad i \geqslant 0, \quad F'(\hat{x}) h = 0\}, \tag{2}$$

найдутся такие множители Лагранжа $(y^*(h_0), \lambda(h_0), \lambda_0(h_0)) \in D$, *что*

$$\mathscr{L}_{xx}(\hat{x}, y^*(h_0), \lambda(h_0), \lambda_0(h_0))[h_0, h_0] \geqslant 0. \tag{3}$$

Доказательство. Обозначим для краткости $F'(\hat{x}) = \Lambda$, $f_i'(\hat{x}) = x_i^*$.

А) Рассмотрим отображение $\varphi\colon \sigma \to X^*$ симплекса $\sigma = \left\{(\lambda, \lambda_0) \mid \lambda_i \geqslant 0, \ \sum_{i=0}^{m} \lambda_i = 1\right\}$, определяемое равенством $\varphi(\lambda, \lambda_0) = \sum_{i=0}^{m} \lambda_i x_i^*$. Тогда

$$\begin{aligned}(y^*, \lambda, \lambda_0) \in D \Leftrightarrow \sum_{i=0}^{m} \lambda_i x_i^* + y^* \circ \Lambda &= \\ = \varphi(\lambda, \lambda_0) + \Lambda^* y^* = 0 \Rightarrow \varphi(\lambda, \lambda_0) &\in \\ \in \operatorname{Im} \Lambda^* \Rightarrow (\lambda, \lambda_0) \in \sigma_1 &= \varphi^{-1}(\operatorname{Im} \Lambda^*).\end{aligned}$$

По условию $\operatorname{Im} \Lambda = Y$, а следовательно (п. 2.1.7), $\operatorname{Im} \Lambda^* = (\operatorname{Ker} \Lambda)^\perp$ и потому замкнут. Значит, σ_1 — замкнутое подмножество компакта и само компакт.

Заметим теперь, что $\operatorname{Ker} \Lambda^* = \{0\}$. Действительно:

$$\begin{aligned}h^* \in \operatorname{Ker} \Lambda^* \Rightarrow \langle \Lambda^* h^*, x\rangle = 0, \ \forall x \Rightarrow \langle h^*, \Lambda x\rangle &= \\ = 0, \ \forall x \Rightarrow h^* \in (\operatorname{Im} \Lambda)^\perp \Rightarrow h^* &= 0.\end{aligned}$$

Так как замкнутое подпространство $\operatorname{Im} \Lambda^*$ банахова пространства само есть банахово пространство, по теореме Банаха существует обратное к Λ^* отображение

$\Gamma : \operatorname{Im}\Lambda^* \to Y^*$. Теперь

$$(y^*, \lambda, \lambda_0) \in D \Leftrightarrow (\lambda, \lambda_0) \in \sigma_1,$$
$$\varphi(\lambda, \lambda_0) + \Lambda^* y^* = 0 \Leftrightarrow (\lambda, \lambda_0) \in \sigma_1,$$
$$y^* = -\Gamma\varphi(\lambda, \lambda_0).$$

Следовательно, D есть образ компакта σ_1 при непрерывном отображении $(\lambda, \lambda_0) \mapsto (-\Gamma\varphi(\lambda, \lambda_0), \lambda_1, \lambda_0)$ и, значит, является компактом.

Непустота D следует из принципа Лагранжа, доказанного в § 3.2. Выпуклость D является очевидным следствием условий (1). Впрочем, ее можно вывести из выпуклости σ и $\operatorname{Im}\Lambda^*$, заметив, что рассмотренные выше отображения были линейными. Этим закончено доказательство утверждения а).

Б) Как и в п. 3.3.1, заменим задачу с неравенствами задачей с равенствами

$$f(x) = \max\{f_0(x) - f_0(\hat{x}), f_1(x), \ldots, f_m(x)\} \to \inf;\ F(x) = 0. \quad (\mathfrak{z}')$$

Если $\hat{x} \in \operatorname{locmin} \mathfrak{z}$, то $\hat{x} \in \operatorname{locmin} \mathfrak{z}'$. Действительно,

$$\hat{x} \notin \operatorname{locmin} \mathfrak{z}' \Rightarrow \forall \varepsilon > 0,\ \exists x_\varepsilon:\ \|x_\varepsilon - \hat{x}\| < \varepsilon,$$
$$F(x_\varepsilon) = 0, \quad f(x_\varepsilon) < 0 \Rightarrow f_0(x_\varepsilon) < f_0(\hat{x}), \quad f_i(x_\varepsilon) < 0,$$
$$i = 1, \ldots, m,\ F(x_\varepsilon) = 0 \Rightarrow \hat{x} \notin \operatorname{locmin} \mathfrak{z}.$$

Далее исследуем задачу $(\mathfrak{z}')$.

В) Пусть $h_0 \in L$. Ввиду компактности D найдутся такие $(\hat{y}^*, \hat{\lambda}, \hat{\lambda}_0)$, что

$$\Psi(h_0) = \mathscr{L}_{xx}(\hat{x}, \hat{y}^*, \hat{\lambda}, \hat{\lambda}_0)[h_0, h_0] =$$
$$= \max_{(y^*, \lambda, \lambda_0) \in D} \mathscr{L}_{xx}(\hat{x}, y^*, \lambda, \lambda_0)[h_0, h_0] =$$
$$= \max\left\{\sum_{i=0}^{m} \lambda_i f_i''(\hat{x})[h_0, h_0] + \langle y^*, F''(\hat{x})[h_0, h_0]\rangle \mid \lambda_i \geqslant 0,\right.$$
$$\left.\sum_{i=0}^{n} \lambda_i = 1,\ \sum_{i=0}^{m} \lambda_i x_i^* + \Lambda^* y^* = 0\right\}.$$

Утверждение б) эквивалентно неравенству $\Psi(h_0) \geqslant 0$.

Предположим, что $\Psi(h_0) < 0$. Обозначив

$$a_i(\lambda) = \frac{\lambda^2}{2} f_i''(\hat{x})[h_0, h_0], \quad i = 0, \ldots, m,$$
$$y(\lambda) = \frac{\lambda^2}{2} F''(\hat{x})[h_0, h_0], \tag{4}$$

рассмотрим задачу

$$\max_{0\leqslant i\leqslant m}(a_i(\lambda)+\langle x_i^*,\ x\rangle)\to\inf;\quad \Lambda x+y(\lambda)=0. \qquad (5)$$

Проверим, что к этой задаче применима лемма о минимаксе из п. 3.3.4. Действительно,

$$\max_{0\leqslant i\leqslant m}\langle x_i^*,\ x\rangle\geqslant 0,\quad \forall x\in\operatorname{Ker}\Lambda$$

— необходимое условие минимальности $\hat{x}$ в задаче ($\mathfrak{z}$) (лемма 1 п. 3.2.4). Кроме того, по условию оператор Λ сюръективен.

По лемме о минимаксе найдется элемент $x_0(\lambda)$, обладающий свойствами:

$$\max_{0\leqslant i\leqslant m}(a_i(\lambda)+\langle x_i^*,\ x_0(\lambda)\rangle)=S(a(\lambda),\ y(\lambda)), \qquad (6)$$

где $S(a(\lambda),\ y(\lambda))$ — значение задачи (5);

$$\|x_0(\lambda)\|\leqslant C_1\{\max|a_i(\lambda)|+\|y(\lambda)\|\}\leqslant C\lambda^2. \qquad (7)$$

При этом согласно (8) п. 3.3.4

$$S(a(\lambda),\ y(\lambda))=$$
$$=\frac{\lambda^2}{2}\max\left\{\sum_{i=0}^{m}\lambda_i f_i''(\hat{x})[h_0,\ h_0]+\langle y^*,\ F''(\hat{x})[h_0,\ h_0]\rangle\ \middle|\ \lambda_i\geqslant 0,\right.$$
$$\left.\sum_{i=0}^{m}\lambda_i=1,\quad \sum_{i=0}^{m}\lambda_i x_i^*+\Lambda^*y^*=0\right\}=\frac{\lambda^2}{2}\Psi(h_0). \qquad (8)$$

Но тогда из (7) вытекает, что $\|\lambda h_0+x_0(\lambda)\|=O(\lambda)$. По формуле Тейлора (п. 2.2.5)

$$F(\hat{x}+x_0(\lambda)+\lambda h_0)=F(\hat{x})+F'(\hat{x})x_0(\lambda)+$$
$$+\frac{1}{2}F''(\hat{x})[\lambda h_0+x_0(\lambda),\ \lambda h_0+x_0(\lambda)]+o(\lambda^2)=$$
$$=F''(\hat{x})[\lambda h_0,\ x_0(\lambda)]+\frac{1}{2}F''(\hat{x})[x_0(\lambda),\ x_0(\lambda)]+$$
$$+o(\lambda^2)=o(\lambda^2)$$

(напомним, что

$$h_0\in L\subset\operatorname{Ker}F'(\hat{x})$$

и

$$F'(\hat{x})x_0(\lambda)+(\lambda^2/2)F''(\hat{x})[h_0,\ h_0]=0.)$$

При доказательстве теоремы Люстерника в п. 2.3.5 было построено отображение φ: $U\to X$ окрестности $U\ni\hat{x}$

такое, что

$$F(x+\varphi(x)) \equiv 0 \quad \text{и} \quad \|\varphi(x)\| \leqslant K\|F(x)\|.$$

Полагая $r(\lambda) = \varphi(\hat{x}+x_0(\lambda)+\lambda h_0)$, имеем

$$F(\hat{x}+x_0(\lambda)+\lambda h_0+r(\lambda)) \equiv 0,$$

$$\|r(\lambda)\| \leqslant K\|F(\hat{x}+x_0(\lambda)+\lambda h_0)\| = o(\lambda^2).$$

Но тогда, применяя формулу Тейлора к f_i и используя (6) и (8), получим

$$f(\hat{x}+x_0(\lambda)+\lambda h_0+r(\lambda)) =$$

$$= \max(f_0(\hat{x}+x_0(\lambda)+\lambda h_0+r(\lambda)) - f_0(\hat{x}),$$

$$f_i(\hat{x}+x_0(\lambda)+\lambda h_0+r(\lambda)), \quad i=1, \ldots, m) =$$

$$= \max_{0 \leqslant i \leqslant m}\left(\langle x_i^*, x_0(\lambda)\rangle + \frac{\lambda^2}{2} f_i''(\hat{x})[h_0, h_0] + o(\lambda^2)\right) \leqslant$$

$$\leqslant \max_{0 \leqslant i \leqslant m}\left(\langle x_i^*, x_0(\lambda)\rangle + \frac{\lambda^2}{2} f_i''(\hat{x})[h_0, h_0]\right) + o(\lambda^2) =$$

$$= \frac{\lambda^2}{2}\Psi(h_0) + o(\lambda^2) < 0$$

при малых λ^2. Итак, $\Psi(h_0) < 0 \Rightarrow \hat{x} \notin \operatorname{locmin} \mathfrak{z}' \Rightarrow \hat{x} \notin \operatorname{locmin} \mathfrak{z}$. Противоречие доказывает теорему. ■

3.4.3. Достаточные условия экстремума для гладких задач с равенствами и неравенствами. Как и в предыдущем пункте, исследуем задачу

$$f_0(x) \to \inf; \quad F(x) = 0, \quad f_i(x) \leqslant 0. \qquad (\mathfrak{z})$$

Теорема. *Пусть X, Y — банаховы пространства, U — открытое множество в X, $\hat{x} \in U$, функции $f_i\colon U \to \mathbf{R}$, $i = 0, 1, \ldots, m$, и отображение $F\colon U \to Y$ — строго дифференцируемы в точке $\hat{x}$ (допустимой в ($\mathfrak{z}$)) и имеют в этой точке вторую производную Фреше, причем $f_i(\hat{x}) = 0$, $i = 1, \ldots, m$.*

Предположим, что существуют множители Лагранжа $\hat{\lambda} \in \mathbf{R}^{m*}$, $\hat{y}^* \in Y^*$ и число $\alpha > 0$ такие, что $\hat{\lambda}_i > 0$

$$\mathscr{L}_x(\hat{x}, \hat{y}^*, \hat{\lambda}, 1) = f_0'(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i f_i'(\hat{x}) + F'^*(\hat{x})\hat{y}^* = 0 \quad (1)$$

и

$$\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, \hat{\lambda}, 1)[h, h] \geqslant 2\alpha\|h\|^2 \qquad (2)$$

для любого h, лежащего в подпространстве

$$L=\{h \mid \langle f_i'(\hat{x}),\ h\rangle=0,\quad i\geqslant 1,\quad F'(\hat{x})\,h=0\}. \tag{3}$$

Тогда $\hat{x}$ доставляет локальный минимум в задаче ($\mathfrak{z}$).

Доказательство. А) Пусть $\hat{x}+h$—допустимый элемент, т. е.

$$f_i(\hat{x}+h)\leqslant 0,\quad i\geqslant 1,\quad F(\hat{x}+h)=0. \tag{4}$$

Проведем двумя способами оценку величины $f_0(\hat{x}+h)$. Имеем

$$f_0(\hat{x}+h)=f_0(\hat{x}+h)+\sum_{i=1}^{m}\hat{\lambda}_i f_i(\hat{x}+h)+$$
$$+\langle \hat{y}^*,\ F(\hat{x}+h)\rangle-\sum_{i=1}^{m}\hat{\lambda}_i f_i(\hat{x}+h)=$$
$$=\mathscr{L}(\hat{x}+h,\ \hat{y}^*,\ \hat{\lambda},\ 1)-\sum_{i=1}^{m}\hat{\lambda}_i f_i(\hat{x}+h). \tag{5}$$

Первый способ оценки основан на прямом разложении:

$$f_0(\hat{x}+h)=\mathscr{L}(\hat{x}+h,\ \hat{y}^*,\ \hat{\lambda},\ 1)-\sum_{i=1}^{m}\hat{\lambda}_i f_i(\hat{x}+h)=$$
$$=\mathscr{L}(\hat{x},\ \hat{y}^*,\ \hat{\lambda},\ 1)+\mathscr{L}_x(\hat{x},\ \hat{y}^*,\ \hat{\lambda},\ 1)[h]-$$
$$-\sum_{i=1}^{m}\hat{\lambda}_i f_i(\hat{x})-\sum_{i=1}^{m}\langle\hat{\lambda}_i f_i'(\hat{x}),\ h\rangle+r_1(h)=$$
$$=f_0(\hat{x})-\sum_{i=1}^{m}\langle\hat{\lambda}_i f_i'(\hat{x}),\ h\rangle+r_1(h), \tag{6}$$

где остаточный член $r_1(h)$ есть $O(\|h\|)^2$.

Второй способ основан на том, что вследствие неравенств $\hat{\lambda}_i>0$, $f_i(\hat{x}+h)\leqslant 0$ сумма $\sum_{i=1}^{m}\hat{\lambda}_i f_i(\hat{x}+h)\leqslant 0$ и, следовательно,

$$f_0(\hat{x}+h)\geqslant\mathscr{L}(\hat{x}+h,\ \hat{y}^*,\ \hat{\lambda},\ 1)=$$
$$=\mathscr{L}(\hat{x},\ \hat{y}^*,\ \hat{\lambda},\ 1)+\mathscr{L}_x(\hat{x},\ \hat{y}^*,\ \hat{\lambda},\ 1)[h]+$$
$$+\frac{1}{2}\mathscr{L}_{xx}(\hat{x},\ \hat{y}^*,\ \hat{\lambda},\ 1)[h,\ h]+r_2(h)=$$
$$=f_0(\hat{x})+B(h,\ h)+r_2(h), \tag{7}$$

где через $B(h, h)$ обозначена квадратичная форма $\frac{1}{2}\mathscr{L}_{xx}(\hat{x}, \hat{y}^*, \hat{\lambda}, 1)[h, h]$, а остаточный член $r_2(h) = o(\|h\|^2)$.

Б) Обозначим через K конус, состоящий из тех h, для которых $\langle f_i'(\hat{x}), h\rangle \leqslant 0$, $i \geqslant 1$, $F'(\hat{x})h = 0$.

В силу леммы Хоффмана произвольное h можно разложить в сумму $h_1 + h_2$, где $h_1 \in K$, а h_2 допускает оценку:

$$\|h_2\| \leqslant C_1 \left\{ \sum_{i=1}^{m} \langle f_i'(\hat{x}), h\rangle_+ + \|F'(\hat{x})h\| \right\}. \tag{8}$$

Воспользовавшись далее замечанием, сделанным в п. 3.3.4 после леммы Хоффмана, и формулой (5′) п. 3.3.4 h_1 можно также разложить в сумму $h_1 = h_1' + h_1''$, где $h_1' \in L$, а h_1'' допускает оценку

$$\|h_1''\| \leqslant C_2 \left\{ \sum_{i=1}^{m} \left| \langle f_i'(\hat{x}), h_1\rangle \right| \right\} = C_2 \left\{ -\sum_{i=1}^{m} \langle f_i'(\hat{x}), h_1''\rangle \right\}. \tag{9}$$

Мы использовали то, что $F'(\hat{x})h_1 = 0$ (ибо $h_1 \in K$) и то, что

$$|\langle f_i'(\hat{x}), h_1\rangle| = |\langle f_i'(\hat{x}), h_1' + h_1''\rangle| = |\langle f_i'(\hat{x}), h_1''\rangle| =$$
$$= -\langle f_i'(\hat{x}), h_1''\rangle \text{ (ибо } \langle f_i'(\hat{x}), h_1'\rangle = 0, \text{ а } \langle f_i'(\hat{x}), h_1\rangle \leqslant 0).$$

В) Из (4) получаем

$$\begin{gathered} 0 = F(\hat{x}+h) - F(\hat{x}) = F'(\hat{x})h + r_0(h), \\ 0 \geqslant f_i(\hat{x}+h) = \langle f_i'(\hat{x}), h\rangle + \rho_i(h), \quad i = 1, \ldots, m, \end{gathered} \tag{10}$$

где $\|r_0(h)\|$ и $|\rho_i(h)|$ — величины порядка $O(\|h\|^2)$. Используя эти оценки в (8), находим, что если $\hat{x}+h$ — допустимый элемент, то

$$\|h_2\| \leqslant C_1 \left\{ \sum_{i=1}^{m} |\rho_i(h)| + \|r_0(h)\| \right\}. \tag{11}$$

Фиксировав $\varepsilon_1 \in (0, 1]$, о величине которого мы позаботимся далее, выберем $\delta \in (0, \|B\|^{-1})$ так, чтобы из неравенства $\|h\| \leqslant \delta$ следовали бы неравенства

$$\sum_{i=1}^{m} |\rho_i(h)| + \sum_{j=0}^{1} \|r_j(h)\| \leqslant \varepsilon_1 \|h\|, \quad \|r_2(h)\| \leqslant \frac{\alpha}{2}\|h\|^2. \tag{12}$$

Далее выберем число $A > 0$ так, чтобы

$$AC_2^{-1}\left(\min_i \hat{\lambda}_i\right) - C_1 \max_i (\hat{\lambda}_i \|f_i'(\hat{x})\|) - 1 \geqslant 0, \tag{13}$$

и, наконец, выберем величину ε_1 так, чтобы для $\varepsilon = (C_1+A)\varepsilon_1$ выполнялись бы неравенства

$$\varepsilon < 1, \ \alpha(1-\varepsilon)^2 - 2\|B\|\varepsilon(1+\varepsilon) - \|B\|\varepsilon^2 - \alpha/2 \geqslant 0. \qquad (14)$$

Г) Завершение доказательства. Пусть $\hat{x}+h$ — допустимый элемент. Представим h, как и в Б), в виде суммы $h = h_1' + h_1'' + h_2$. Возможны два случая: а) $\|h_1''\| > A\varepsilon_1\|h\|$ и б) $\|h_1''\| \leqslant A\varepsilon_1\|h\|$ В случае а) имеем из (9) при $\|h\| \leqslant \delta$

$$A\varepsilon_1\|h\| < \|h_1''\| < C_2\left(-\sum_{i=1}^{m}\langle f_i'(\hat{x}), h_1''\rangle\right). \qquad (15)$$

Тогда вследствие (6), (15), (11), (12) и (13) получаем

$$f_0(\hat{x}+h) \overset{(6)}{=}$$

$$= f_0(\hat{x}) - \sum_{i=1}^{m}\langle\hat{\lambda}_i f_i'(\hat{x}), h_1''+h_2\rangle + r_1(h) \overset{(15),\,(11),\,(12)}{\geqslant}$$

$$\geqslant f_0(\hat{x}) + AC_2^{-1}\left(\min_i \hat{\lambda}_i\right)\varepsilon_1\|h\| -$$

$$- \max_i(\hat{\lambda}_i\|f_i'(\hat{x})\|)C_1\varepsilon_1\|h\| - \varepsilon_1\|h\| = f_0(\hat{x}) +$$

$$+ \varepsilon_1\|h\|\left(AC_2^{-1}\left(\min_i \hat{\lambda}_i\right) - C_1\max_i(\hat{\lambda}_i\|f_i'(\hat{x})\|) - 1\right) \overset{(13)}{\geqslant} f_0(\hat{x}).$$

Пусть теперь имеет место случай б). Тогда $\|h_1''\| \leqslant A\varepsilon_1\|h\|$, значит, в силу (11) и (12), если обозначить $h_2' = h_1'' + h_2$, то $\|h_2'\| = \|h_1''+h_2\| \leqslant (A+C_1)\varepsilon_1\|h\| = \varepsilon\|h\|$. Тогда $h = h_1' + h_2'$, где $(1-\varepsilon)\|h\| \leqslant \|h_1'\| \leqslant (1+\varepsilon)\|h\|$. Теперь применим эти неравенства, а также (7), (2), (12) и получаем

$$f_0(\hat{x}+h) \overset{(7)}{\geqslant} f_0(\hat{x}) + B(h, h) + r_2(h) =$$

$$= f_0(\hat{x}) + B(h_1'+h_2', h_1'+h_2') + r_2(h) =$$

$$= f_0(\hat{x}) + B(h_1', h_1') + 2B(h_1', h_2') + B(h_2', h_2') + r_2(h) \overset{(2),\,(12)}{\geqslant}$$

$$\geqslant f_0(\hat{x}) + \left(\alpha(1-\varepsilon)^2 - 2\|B\|\varepsilon(1+\varepsilon) - \|B\|\varepsilon^2 - \frac{\alpha}{2}\right)\|h\|^2 \geqslant$$

$$\geqslant f_0(\hat{x}).$$

Теорема доказана.

ГЛАВА IV

ПРИНЦИП ЛАГРАНЖА В ЗАДАЧАХ КЛАССИЧЕСКОГО ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ И ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Предмет этой главы ясен из ее заглавия. Особое внимание мы уделяем, с одной стороны, выявлению сходства двух разновидностей теории, подчеркивая это едиными обозначениями, а с другой — выяснению различий в классической и современной постановках задач. Сначала мы рассмотрим необходимые условия в так называемой *задаче Лагранжа*. К этой задаче могут быть сведены многие остальные задачи классического вариационного исчисления. Затем будет выведен принцип максимума Понтрягина, являющийся одним из наиболее важных средств в современных задачах оптимального управления. Остальная часть главы посвящена более специальным классам задач и выводу следствий из общей теории. Менее подробно мы останавливаемся на рассмотрении достаточных условий экстремума, ограничиваясь только частными ситуациями.

§ 4.1. Принцип Лагранжа для задачи Лагранжа

4.1.1. Постановка задачи и формулировка теоремы. Зафиксируем замкнутый отрезок $\Delta=[\alpha, \beta]\subset\mathbf{R}$ и рассмотрим банахово пространство

$$\Xi=C^1(\Delta, \mathbf{R}^n)\times C(\Delta, \mathbf{R}^r)\times\mathbf{R}\times\mathbf{R},$$

состоящее из элементов $\xi=(x(\cdot), u(\cdot), t_0, t_1)$.

В этом пространстве рассмотрим задачу:

$$\mathscr{B}_0(x(\cdot),\ u(\cdot),\ t_0,\ t_1)=$$
$$=\int_{t_0}^{t_1} f_0(t,\ x(t),\ u(t))\,dt+\psi_0(t_0,\ x(t_0),\ t_1,\ x(t_1))\to \text{extr}, \quad (1)$$
$$\dot{x}=\varphi(t,\ x(t),\ u(t)), \quad (2)$$
$$\mathscr{B}_i(x(\cdot),\ u(\cdot),\ t_0,\ t_1)=$$
$$=\int_{t_0}^{t_1} f_i(t,\ x(t),\ u(t))\,dt+\psi_i(t_0,\ x(t_0),\ t_1,\ x(t_1))\lesseqgtr 0. \quad (3)$$

При этом в (1)—(3) f_i: $V\to\mathbf{R}$, ψ_i: $W\to\mathbf{R}$, $i=0,\ 1,\ \ldots$ $\ldots,\ m$; φ: $V\to\mathbf{R}^n$, где V и W — открытые множества в пространствах $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^r$ и $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}\times\mathbf{R}^n$ соответственно. Все перечисленные функции предполагаются по крайней мере непрерывными. Знак $\lesseqgtr$ имеет тот же смысл, что и в § 3.2.

Задачу (1)—(3) мы называем *задачей Лагранжа в понтрягинской форме*. Частные случаи этой задачи обсуждались в §§ 1.4, 1.5 и в § 3.1. Функционалы того же типа, что и $\mathscr{B}_i$, содержащие как интегральную, так и терминальную части, были ранее названы функционалами Больца (см. пп. 1.4.2 и 3.1.3). Если в ограничении $\mathscr{B}_i\gtreqless 0$ терминальная часть является константой, т. е. оно принимает вид $\int_{t_0}^{t_1} f_i(t,\ x,\ u)\,dt\gtreqless a_i$, то вслед за Эйлером мы называем такое ограничение *изопериметрическим*. Если, наоборот, отсутствует интегральный член, то ограничение $\psi_i(t_0,\ x(t_0),\ t_1,\ x(t_1))\gtreqless 0$ называется *граничным*, или *краевым условием*.

Ограничение (2) называется *дифференциальной связью*. Такой вид дифференциальной связи — разрешенный относительно $\dot{x}$ — является характерным признаком понтрягинской формы задачи. Лагранж задает дифференциальную связь уравнением $\psi(t,\ x,\ \dot{x})=0$ (вкратце об этом говорилось в п. 1.5.1).

Наконец, в отличие от гл. I, отрезок времени $[t_0,\ t_1]$ в задаче (1)—(3) не предполагается фиксированным.

Четверку $\xi=(x(\cdot),\ u(\cdot),\ t_0,\ t_1)\in\Xi$ будем называть *управляемым процессом* в задаче (1)—(3), если $t_0,\ t_1\in\operatorname{int}\Delta$, $(t,\ x(t),\ u(t))\in V$ для $t\in\Delta$ и всюду на $[t_0,\ t_1]$ удовлетворяется дифференциальная связь, т. е. $\dot{x}(t)=\varphi(t,\ x(t),$

$u(t)$), и *допустимым управляемым процессом*, если эта четверка является управляемым процессом и, кроме того, выполнены условия (3).

Четверка $\hat{\xi}=(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1)$ называется *оптимальным в слабом смысле процессом*, или *слабым экстремумом* в задаче (1)—(3), если она является локальным экстремумом в пространстве Ξ, т. е. если найдется такое $\varepsilon>0$, что для любого допустимого управляемого процесса ξ, удовлетворяющего условию $\|\xi-\hat{\xi}\|_{\Xi}<\varepsilon$, выполняется одно из неравенств: $\mathscr{B}_0(\xi)\geqslant\mathscr{B}_0(\hat{\xi})$ в случае минимума и $\mathscr{B}_0(\xi)\leqslant\leqslant\mathscr{B}_0(\hat{\xi})$ в случае максимума.

Функцией Лагранжа этой задачи называется функция

$$\mathscr{L}(x(\cdot),\ u(\cdot),\ t_0,\ t_1;\ p(\cdot),\ \lambda,\ \lambda_0)=\int_{t_0}^{t_1} L\,dt+l, \tag{4}$$

где

$$\lambda_0\in\mathbf{R},\quad p(\cdot)\in C^1(\Delta,\ \mathbf{R}^{n*}),\quad \lambda=(\lambda_1,\ \ldots,\ \lambda_m)\in\mathbf{R}^{m*}$$

— *множители Лагранжа*,

$$L(t,\ x,\ \dot{x},\ u)=\sum_{i=0}^{m}\lambda_i f_i(t,\ x,\ u)+p(t)(\dot{x}-\varphi(t,\ x,\ u)) \tag{5}$$

— *лагранжиан*, или *интегрант*, и

$$l(t_0,\ x_0,\ t_1,\ x_1)=\sum_{i=0}^{m}\lambda_i\psi_i(t_0,\ x_0,\ t_1,\ x_1) \tag{6}$$

— *терминант*.

Далее мы в этом и следующем параграфах пользуемся, как и ранее, следующими сокращениями:

$$\hat{L}_{\dot{x}}(t)=L_{\dot{x}}(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t)),$$
$$\hat{L}_x(t)=L_x(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t)),$$
$$\hat{L}_u(t)=L_u(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t)),$$
$$\hat{l}_{x_k}=l_{x_k}(\hat{t}_0,\ \hat{x}(t_0),\ \hat{t}_1,\ \hat{x}(t_1)),$$
$$\hat{\mathscr{L}}_{t_k}=\mathscr{L}_{t_k}(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1;\ \hat{p}(\cdot),\ \hat{\lambda},\ \hat{\lambda}_0)\ \text{и т. п.}$$

Теорема Эйлера—Лагранжа. *Пусть функции* $f_i\colon V\to\mathbf{R}$, $i=0,\ 1,\ \ldots,\ m$, $\varphi\colon V\to\mathbf{R}^n$ *и их частные производные по* x *и* u *непрерывны в открытом множестве* V *пространства* $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^r$, *а функции* $\psi_i\colon W\to\mathbf{R}$, $i=$

$=0, 1, \ldots, m$, *непрерывно дифференцируемы в открытом множестве* W *пространства* $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}\times\mathbf{R}^n$, *и пусть* $\hat{x}(\cdot)\in C^1(\Delta, \mathbf{R}^n)$, $\hat{u}(\cdot)\in C(\Delta, \mathbf{R}^r)$, $\hat{t}_0$, $\hat{t}_1\in \operatorname{int}\Delta$ *таковы, что*

$$(t, \hat{x}(t), \hat{u}(t))\in V, \quad t\in\Delta, \quad (\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1))\in W.$$

Если $\hat{\xi}=(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$ *является оптимальным в слабом смысле процессом для задачи* (1)—(3), *то найдутся множители Лагранжа* $\hat{\lambda}_0\geqslant 0$ *в случае задачи на минимум и* $\leqslant 0$ *в задаче на максимум*, $\hat{p}(\cdot)\in C^1(\Delta, \mathbf{R}^{n*})$, $\hat{\lambda}=(\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$, *не равные одновременно нулю и такие, что:*

а) *выполнены условия стационарности функции Лагранжа:*

по $x(\cdot)$ $(\hat{\mathscr{L}}_{x(\cdot)}=0)$:

$$-\frac{d}{dt}\hat{L}_{\dot{x}}(t)+\hat{L}_x(t)\equiv 0, \tag{7}$$

$$\hat{L}_{\dot{x}}(\hat{t}_k)=(-1)^k\hat{l}_{x_k}, \quad k=0, 1; \tag{8}$$

по $u(\cdot)$ $(\hat{\mathscr{L}}_{u(\cdot)}=0)$:

$$\hat{L}_u(t)\equiv 0; \tag{9}$$

по t_k

$$\hat{\mathscr{L}}_{t_k}=0, \quad k=0, 1; \tag{10}$$

б) *выполнено условие согласования знаков*:

$$\hat{\lambda}_i\gtreqless 0, \quad i=1, \ldots, m; \tag{11}$$

в) *выполнены условия дополняющей нежесткости:*

$$\hat{\lambda}_i\mathscr{B}_i(\hat{\xi})=0, \quad i=1, \ldots, m. \tag{12}$$

Как и в §§ 3.1, 3.2, неравенства (11) означают, что $\hat{\lambda}_i\geqslant 0$, если в условии (3) $\mathscr{B}_i(\xi)\leqslant 0$, $\hat{\lambda}_i\leqslant 0$, если в (3) $\mathscr{B}_i(\xi)\geqslant 0$ и $\hat{\lambda}_i$ может иметь любой знак, если $\mathscr{B}_i(\xi)=0$.

Утверждение о существовании множителей Лагранжа, удовлетворяющих совокупности условий а)—в), кратко называется нами *принципом Лагранжа для задачи Лагранжа* (1)—(3).

Покажем, что утверждение теоремы находится в полном соответствии с общим принципом Лагранжа, о котором говорилось в гл. 1 и в п. 3.1.5. Действительно, функция Лагранжа $\mathscr{L}$ является функцией трех аргументов: $x(\cdot)$, $u(\cdot)$ и (t_0, t_1). Таким образом, согласно общему принципу Лагранжа, надлежит рассмотреть три задачи:

$$(\alpha) \quad \mathscr{L}(x(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1; \hat{p}(\cdot), \hat{\lambda}, \hat{\lambda}_0) \to \mathrm{extr},$$

$$(\beta) \quad \mathscr{L}(\hat{x}(\cdot), u(\cdot), \hat{t}_0, \hat{t}_1; \hat{p}(\cdot), \hat{\lambda}, \hat{\lambda}_0) \to \mathrm{extr},$$

$$(\gamma) \quad \mathscr{L}(\hat{x}(\cdot), \hat{u}(\cdot), t_0, t_1; \hat{p}(\cdot), \hat{\lambda}, \hat{\lambda}_0) \to \mathrm{extr}.$$

Задача (α) является элементарной задачей Больца, и условия стационарности (7) и (8) написаны в полном соответствии с п. 1.4.2 и теоремой п. 3.1.3. Задача (β) также является задачей Больца, но вырожденной, поскольку $\dot{u}$ не входит в лагранжиан, а $u(t_0)$ и $u(t_1)$ — в терминант. Уравнение Эйлера — Лагранжа поэтому превращается в (9), а условия трансверсальности становятся бессодержательными: $0=0$. Наконец, (γ) является элементарной задачей и (10) есть просто теорема Ферма (см. пп. 1.3.1 и 3.1.1).

Условия дополняющей нежесткости и условия согласования знаков также, как мы знаем, написаны в соответствии с общим принципом Лагранжа, примененным к ограничениям типа неравенств (п. 3.1.5)

Таким образом, основную теорему этого параграфа можно сформулировать так: *если функции, входящие в постановку задачи, обладают достаточной гладкостью, то для локального экстремума выполнен принцип Лагранжа.*

Условия стационарности можно переписать в более развернутом виде. В соответствии с (5) получаем:

из (7)

$$\frac{d\hat{p}(t)}{dt} = -\hat{p}(t)\hat{\varphi}_x(t) + \sum_{i=0}^{m} \hat{\lambda}_i \hat{f}_{ix}(t) \tag{7a}$$

(иногда (7a) называют *сопряженным уравнением*);

из (9)

$$0 = -\hat{p}(t)\hat{\varphi}_u(t) + \sum_{i=0}^{m} \hat{\lambda}_i \hat{f}_{iu}(t); \tag{9a}$$

из (8)

$$\hat{p}(\hat{t}_k) = (-1)^k \hat{l}_{x_k}, \quad k=0,\ 1; \tag{8a}$$

и, наконец, из (10)

$$0=(-1)^{k-1}\hat{L}(\hat{t}_k)+\hat{l}_{t_k}+\hat{l}_{x_k}\dot{\hat{x}}(\hat{t}_k)=$$

$$=(-1)^{k-1}\left[\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_i(\hat{t}_k)+\hat{p}(\hat{t}_k)(\dot{\hat{x}}(\hat{t}_k)-\hat{\varphi}(\hat{t}_k))\right]+$$

$$+\hat{l}_{t_k}+(-1)^k\hat{p}(\hat{t}_k)\dot{\hat{x}}(\hat{t}_k)=$$

$$=(-1)^{k-1}\left[\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_i(\hat{t}_k)-\hat{p}(\hat{t}_k)\hat{\varphi}(\hat{t}_k)\right]+\hat{l}_{t_k}$$

или

$$\hat{p}(\hat{t}_k)\hat{\varphi}(\hat{t}_k)-\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_i(\hat{t}_k)=(-1)^{k-1}\hat{l}_{t_k}. \tag{10a}$$

Функцию

$$H(t,\ x,\ p,\ u)=L_{\dot{x}}\dot{x}-L=$$

$$=p\varphi(t,\ x,\ u)-\sum_{i=0}^{m}\hat{\lambda}_i f_i(t,\ x,\ u) \tag{13}$$

мы будем называть *функцией Понтрягина* рассматриваемой задачи. С ее помощью уравнения (7а)—(10а) и уравнение (2) можно записать еще и так:

$$\frac{d\hat{x}}{dt}=\hat{H}_p,\quad \frac{d\hat{p}}{dt}=-\hat{H}_x,\quad \hat{H}_u=0 \tag{14}$$

и

$$\hat{p}(\hat{t}_k)=(-1)^k\hat{l}_{x_k},\quad \hat{H}(\hat{t}_k)=(-1)^{k-1}\hat{l}_{t_k},\quad k=0,\ 1. \tag{15}$$

Замечание. Иногда функцию (13) называют функцией Гамильтона. Мы вкладываем в термин «функция Гамильтона» или «гамильтониан» другой смысл, более естественный с точки зрения классической механики, и называем так функцию

$$\mathcal{H}(t,\ x,\ p)=H(t,\ x,\ p,\ u(t,\ x,\ p)), \tag{16}$$

где $u(t,\ x,\ p)$—неявная функция, определяемая из уравнения $H_u(t,\ x,\ p,\ u)=0$. При выполнении стандартных условий применимости теоремы о неявной функции уравнения (14) принимают обычный вид канонической гамильтоновой системы

$$\frac{d\hat{x}}{dt}=\mathcal{H}_p,\quad \frac{d\hat{p}}{dt}=-\mathcal{H}_x. \tag{17}$$

Полезно заметить, что набор условий, который доставляет сформулированная теорема, в некотором смысле является полным. Действительно, для определения неизвестных функций $(x(\cdot), p(\cdot), u(\cdot))$ мы имеем систему из дифференциальных уравнений (2) и (7а) и конечного уравнения (9а). Выражая из последнего (разумеется, когда это можно сделать, например, если выполнены условия теоремы о неявной функции) $u(\cdot)$ через $x(\cdot)$ и $p(\cdot)$, мы получаем систему из $2n$ скалярных дифференциальных уравнений (эквивалентную системе (17)). Ее общее решение зависит от $2n$ произвольных постоянных и еще от множителей Лагранжа λ_i, среди которых m независимых. Добавляя сюда еще t_0 и t_1, мы получаем всего $2n+m+2$ неизвестных. Для их определения мы имеем $2n+2$ условия трансверсальности (15) и m условий дополняющей нежесткости (12).

Таким образом, число неизвестных совпадает с числом уравнений. Именно это и имелось в виду, когда выше было сказано о «полноте» набора условий. Разумеется, разрешимости полученной системы уравнений указанное обстоятельство никак не гарантирует.

4.1.2. Редукция задачи Лагранжа к гладкой задаче. Обозначим через Y пространство $C(\Delta, \mathbf{R}^n)$ и запишем задачу (1)—(3) п. 4.1.1 в виде

$$\mathscr{B}_0(\xi)\to \mathrm{extr};\quad \Phi(\xi)=0,\quad \mathscr{B}_i(\xi)\lesseqgtr 0,\quad i=1, \ldots, m, \tag{1}$$

где

$$\Phi(\xi)(t)=\dot{x}(t)-\varphi(t, x(t), u(t)). \tag{2}$$

Отображение Φ и функционалы $\mathscr{B}_i$ определены в области

$$\mathscr{V}=\{(x(\cdot), u(\cdot), t_0, t_1) \mid (t, x(t), u(t))\in V,\ t\in\Delta;$$
$$t_0, t_1\in \operatorname{int}\Delta,\ (t_0, x(t_0), t_1, x(t_1))\in W\}$$

пространства Ξ.

Задача (1) имеет тот же вид, что задача п. 3.2.1 предыдущей главы, для которой принцип Лагранжа был уже доказан. Предположения соответствующей теоремы распадались на три группы: *банаховость* основных пространств, *гладкость* отображений и *замкнутость образа* бесконечномерного отображения. Проверим, что в задаче (1) все эти требования выполняются.

Банаховость пространств Ξ и Y очевидна — пространства C^1 и C являются основными для нас примерами банаховых пространств (см. пп. 2.1.1 и 2.1.2).

Гладкость (а именно непрерывная дифференцируемость по Фреше) функционалов $\mathscr{B}_i$, $i=0, 1, \ldots, m$, следует из результатов § 2.4.

Действительно, с интегральной частью мы можем поступать точно так же, как и в предложении 2 п. 2.4.2, заменив всюду $\dot{x}(\cdot)$ на $u(\cdot)$, а дифференцируемость терминальной части доказана в п. 2.4.3 (это оператор краевых условий). При этом, согласно формулам (9) п. 2.4.2 и (3) п. 2.4.3, имеем для $\hat{\xi}=(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$, $\eta = (h(\cdot), v(\cdot), \tau_0, \tau_1)$

$$\mathscr{B}_i'(\hat{\xi})[\eta]=\int_{\hat{t}_0}^{\hat{t}_1}(\hat{f}_{ix}(t)h(t)+\hat{f}_{iu}(t)v(t))\,dt+ \\ +\hat{f}_i(\hat{t}_1)\tau_1-\hat{f}_i(\hat{t}_0)\tau_0+\hat{l}_{t_0}\tau_0+\hat{l}_{t_1}\tau_1+\hat{l}_{x_0}h(\tau_0)+ \\ +\hat{l}_{x_1}h(\hat{t}_1)+\hat{l}_{x_0}\dot{\hat{x}}(\hat{t}_0)\tau_0+\hat{l}_{x_1}\dot{\hat{x}}(\hat{t}_1)\tau_1 \tag{3}$$

(ср. лемму п. 3.1.3).

Отображение (2) также непрерывно дифференцируемо по Фреше (см. предложение 3 п. 2.4.1). Его производная, согласно (14) п. 2.4.1, имеет вид

$$\Phi'(\hat{\xi})[\eta](t)=\dot{h}(t)-\hat{\varphi}_x(t)h(t)-\hat{\varphi}_u(t)v(t). \tag{4}$$

Из непрерывной дифференцируемости по Фреше функционалов $\mathscr{B}_i$ и отображения Φ следует их строгая дифференцируемость, требуемая в теореме о принципе Лагранжа.

Замкнутость образа отображения Φ имеет место в силу его регулярности: образ $\Phi'(\hat{\xi})\Xi$ совпадает с Y. Действительно, взяв произвольное $y(\cdot)\in Y = C(\Delta, \mathbf{R}^n)$, положим $v(\cdot)=0$, $\tau_0=\tau_1=0$. Уравнение

$$\Phi'(\hat{\xi})[h(\cdot), 0, 0, 0](\cdot)=y(\cdot)$$

эквивалентно линейному дифференциальному уравнению $\dot{h}(t)-\hat{\varphi}_x(t)h(t)=y(t)$ с непрерывными коэффициентами, которое имеет решение $h(\cdot)\in C^1(\Delta, \mathbf{R}^n)$ в силу теоремы существования для линейных систем (п. 2.5.4). Таким образом, выполнены все условия теоремы п. 3.2.1.

Для определенности будем считать, что (1) — задача на минимум. Согласно принципу Лагранжа, если $\hat{\xi} = (\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$ — локальная минималь (1), то найдутся множители Лагранжа $\hat{\lambda}_0$, $\hat{\lambda}\in\mathbf{R}^{m*}$ и элемент $\hat{y}^*\in$

$\in Y^* = (C(\Delta, \mathbf{R}^n))^*$ такие, что для функции Лагранжа

$$\tilde{\mathscr{L}}(\xi; y^*, \lambda, \lambda_0) = \tilde{\mathscr{L}}(x(\cdot), u(\cdot), t_0, t_1; y^*, \lambda, \lambda_0) =$$
$$= \int_{t_0}^{t_1} \left(\sum_{i=0}^{m} \lambda_i f_i(t, x, u) \right) dt + \sum_{i=0}^{m} \lambda_i \psi_i(t_0, x(t_0), t_1, x(t_1)) +$$
$$+ y^* \circ \Phi \quad (5)$$

выполнены условия согласования знаков, дополняющей нежесткости и стационарности

$$\hat{\tilde{\mathscr{L}}}_\xi = 0 \Longleftrightarrow \hat{\tilde{\mathscr{L}}}_{x(\cdot)} = 0, \quad \hat{\tilde{\mathscr{L}}}_{u(\cdot)} = 0, \quad \hat{\tilde{\mathscr{L}}}_{t_k} = 0, \quad k = 0, 1.$$

Три соотношения, из которых складывается правая часть, мы называем далее *стационарностью* по $x(\cdot)$, $u(\cdot)$ и по t_k соответственно.

Однако мы замечаем, что функция (5) не тождественна функции Лагранжа (4) п. 4.1.1. Воспользовавшись теоремой Рисса об общем виде непрерывного линейного функционала на пространстве $C(\Delta, \mathbf{R}^n)$ (п. 2.1.9), мы можем утверждать лишь, что

$$\langle y^*, \eta(\cdot)\rangle = \int_\Delta d\nu(t)\, \eta(t),$$

где $\nu(t) = (\nu_1(t), \ldots, \nu_n(t))$—вектор-строчка из канонических функций ограниченной вариации. Подставляя это в (5) и учитывая обозначение (6) из п. 4.1.1, получаем

$$\tilde{\mathscr{L}} = \int_{t_0}^{t_1} \left(\sum_{i=0}^{m} \lambda_i f_i(t, x(t), u(t)) \right) dt +$$
$$+ \int_\Delta d\nu(t) \{\dot{x}(t) - \varphi(t, x(t), u(t))\} +$$
$$+ l(t_0, x(t_0), t_1, x(t_1)). \quad (6)$$

Сравнивая эту формулу с (4)—(6) из п. 4.1.1, мы видим, что функция Лагранжа $\mathscr{L}$ получается из $\tilde{\mathscr{L}}$, если функция $\nu(\cdot)$ абсолютно непрерывна, а ее производная равна нулю вне $[t_0, t_1]$ и совпадает с $p(\cdot)$ в этом отрезке. Ситуация напоминает здесь п. 1.4.1. Подобно тому как лемма Дюбуа-Реймона позволила нам там вывести уравнение Эйлера без дополнительного предположения о дифференцируемости функции $t \to p(t) = \hat{L}_{\dot{x}}(t)$, подходящее обобщение этой леммы позволит нам сейчас доказать абсолютную непрерывность функции $\nu(\cdot)$.

4.1.3. Обобщенная лемма Дюбуа-Реймона. *Пусть на отрезке* $[\alpha, \beta]$ *векторная функция* $a(\cdot)\colon [\alpha, \beta] \to \mathbf{R}^{n*}$ *интегрируема по Лебегу, матричная функция* $b(\cdot)\colon [\alpha, \beta] \to \mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$ *непрерывна,* $\nu(\cdot)\colon [\alpha, \beta] \to \mathbf{R}^{n*}$ *— каноническая векторная функция ограниченной вариации;* t_k *— различные точки интервала* (α, β), $c_k \in \mathbf{R}^{n*}$, $k=0, 1, \ldots, p$.

Если для всех $h(\cdot) \in C^1([\alpha, \beta], \mathbf{R}^n)$ *имеет место равенство*

$$\int_\alpha^\beta a(t)\, h(t)\, dt + \int_\alpha^\beta d\nu(t)\, \{\dot{h}(t) - b(t)\, h(t)\} + \sum_{k=0}^{p} c_k h(t_k) = 0, \tag{1}$$

то функция $\nu(\cdot)$ *абсолютно непрерывна; ее производная* $p(\cdot)$ *непрерывна на* $[\alpha, \beta]$, *кроме, быть может, точек* t_k, *в которых она имеет пределы слева и справа, абсолютно непрерывна в любом интервале, не содержащем точек* t_k, *и при этом*

$$\dot{p}(t) = a(t) - p(t)\, b(t) \ \textit{почти всюду}, \tag{2}$$

$$p(\alpha) = p(\beta) = 0, \tag{3}$$

$$p(t_k + 0) - p(t_k - 0) = c_k. \tag{4}$$

Доказательство. Положив $h(t) = \gamma + \int_\alpha^t \eta(s)\, ds$, мы получаем из (1)

$$\left\{\int_\alpha^\beta a(t)\, dt - \int_\alpha^\beta d\nu(t)\, b(t) + \sum_{k=0}^{p} c_k\right\}\gamma + \int_\alpha^\beta a(t) \int_\alpha^t \eta(s)\, ds\, dt +$$
$$+ \int_\alpha^\beta d\nu(t)\, \eta(t) - \int_\alpha^\beta d\nu(t)\, b(t) \int_\alpha^t \eta(s)\, ds + \sum_{k=0}^{p} c_k \int_\alpha^{t_k} \eta(s)\, ds = 0.$$

Так как $\gamma \in \mathbf{R}^n$ и $\eta(\cdot) \in C([\alpha, \beta], \mathbf{R}^n)$ можно выбирать произвольно и независимо друг от друга, то, во-первых,

$$\int_\alpha^\beta a(t)\, dt - \int_\alpha^\beta d\nu(t)\, b(t) + \sum_{k=0}^{p} c_k = 0 \tag{5}$$

(ибо можно положить $\eta(\cdot) = 0$, γ — любое), а во-вторых, для любого $\eta(\cdot)\in C([\alpha,\beta], \mathbf{R}^n)$ имеет место равенство

$$\int_\alpha^\beta a(t)\int_\alpha^t \eta(s)\,ds\,dt+\int_\alpha^\beta d\nu(t)\,\eta(t)-$$
$$-\int_\alpha^\beta d\nu(t)\,b(t)\int_\alpha^t \eta(s)\,ds+\sum_{k=0}^p c_k\int_\alpha^{t_k}\eta(s)\,ds=0. \quad (6)$$

Меняя в первом и третьем члене порядок интегрирования в соответствии с формулой Дирихле ((5) п. 2.1.9), обозначая буквой s переменную интегрирования во втором члене, полагая

$$\chi_{[\alpha,\,t]}(s)=\begin{cases}1, & \alpha\leqslant s\leqslant t,\\ 0, & t<s\leqslant\beta\end{cases} \quad (7)$$

и используя равенство

$$\int_\alpha^\beta f(t)\,d\mu(t)=\int_\alpha^\beta f(t)\,\mu'(t)\,dt, \quad (8)$$

справедливое для абсолютно непрерывных функций [КФ, стр. 359], преобразуем (6) к виду

$$\int_\alpha^\beta\left(\int_s^\beta a(t)\,dt\right)\eta(s)\,ds+\int_\alpha^\beta d\nu(s)\,\eta(s)-\int_\alpha^\beta\left(\int_s^\beta d\nu(t)\,b(t)\right)\eta(s)\,ds+$$
$$+\int_\alpha^\beta\sum_{k=0}^p c_k\chi_{[\alpha,t_k]}(s)\,\eta(s)\,ds=\int_\alpha^\beta d\Lambda(s)\,\eta(s)+\int_\alpha^\beta d\nu(s)\,\eta(s)=0, \quad (9)$$

где

$$\Lambda(s)=\int_\alpha^s\left\{\int_\sigma^\beta a(t)\,dt-\int_\sigma^\beta d\nu(t)\,b(t)+\sum_{k=0}^p c_k\chi_{[\alpha,\,t_k]}(\sigma)\right\}d\sigma.$$

Функция $\Lambda(\cdot)$ абсолютно непрерывна на $[\alpha,\beta]$ и $\Lambda(\alpha)=0$, следовательно, это каноническая функция, а так как $\nu(\cdot)$—также каноническая функция ограниченной вариации, то из (9) и свойства единственности в теореме Рисса (п. 2.1.9) $\Lambda(s)+\nu(s)\equiv 0$. Следовательно, $\nu(\cdot)$ абсолютно непрерывна, а ее производная $p(s)=$ $=\nu'(s)$ при $s\neq t_k$ равна (снова используем (8))

$$p(s)=\nu'(s)=-\Lambda'(s)=$$
$$=-\int_s^\beta a(t)\,dt+\int_s^\beta p(t)\,b(t)\,dt-\sum_{k=0}^p c_k\chi_{[\alpha,\,t_k]}(s).$$

Из этой формулы видно, в свою очередь, что $p(\cdot)$ непрерывна всюду, кроме точек t_k, где справедливо (4), и что имеет место (2). Равенство $p(\beta)=0$ очевидно, а $p(\alpha)=0$ эквивалентно (5), так что доказано и (3). ■

4.1.4. Вывод условий стационарности. Итак, мы показали, что если $\hat{\xi}$ доставляет локальный минимум в задаче (1)—(3) п. 4.1.1, то найдутся такие множители Лагранжа $\hat{\lambda}_0 \geqslant 0$, $\hat{\lambda} \in \mathbf{R}^{m*}$ и $\hat{y}^* \in C(\Delta, \mathbf{R}^n)^*$, что выполнены условия согласования знаков и дополняющей нежесткости для λ и условия стационарности

$$\text{а) } \hat{\bar{\mathscr{L}}}_{x(\,)}=0, \quad \text{б) } \hat{\bar{\mathscr{L}}}_{u(\,)}=0, \quad \text{в) } \hat{\bar{\mathscr{L}}}_{t_k}=0, \quad k=0,\ 1, \tag{1}$$

где $\bar{\mathscr{L}}$ определяется равенством (6) п. 4.1 2, а $\hat{\bar{\mathscr{L}}}_{x(\,)}$, $\hat{\bar{\mathscr{L}}}_{u(\,)}$ и $\hat{\bar{\mathscr{L}}}_{t_k}$ получаются из $\bar{\mathscr{L}}_{x(\,)}$, $\bar{\mathscr{L}}_{u(\,)}$, $\bar{\mathscr{L}}_{t_k}$ подстановкой $x(\cdot)=\hat{x}(\cdot)$, $u(\cdot)=\hat{u}(\cdot)$, $t_k=\hat{t}_k$, $k=0,\ 1$.

Переходим к расшифровке условий (1). При дифференцировании функции $\bar{\mathscr{L}}$ мы будем иметь в виду, что первый и третий члены формулы (6) п. 4.1.2 образуют функционал Больца, производная которого дается формулой (3) того же пункта, а второй член является суперпозицией оператора дифференциальной связи и линейного отображения $\eta(\cdot)\longrightarrow\int\limits_{\Delta} d\nu(t)\,\eta(t)$, и его производная получается поэтому интегрированием формулы (4) п. 4.1.2 по $d\nu(t)$.

А) Стационарность по $x(\cdot)$. Дифференцируя $\bar{\mathscr{L}}$ по $x(\cdot)$, получаем в соответствии с (1а), что при всех $h(\cdot)\in C^1(\Delta, \mathbf{R}^n)$ имеет место равенство

$$0=\hat{\bar{\mathscr{L}}}_{x(\,)}[h(\cdot)]=\int\limits_{\hat{t}_0}^{\hat{t}_1}\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_{ix}(t)\,h(t)\,dt+$$
$$+\int\limits_{\Delta} d\nu(t)\,\{\dot{h}(t)-\hat{\varphi}_x(t)\,h(t)\}+\hat{l}_{x_0}h(\hat{t}_0)+\hat{l}_{x_1}h(\hat{t}_1). \tag{2}$$

Рассмотрим сначала случай $\hat{t}_1>\hat{t}_0$. Согласно обобщенной лемме Дюбуа-Реймона из (2) следует, что функция $\hat{\nu}(\cdot)$ абсолютно непрерывна на Δ, а ее производная $p(\cdot)=$ $=\hat{\nu}'(\cdot)$ непрерывна на Δ всюду, кроме, быть может, точек $\hat{t}_0$ и $\hat{t}_1$, где она имеет пределы слева и справа, и выполняются соотношения (2)—(4) п. 4.1.3, которые применительно к нашей ситуации означают, что

$$dp/dt = -p(t)\,\hat{\varphi}_x(t) + \chi_{[\hat{t}_0,\,\hat{t}_1]}(t)\sum_{i=0}^{m}\hat{\lambda}_i \hat{f}_{ix}(t) \tag{3}$$

при $t \neq \hat{t}_0,\ \hat{t}_1,\ p(\alpha) = p(\beta) = 0$, где α и β — концы Δ, и что

$$p(\hat{t}_0 + 0) - p(\hat{t}_0 - 0) = \hat{l}_{x_0}, \quad p(\hat{t}_1 + 0) - p(\hat{t}_1 - 0) = \hat{l}_{x_1} \tag{4}$$

(здесь $\chi_{[\hat{t}_0,\,\hat{t}_1]}$ — характеристическая функция, равная 1 на $[\hat{t}_0,\ \hat{t}_1]$ и 0 вне этого отрезка). Из уравнения (3) видно, что $p(\cdot)$ удовлетворяет однородному линейному дифференциальному уравнению на полуинтервалах $[\alpha,\ \hat{t}_0)$ и $(\hat{t}_1,\ \beta]$. Поскольку она обращается в нуль в концах α и β этих полуинтервалов, по теореме единственности п. 2.5 3 $p(t) \equiv 0$ на $\Delta \setminus [\hat{t}_0,\ \hat{t}_1]$ и $p(\hat{t}_0 - 0) = p(\hat{t}_1 + 0) = 0$.

Обозначим теперь через $\hat{p}(\cdot)$ решение уравнения (7а) п. 4.1.1, совпадающее на интервале $(\hat{t}_0,\ \hat{t}_1)$ с $p(\cdot)$ (в этом интервале уравнения (7а) п. 4.1.1 и (3) тождественны). Поскольку (7а) линейно, решение $\hat{p}(\cdot)$ определено на всем Δ и принадлежит $C^1(\Delta,\ \mathbf{R}^{n*})$ (п. 2.5 4). Для этого решения $\hat{p}(\hat{t}_0) = p(\hat{t}_0 + 0)$ и $\hat{p}(\hat{t}_1) = p(\hat{t}_1 - 0)$, а так как $p(\hat{t}_0 - 0) = 0$ и $p(\hat{t}_1 + 0) = 0$, то (4) превращается в условие (8а) п. 4.1.1. Таким образом, для $\hat{p}(\cdot)$ условия стационарности (7) и (8) (или (7а) и (8а)) из п. 4.1.1 справедливы.

При $\hat{t}_1 < \hat{t}_0$ мы вместо (3) получим уравнение

$$dp/dt = -p(t)\,\hat{\varphi}_x(t) - \chi_{[\hat{t}_1,\,\hat{t}_0]}(t)\sum_{i=0}^{m}\hat{\lambda}_i \hat{f}_{ix}(t),$$

решение $p(\cdot)$ которого равно нулю на $[\alpha,\ \hat{t}_1)$ и $(\hat{t}_0,\ \beta]$. В этом случае надо обозначить через $\hat{p}(\cdot)$ то решение уравнения (7а) п. 4.1.1, которое совпадает с $(-1)\,p(\cdot)$ на $(\hat{t}_1,\ \hat{t}_0)$, после чего условия (7), (8) п. 4.1.1 будут выполнены.

Наконец, при $\hat{t}_1 = \hat{t}_0$ мы снова можем воспользоваться обобщенной леммой Дюбуа-Реймона, но теперь уже $\hat{v}'(t) = p(t) \equiv 0$ всюду на Δ, кроме, быть может, точки $\hat{t}_0 = \hat{t}_1$, где должно еще выполняться условие $p(\hat{t}_1 + 0) - p(\hat{t}_1 - 0) = \hat{l}_{x_0} + \hat{l}_{x_1} = 0$. Отсюда $\hat{l}_{x_0} + \hat{l}_{x_1} = 0$. Беря в качестве $\hat{p}(\cdot)$ то решение уравнения (7а) п. **4.1.1**, для которого $\hat{p}(\hat{t}_0)=\hat{l}_{x_0}$, мы снова убеждаемся в том, что условия (7), (8) п. 4.1.1 выполнены.

Б) Стационарность по $u(\cdot)$. Дифференцируя $\tilde{\mathscr{L}}$ по $u(\cdot)$ и учитывая, что $\nu(\cdot)$ абсолютно непрерывна, а ее производная $p(\cdot)=\nu'(\cdot)$ отлична от нуля лишь в интервале $(\hat{t}_0, \hat{t}_1)$, а в этом интервале совпадает с $\hat{p}(\cdot)$, имеем

$$0=\hat{\tilde{\mathscr{L}}}_u[v(\cdot)]=\int\limits_{\hat{t}_0}^{\hat{t}_1}\left\{\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_{iu}(t)-\hat{p}(t)\hat{\varphi}_u(t)\right\}v(t)\,dt.$$

Так как это равенство имеет место при всех $v(\cdot)\in C(\Delta, \mathbf{R}^n)$, то, снова сославшись на единственность в теореме Рисса (п. 2.1.9), мы получаем

$$\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_{iu}(t)-\hat{p}(t)\hat{\varphi}_u(t)=0,$$

что совпадает с уравнением (9а), а следовательно, и с условием (9) п.4.1.1.

В) Стационарность по t_k. Поскольку

$$\tilde{\mathscr{L}}(\hat{x}(\cdot), \hat{u}(\cdot), t_0, t_1; \ldots)-\mathscr{L}(\hat{x}(\cdot), \hat{u}(\cdot), t_0, t_1; \ldots)=$$
$$=\int\limits_{\Delta}d\nu(t)\{\hat{x}(t)-\varphi(t, \hat{x}(t), \hat{u}(t))\}-$$
$$-\int\limits_{t_0}^{t_1}\hat{p}(t)\{\dot{\hat{x}}(t)-\varphi(t, \hat{x}(t), \hat{u}(t))\}\,dt\equiv 0,$$

то производные этих двух функций по t_k в точке $\hat{\xi}=(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$ совпадают, а так как имеет место равенство $\hat{\tilde{\mathscr{L}}}_{t_k}=0$, то выполняется и условие (10) п. 4.1.1.

4.1.5. Задача со старшими производными. Уравнение Эйлера — Пуассона. Снова зафиксируем замкнутый отрезок $\Delta\subset\mathbf{R}$ и будем рассматривать пространство

$$\Xi_s=C^s(\Delta, \mathbf{R}^n)\times\mathbf{R}\times\mathbf{R},$$

элементами которого являются тройки $\xi=(x(\cdot), t_0, t_1)$, где функция $x(\cdot):\Delta\to\mathbf{R}^n$ непрерывно дифференцируема до порядка s включительно; $t_0, t_1\in\operatorname{int}\Delta$. Говоря о *задаче со старшими производными* (в классическом вариационном исчислении), будем иметь в виду экстремальную задачу

$$\mathscr{P}_0(\xi) \to \mathrm{extr}; \quad \mathscr{P}_i(\xi) \leqslant 0, \quad i=1, \ldots, m, \tag{1}$$

где

$$\mathscr{P}_i(\xi) = \int_{t_0}^{t_1} f_i(t, x(t), \ldots, x^{(s)}(t))\,dt +$$
$$+ \psi_i(t_0, x(t_0), \ldots, x^{(s-1)}(t_0), t_1, x(t_1), \ldots, x^{(s-1)}(t_1)) \tag{2}$$

в пространстве Ξ_s. При этом в (1), (2) $f_i: V \to \mathbf{R}$, $\psi_i: W \to \mathbf{R}$, $i=0, 1, \ldots, m$, V и W — открытые множества в пространствах $\mathbf{R} \times (\mathbf{R}^n)^{s+1}$ и $\mathbf{R} \times (\mathbf{R}^n)^s \times \mathbf{R} \times (\mathbf{R}^n)^s$ соответственно, f_i и ψ_i по крайней мере непрерывны. Тройка $\xi = (x(\cdot), t_0, t_1) \in \Xi_s$ является *допустимой* в задаче (1), (2), если $(t, x(t), \ldots, x^{(s)}(t)) \in V$, $t \in \Delta$ и

$$(t_0, x(t_0), \ldots, x^{(s-1)}(t_0), t_1, x(t_1), \ldots, x^{(s-1)}(t_1)) \in W.$$

Допустимую тройку $\hat{\xi} = (\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)$ мы называем *(локальным) решением* задачи (1), (2), если она доставляет функционалу $\mathscr{P}_0$ локальный экстремум в пространстве Ξ_s, т. е. если существует такое $\varepsilon > 0$, что для всех допустимых $\xi = (x(\cdot), t_0, t_1)$, для которых

$$\|\xi - \hat{\xi}\|_{\Xi_s} < \varepsilon \Leftrightarrow |t_k - \hat{t}_k| < \varepsilon, \quad k=0, 1;$$
$$\|x - \hat{x}(\cdot)\|_{C^s(\Delta, R^n)} < \varepsilon,$$

выполняется одно из неравенств: $\mathscr{P}_0(\xi) \geqslant \mathscr{P}_0(\hat{\xi})$ в случае минимума и $\mathscr{P}_0(\xi) \leqslant \mathscr{P}_0(\hat{\xi})$ в случае максимума.

Покажем, что задача со старшими производными (1), (2) сводится к задаче Лагранжа. С этой целью обозначим

$$\boldsymbol{x} = (x_1, \ldots, x_s) \in (\mathbf{R}^n)^s,$$
$$\boldsymbol{x}(\cdot) = (x(\cdot), \dot{x}(\cdot), \ldots, x^{(s-1)}(\cdot)),$$
$$u(\cdot) = x^{(s)}(\cdot), \quad \boldsymbol{p}(\cdot) = (p_1(\cdot), \ldots, p_s(\cdot)).$$

Тогда задача (1), (2) приобретает следующий стандартный вид задачи п. 4.1.1:

$$\int_{t_0}^{t_1} f_0(t, \boldsymbol{x}(t), u(t))\,dt + \psi_0(t_0, \boldsymbol{x}(t_0), t_1, \boldsymbol{x}(t_1)) \to \mathrm{extr}, \tag{1'}$$

$$\int_{t_0}^{t_1} f_i(t, \boldsymbol{x}(t), u(t))\,dt + \psi_i(t_0, \boldsymbol{x}(t_0), t_1, \boldsymbol{x}(t_1)) \leqslant 0,$$
$$i=1, \ldots, m, \tag{2'}$$

$$\dot{x}_j = x_{j+1}, \quad j=1, \ldots, s-1, \quad x_s = u \tag{3}$$

с функцией Лагранжа

$$\mathscr{L}(\boldsymbol{x}(\cdot),\ u(\cdot),\ t_0,\ t_1;\ \boldsymbol{p}(\cdot),\ \lambda,\ \lambda_0)=\int_{t_0}^{t_1} L\,dt+l, \qquad (4)$$

где

$$L(t,\ \boldsymbol{x},\ \dot{\boldsymbol{x}},\ u)=\sum_{i=0}^{m}\lambda_i f_i(t,\ \boldsymbol{x},\ u)+\sum_{j=1}^{s-1}p_j(t)(\dot{x}_j-x_{j+1})+p_s(t)(\dot{x}_s-u) \qquad (5)$$

и

$$l(t_0,\ x_0,\ \dot{x}_0,\ \ldots,\ x_0^{(s-1)},\ t_1,\ x_1,\ \dot{x}_1,\ \ldots,\ x_1^{(s-1)})=$$
$$=\sum_{i=0}^{m}\lambda_i\psi_i(t_0,\ x_0,\ \dot{x}_0,\ \ldots,\ x_0^{(s-1)},\ t_1,\ x_1,\ \dot{x}_1,\ \ldots,\ x_1^{(s-1)}). \qquad (6)$$

Перефразируя на этот случай теорему п. 4.1.2, получаем необходимое условие экстремума.

Теорема. *Пусть функции* f_i: $V\to\mathbf{R}$, $i=0,\ 1,\ \ldots,\ m$, *и их частные производные по* $x,\ \ldots,\ x^{(s)}$ *непрерывны в открытом множестве* $U\subset\mathbf{R}\times(\mathbf{R}^n)^{s+1}$, *а функции* ψ_i: $W\to\mathbf{R}$, $i=0,\ 1,\ \ldots,\ m$, *непрерывно дифференцируемы в открытом множестве* $W\subset\mathbf{R}\times(\mathbf{R}^n)^s\times\times\mathbf{R}\times(\mathbf{R}^n)^s$, *и пусть* $\hat{x}(\cdot)\in C^s(\Delta,\ \mathbf{R}^n)$, $\hat{t}_0,\ \hat{t}_1\in\operatorname{int}\Delta$ *таковы, что*

$$(t,\ \hat{x}(t),\ \ldots,\ \hat{x}^{(s)}(t))\in V,\qquad t\in\Delta;$$
$$(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \ldots,\ \hat{x}^{(s-1)}(\hat{t}_0),\ \hat{t}_1,\ \hat{x}(\hat{t}_1),\ \ldots,\ \hat{x}^{(s-1)}(\hat{t}_1))\in W.$$

Если $\hat{\xi}=(\hat{x}(\cdot),\ \hat{t}_0,\ \hat{t}_1)$ *является локальным решением задачи* (1), (2), *то найдутся множители Лагранжа:* $\hat{\lambda}_0\geqslant 0$ *в задаче на минимум и* $\leqslant 0$ *в задаче на максимум,* $\hat{p}(\cdot)=(\hat{p}_1(\cdot),\ \ldots,\ \hat{p}_s(\cdot))\in C^1(\Delta,\ (\mathbf{R}^n)^{s*})$, $\hat{\lambda}=(\hat{\lambda}_1,\ \ldots,\ \hat{\lambda}_n)$, *не равные нулю одновременно и такие, что:*

а) *выполнены условия стационарности функции Лагранжа:*

по $x(\cdot)$ $(\hat{\mathscr{L}}_{x(\cdot)}=0,\ \hat{\mathscr{L}}_{u(\cdot)}=0)$

$$\left.\begin{aligned}&\dot{\hat{p}}_1(t)=f_x(t,\ \hat{x}(t),\ \ldots,\ \hat{x}^{(s)}(t)),\\&\dot{\hat{p}}_j(t)=f_{x^{(j-1)}}(t,\ \hat{x}(t),\ \ldots,\ \hat{x}^{(s)}(t))-\hat{p}_{j-1}(t),\\&\qquad\qquad j=2,\ \ldots,\ s,\\&0=f_{x^{(s)}}(t,\ \hat{x}(t),\ \ldots,\ \hat{x}^{(s)}(t))-\hat{p}_s(t),\end{aligned}\right\}\qquad (7)$$

$$\hat{p}_j(\hat{t}_k)=(-1)^k\hat{l}_{x_k^{(j-1)}},\qquad k=0,\ 1; \qquad (8)$$

по t_k $(\hat{\mathscr{L}}_{t_k}=0)$

$$\hat{H}(\hat{t}_k)=(-1)^{k+1}\hat{l}_{t_k},\quad k=0,\ 1; \tag{9}$$

б) *выполнено условие согласования знаков*

$$\hat{\lambda}_i \gtrless 0,\quad i=1,\ \ldots,\ m; \tag{10}$$

в) *выполнены условия дополняющей нежесткости*

$$\hat{\lambda}_i \mathscr{P}_i(\hat{\xi})=0,\quad i=1,\ \ldots,\ m. \tag{11}$$

Здесь

$$f(t,\ x,\ \dot{x},\ \ldots,\ x^{(s)})=\sum_{i=0}^{m}\hat{\lambda}_i f_i(t,\ x,\ \ldots,\ x^{(s)}), \tag{12}$$

$$\hat{H}(t)=\sum_{j=1}^{s}\hat{p}_j(t)\,\hat{x}^{(j)}(t)-f(t,\ \hat{x}(t),\ \ldots,\ \hat{x}^{(s)}(t)) \tag{13}$$

и, как обычно,

$$\hat{l}_{x_k^{(j)}}=l_{x_k^{(j)}}(\hat{t}_0,\ \hat{x}(\hat{t}_0),\ \ldots,\ \hat{x}^{(s-1)}(\hat{t}_0),\ \hat{t}_1,\ \hat{x}(\hat{t}_1),\ \ldots,\ \hat{x}^{(s-1)}(\hat{t}_1))$$

и т. п., а неравенства (10) означают, что $\hat{\lambda}_i\geqslant 0$, если в задаче (1) $\mathscr{P}_i(\xi)\leqslant 0$, $\hat{\lambda}_i\leqslant 0$, если в (1) $\mathscr{P}_i(\xi)\geqslant 0$ и $\hat{\lambda}_i$ может иметь любой знак, если в (1) $\mathscr{P}_i(\xi)=0$.

Подробные выкладки, приводящие условия (7)—(12) п. 4.1.1 к аналогичным условиям этой теоремы, предоставляются читателям в качестве упражнения.

Для частного случая задачи со старшими производными, в котором терминальные члены в (1′) и (2′) суть константы и концы закреплены, утверждение теоремы можно представить в более простой форме. Условия трансверсальности (8) и (9) здесь опускаются, а систему (7) можно записать в виде одного уравнения

$$\hat{f}_x(t)-\frac{d}{dt}\left(\hat{f}_{\dot{x}}(t)-\frac{d}{dt}\left(\hat{f}_{\ddot{x}}-\ldots-\frac{d}{dt}\hat{f}_{x^{(s)}}(t)\right)\right)=0. \tag{14}$$

Наконец, если предположить существование нужных производных, уравнение (13) можно записать в виде уравнения Эйлера—Пуассона

$$\hat{f}_x(t)-\frac{d}{dt}\hat{f}_{\dot{x}}+\frac{d^2}{dt^2}\hat{f}_{\ddot{x}}(t)-\ldots-(-1)^s\frac{d^s}{dt^s}\hat{f}_{x^{(s)}}(t)=0. \tag{15}$$

Общее решение этого уравнения порядка $2s$ содержит $2s$ произвольных постоянных и зависит еще от чисел $\lambda_0,\ \ldots,\ \lambda_m$, одно из которых можно выбирать произвольно. Всего в нашем распоряжении оказывается

$2s+m$ неизвестных. Для их определения мы используем $2s$ краевых условий (для закрепленных концов) и еще m условий получаем из условий дополняющей нежесткости и неравенств (1) и (10) (проверьте, что всего из них можно извлечь ровно m равенств). Аналогично можно проверить, что и в общем случае число уравнений, которые дает теорема, совпадает с числом неизвестных.

§ 4.2. Принцип максимума Понтрягина

4.2.1. Постановка задачи оптимального управления. Экстремальная задача, которой посвящен этот параграф, по форме почти совпадает с задачей Лагранжа (1)—(3) п. 4.1.1. Здесь мы встретимся с тем же функционалом, той же дифференциальной связью и с теми же ограничениями типа неравенств, что и раньше. Отличие новой задачи состоит в том, что в ее формулировке явно выделено ограничение на управление вида $u(t) \in \mathfrak{U}$, где $\mathfrak{U}$ — некоторое топологическое пространство.

На первый взгляд это не прибавляет ничего нового и даже выглядит менее общим. Действительно, в задаче Лагранжа мы имели дело с условием $(t, x(t), u(t)) \in V$, которому должны были удовлетворять допустимые процессы. В частном случае, беря $V = G \times \mathfrak{U}$, мы расщепляем это условие на два: $(t, x(t)) \in G$ и $u(t) \in \mathfrak{U}$.

Однако разделение переменных на фазовые и управление, и выделение ограничений на управление оказалось весьма знаменательным. Вместе с ним выделилась и новая ветвь в теории экстремальных задач, быстро завоевавшая популярность среди прикладников своими возможностями применительно к практическим задачам и позволившая по-новому взглянуть и на классическую теорию.

Чем же все это было вызвано? В первую очередь необходимостью исследования задач технического содержания. Разделение фазовых переменных и управления и их связь при помощи дифференциального уравнения — обычная модель для процесса, развивающегося по законам природы (дифференциальное уравнение!), но испытывающего воздействие человека, управляющего этим процессом в соответствии со своими целями и стремящегося сделать его в каком-то смысле оптимальным. Теперь ясно, что ограничения на управления вида $u(t) \in \mathfrak{U}$ связаны с ограниченными возможностями воздействовать на процесс (скажем, с ограниченностью диапазона поворота рулей управляемого аппарата).

Если принять за $\mathfrak{U}$ открытое множество в $\mathbf{R}^r$, то ничего нового мы действительно не получим. Однако в технических и иных прикладных задачах множество $\mathfrak{U}$ не обязано быть открытым. Зачастую управление может быть даже просто дискретным (включено—выключено). Естественно возникают ограничения, при которых $\mathfrak{U}$ оказывается замкнутым множеством: насколько это существенно, мы видели, решая задачу Ньютона (п. 1.6.2). Уже в простейших задачах оптимального управления встречаются ограничения вида $|u(t)| \leqslant 1$ (см. задачу о быстродействии в п. 1.6.3). В этом последнем случае множеством допустимых управлений оказывается замкнутый шар в сложно устроенном пространстве $L_\infty(\Delta)$, и в типичных случаях оптимальное управление лежит на границе этого шара (из сказанного в п. 1.6.3 видно, что при ограничении $|u(t)| < 1$ задача не имела бы решения: минимальное время достигается при управлении, где почти всюду $|u(t)| = 1$), а эта последняя чрезвычайно негладкая и «многогранная», а точнее, «бесконечногранная». Все это затрудняет применение обычных методов дифференциального исчисления, и условия гладкости по управлению оказываются довольно часто неестественными. В связи с этим мы не будем больше требовать существования производных f_{iu}, φ_u и т. п. и даже само множество $\mathfrak{U}$ возможных значений управления будем считать произвольным топологическим пространством, не обладающим, вообще говоря, линейной структурой. Отсутствие производных по u—второе отличие новой задачи от старой. Желая оттенить это обстоятельство, функционалы, входящие в формулировку задачи, мы обозначаем новой буквой, хотя по виду они совпадают с функционалами Больца.

Наконец, последнее отличие—это отказ от непрерывности управления. В основном это связано с тем, что в классе непрерывных управлений задача часто не имеет решения (например, в большинстве задач, где $\mathfrak{U}$ дискретно или в задаче о быстродействии п. 1.6.3). Следует, впрочем, отметить, что переход к рассмотрению «ломаных экстремалей» (что соответствует кусочно-непрерывным управлениям) делается и в классическом вариационном исчислении (см. п. 1.4.3). Сама же возможность свободного выбора управления внутри множества $\mathfrak{U}$ (а ею мы уже воспользовались, «выпуская» игольчатые вариации при выводе условия Вейерштрасса в п. 1.4.4 и при доказательстве принципа максимума в простейшей ситуации (п. 1.5.4)), порождает скрытую выпуклость по управлениям, с чем связана форма основного условия — принципа максимума (см. (10), (11) в п. 4.2.2), напоминающая соответствующее условие в задачах выпуклого программирования. Здесь эта связь не будет вскрыта в полном объеме. Мы сделаем это в следующем параграфе, правда в частном варианте общей задачи.

Итак, рассмотрим экстремальную задачу

$$\mathcal{J}_0(x(\cdot),\ u(\cdot),\ t_0,\ t_1)=$$
$$=\int_{t_0}^{t_1} f_0(t,\ x(t),\ u(t))\,dt+\psi_0(t_0,\ x(t_0),\ t_1,\ x(t_1))\to\inf, \quad (1)$$

$$\dot{x}=\varphi(t,\ x(t),\ u(t)),\quad u(t)\in\mathfrak{U}, \quad (2)$$

$$\mathcal{J}_i(x(\cdot),\ u(\cdot),\ t_0,\ t_1)=\int_{t_0}^{t_1} f_i(t,\ x(t),\ u(t))\,dt+$$
$$+\psi_i(t_0,\ x(t_0),\ t_1,\ x(t_1))\lesseqgtr 0,\quad i=1,\ 2,\ \ldots,\ m. \quad (3)$$

Здесь f_i: $G\times\mathfrak{U}\to\mathbf{R}$, $\psi_i:W\to\mathbf{R}$, $i=0,\ 1,\ \ldots,\ m$, φ: $G\times\mathfrak{U}\to\mathbf{R}^n$, где G и W — открытые множества в пространствах $\mathbf{R}\times\mathbf{R}^n$ и $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}\times\mathbf{R}^n$ соответственно, $\mathfrak{U}$ — произвольное топологическое пространство. Знак $\lesseqgtr$ имеет тот же смысл, что в §§ 3.2 и 4.1. Все функции f_i, ψ_j, φ предполагается по крайней мере непрерывными.

Четверку $(x(\cdot),\ u(\cdot),\ t_0,\ t_1)$ будем называть *управляемым процессом* в задаче (1)—(3), если:

а) *управление* $u(\cdot)$: $[t_0,\ t_1]\to\mathfrak{U}$ — кусочно-непрерывная функция. Ее значения в точках разрыва существенной роли не играют; для определенности будем считать, что $u(\cdot)$ непрерывна справа для $t_0\leqslant t<t_1$ и слева в точке t_1.

б) *фазовая траектория* $x(\cdot)$: $[t_0,\ t_1]\to\mathbf{R}^n$ непрерывна и ее график лежит в G

$$\Gamma=\{(t,\ x(t))\,|\,t_0\leqslant t\leqslant t_1\}\subset G.$$

в) для всех $t\in[t_0,\ t_1]$, кроме, быть может, точек разрыва управления $u(\cdot)$, функция $x(\cdot)$ удовлетворяет дифференциальному уравнению

$$\dot{x}(t)=\varphi(t,\ x(t),\ u(t))$$

(в этом случае мы говорим, что $x(\cdot)$ соответствует управлению $u(\cdot)$; легко видеть, что в точках разрыва управления $x(\cdot)$ имеет производные слева и справа).

Управляемый процесс называется *допустимым*, если удовлетворяются условия (3).

Допустимый управляемый процесс $(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$ называется (локально) *оптимальным*, если найдется такое $\varepsilon > 0$, что для всякого допустимого управляемого процесса $(x(\cdot), u(\cdot), t_0, t_1)$ такого, что

$$|t_k - \hat{t}_k| < \varepsilon, \ k = 0, 1 \quad \text{и} \quad |x(t) - \hat{x}(t)| < \varepsilon$$
$$\text{для всех } t \in [t_0, t_1] \cap [\hat{t}_0, \hat{t}_1] \tag{4}$$

выполняется неравенство

$$\mathcal{J}_0(x(\cdot), u(\cdot), t_0, t_1) \geqslant \mathcal{J}_0(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1). \tag{5}$$

Заметим, что и здесь произошло изменение по сравнению с задачей Лагранжа: мы не требуем больше близости производных $\dot{x}(t)$ и $\dot{\hat{x}}(t)$. В классическом вариационном исчислении это соответствует переходу от слабого экстремума к сильному (см. п. 1.4.3).

Поучительно сравнить задачу Лагранжа и задачу оптимального управления в том случае, когда $\mathfrak{U} = \mathbf{R}^r$ и в п. 4.1.1 $V = G \times \mathbf{R}^r$. При этом предположении:

Каждый допустимый управляемый процесс $(x(\cdot), u(\cdot), t_0, t_1)$ задачи Лагранжа является таковым и для задачи оптимального управления (точнее, превращается в таковой при ограничении $u(\cdot)$ на $[t_0, t_1]$).

Это означает, что произошло расширение задачи (ср. п. 1.4.3). Следовательно, значение задачи оптимального управления не больше значения задачи Лагранжа ($\inf \mathcal{J}_0 \leqslant \inf \mathcal{B}_0$). Впрочем, при достаточно естественных предположениях относительно функций f_i, ψ_i и φ обе эти величины совпадают (в тривиальном случае, когда $\dot{x} = u$, мы доказали это в лемме «о скруглении углов» в п. 1.4.3).

Если $\inf \mathcal{J}_0 < \inf \mathcal{B}_0$, то, как легко видеть, решения обеих задач могут существовать или нет независимо друг от друга. Если же значения задач равны, то возможно одно из трех:

а) Задача Лагранжа имеет решение: $\inf \mathcal{B}_0 = \mathcal{B}_0(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$. Тогда та же четверка является и решением задачи оптимального управления, поскольку

$$\mathcal{J}_0(x(\cdot),\ u(\cdot),\ t_0,\ t_1) \geqslant \mathcal{B}_0(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1) = \\ = \inf \mathcal{B}_0 = \inf \mathcal{J}_0.$$

б) Задача Лагранжа не имеет решения, а задача оптимального управления имеет: $\inf \mathcal{J}_0$ достигается на кусочно-гладкой кривой $\hat{x}(\cdot)$. Такую ситуацию мы наблюдали в примере (6) п. 1.4.3 (при $\dot{x}=u$ переход от класса C^1 к классу KC^1 как раз и соответствует переходу к задаче оптимального управления).

в) Обе задачи не имеют решения: нижняя грань функционала не достигается и при расширении области его определения. Так обстоит дело в примере Больца (п. 1.4.3).

В предыдущих рассуждениях молчаливо предполагалось, что речь идет о минимизации по всему классу допустимых управляемых процессов. Однако обычно при рассмотрении экстремальных задач мы придерживаемся локальной точки зрения, что, кстати сказать, отражено и в наших определениях. В п. 4.1.1 допустимый управляемый процесс $(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1)$ был назван оптимальным (в слабом смысле), если он доставляет минимум функционалу $\mathcal{B}_0$ в некоторой C^1-окрестности по $x(\cdot)$ и C-окрестности по $u(\cdot)$. Теперь же определение оптимальности предполагает, что минимум $\mathcal{J}_0$ имеет место в более широкой окрестности, описываемой неравенствами (4). Фактически это C-окрестность по $x(\cdot)$, а ограничения на управления не накладываются вовсе. Как уже было сказано, в классической ситуации это был бы сильный минимум. Поэтому локальное решение задачи Лагранжа может не быть оптимальным процессом для задачи оптимального управления. Если же в оптимальном процессе $(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1)$ управление $\hat{u}(\cdot)$ непрерывно, то при наших обычных предположениях относительно функций, входящих в задачу, этот процесс будет и локальным решением задачи Лагранжа.

Ограничения (3) не являются самыми общими, которые приходится рассматривать в задачах оптимального управления. Мы, например, совсем оставляем в стороне так называемые *фазовые ограничения* вида

$$(t,\ x(t)) \in D, \quad t_0 \leqslant t \leqslant t_1, \tag{6}$$

где D — некоторое подмножество в $\mathbf{R}\times\mathbf{R}^n$. Конечно, если D открыто, то, заменив в определении управляемого процесса G на $G\cap D$ и уменьшив, если нужно, ε в (4), мы автоматически учтем ограничения (6). Поэтому основной интерес представляет случай, когда D не является открытым множеством и оптимальная фазовая траектория $x(t)$ выходит для некоторых t на границу D (см. рис. 37; минимизируется длина кривой, заштрихованная часть плоскости не принадлежит D). Это приводит к существенно новым эффектам, в частности, к необходимости

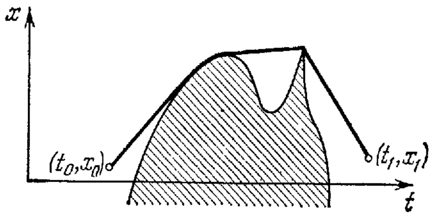


Рис. 37.

ввести в уравнение Эйлера—Лагранжа обобщенную функцию — производную не абсолютно непрерывной меры (или заменить это дифференциальное уравнение интегральным с интегралом Стилтьеса по этой мере).

4.2.2. Формулировка принципа максимума. Принцип Лагранжа в задаче оптимального управления. Рассмотрим снова задачу

$$\mathcal{J}_0(x(\cdot),\ u(\cdot),\ t_0,\ t_1)=\int_{t_0}^{t_1} f_0(t,\ x(t),\ u(t))\,dt+$$
$$+\psi_0(t_0,\ x(t_0),\ t_1,\ x(t_1))\to\inf, \quad (1)$$

$$\dot{x}=\varphi(t,\ x(t),\ u(t)),\quad u(t)\in\mathfrak{U}, \quad (2)$$

$$\mathcal{J}_i(x(\cdot),\ u(\cdot),\ t_0,\ t_1)=\int_{t_0}^{t_1} f_i(t,\ x(t),\ u(t))\,dt+$$
$$+\psi_i(t_0,\ x(t_0),\ t_1,\ x(t_1))\leqslant 0,\quad i=1,\ 2,\ \ldots,\ m. \quad (3)$$

Как и в п. 4.1.1, функцией Лагранжа этой задачи называется функция

$$\mathcal{L}(x(\cdot),\ u(\cdot),\ t_0,\ t_1;\ p(\cdot),\ \lambda,\ \lambda_0)=\int_{t_0}^{t_1} L\,dt+l, \quad (4)$$

где

$$\lambda_0 \in \mathbf{R},\ \lambda = (\lambda_1, \ \ldots, \ \lambda_m) \in \mathbf{R}^{m*},$$

$$L(t, \ x, \ \dot{x}, \ u) = \sum_{i=0}^{m} \lambda_i f_i(t, x, u) + p(t)(\dot{x} - \varphi(t, x, u)), \quad (5)$$

$$l(t_0, x_0, t_1, x_1) = \sum_{i=0}^{m} \lambda_i \psi_i(t_0, \ x_0, \ t_1, \ x_1). \quad (6)$$

Функция $p(\cdot)$: $[t_0, \ t_1] \to \mathbf{R}^{n*}$ предполагается непрерывной. Там, где это будет нужно, мы будем считать, что $p(\cdot)$ определена на более широком промежутке, чем $[t_0, \ t_1]$, продолжая ее за пределы этого отрезка произвольно, но с сохранением непрерывности. Функция $p(\cdot)$ и числа λ_i, λ_0 называются *множителями Лагранжа* задачи (1)—(3), функция (5)—*лагранжианом*, (6)—*терминантом*. Наконец, функцию

$$H(t, \ x, \ u, \ p) = L_{\dot{x}} \dot{x} - L = p\varphi(t, \ x, \ u) - \sum_{i=0}^{m} \lambda_i f_i(t, x, u) \quad (7)$$

мы будем называть *функцией Понтрягина*. Как и обычно, будут использоваться обозначения

$$\hat{L}_x(t) = L_x(t, \ \hat{x}(t), \ \dot{\hat{x}}(t), \ \hat{u}(t)), \ \hat{\varphi}(t) = \varphi(t, \ \hat{x}(t), \ \hat{u}(t)),$$
$$\hat{l}_{x_k} = l_{x_k}(\hat{t}_0, \ \hat{x}(\hat{t}_0), \ \hat{t}_1, \ \hat{x}(\hat{t}_1))$$

и т. д.

Теорема (принцип максимума Понтрягина). *Пусть G—открытое множество в пространстве $\mathbf{R} \times \mathbf{R}^n$, W—открытое множество в пространстве $\mathbf{R} \times \mathbf{R}^n \times \times \mathbf{R} \times \mathbf{R}^n$, $\mathfrak{U}$—произвольное топологическое пространство; функции f_i: $G \times \mathfrak{U} \to \mathbf{R}$, $i = 0, 1, \ldots, m$, и φ: $G \times \mathfrak{U} \to \mathbf{R}^n$ и их частные производные по x непрерывны в $G \times \mathfrak{U}$, а функции ψ_i: $W \to \mathbf{R}$, $i = 1, \ldots, m$, непрерывно дифференцируемы в W.*

Если $(\hat{x}(\cdot), \ \hat{u}(\cdot), \ \hat{t}_0, \ \hat{t}_1)$—оптимальный процесс для задачи (1)—(3), *то найдутся множители Лагранжа*

$$\hat{\lambda}_0 \geqslant 0, \ \hat{p}(\cdot), \quad \hat{\lambda} = (\hat{\lambda}_1, \ \ldots, \ \hat{\lambda}_n),$$

не равные одновременно нулю и такие, что:

а) *выполнены условия стационарности и принцип минимума для функции Лагранжа:*

по $x(\cdot)$ *условие стационарности* ($\mathscr{L}_{x(\cdot)}=0$)

$$-\frac{d}{dt}\hat{L}_{\dot{x}}(t)+\hat{L}_x(t)=0, \qquad (8)$$

$$\hat{L}_{\dot{x}}(\hat{t}_k)=(-1)^k\hat{l}_{x_k}, \quad k=0,\ 1; \qquad (9)$$

по $u(\cdot)$ *принцип минимума в лагранжевой форме*

$$\hat{L}(t)\equiv L(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),\ \hat{u}(t))\equiv$$

$$\equiv\min_{v\in\mathfrak{U}} L(t,\ \hat{x}(t),\ \dot{\hat{x}}(t),v), \qquad (10)$$

или в гамильтоновой (понтрягинской) форме в виде принципа максимума

$$\hat{H}(t)\equiv H(t,\ \hat{x}(t),\ \hat{u}(t),\ \hat{p}(t))\equiv$$

$$\equiv\max_{v\in\mathfrak{U}} H(t,\ \hat{x}(t),\ v,\ \hat{p}(t)); \qquad (11)$$

при этом функция $\hat{H}(t)$ *непрерывна на отрезке* $[\hat{t}_0,\ \hat{t}_1]$;
по t_k *условия стационарности*

$$\hat{\mathscr{L}}_{t_k}=0, \quad k=0,\ 1; \qquad (12)$$

б) *выполнено условие согласования знаков*

$$\lambda_i\gtreqless 0; \qquad (13)$$

в) *выполнены условия дополняющей нежесткости*

$$\hat{\lambda}_i\mathscr{I}_i(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1)=0, \quad i=1,\ 2,\ \ldots,\ m \qquad (14)$$

(как и в предыдущем параграфе неравенства (13) означают, что $\hat{\lambda}_i\geqslant 0$, если в условии (3) $\mathscr{I}_i\leqslant 0$; $\hat{\lambda}_i\leqslant 0$, если в (3) $\mathscr{I}_i\geqslant 0$ и $\hat{\lambda}_i$ может иметь любой знак, если $\mathscr{I}_i=0$).

Утверждение о существовании множителей Лагранжа, удовлетворяющих совокупности условий а)—в), кратко называется нами *принципом Лагранжа для задачи оптимального управления* (1)—(3) или *принципом максимума Понтрягина*. Как и раньше, это утверждение находится в полном соответствии с общим принципом Лагранжа из п. 3.1.5. Поскольку функция Лагранжа $\mathscr{L}$ является функцией трех аргументов: $x(\cdot)$, $u(\cdot)$ и $(t_0,\ t_1)$, надлежит рассмотреть три задачи:

$$(\alpha)\qquad \mathscr{L}(x(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1;\ \hat{p}(\cdot),\ \hat{\lambda},\ \hat{\lambda}_0)\to\inf,$$

$$(\beta)\qquad \mathscr{L}(\hat{x}(\cdot),\ u(\cdot),\ \hat{t}_0,\ \hat{t}_1;\ \hat{p}(\cdot),\ \hat{\lambda},\ \hat{\lambda}_0)\to\inf,$$

$$(\gamma)\qquad \mathscr{L}(\hat{x}(\cdot),\ \hat{u}(\cdot),\ t_0,\ t_1;\ \hat{p}(\cdot),\ \hat{\lambda},\ \hat{\lambda}_0)\to\inf.$$

Задачи (α) и (γ) здесь те же самые, что и в п. 4.1.1. Поэтому и соответствующие условия стационарности по $x(\cdot)$ и по t_k должны выглядеть здесь так же, как и в задаче Лагранжа и это действительно так. Что же касается задачи (β), то это элементарная задача оптимального управления из п. 3.1.4, и принцип минимума (10) вполне соответствует условию (1) этого пункта. Таким образом, приведенная выше формулировка теоремы действительно согласуется с общим принципом Лагранжа.

Обычно условия стационарности (8), (9) и (12) записываются в другой форме: (8)—как уравнение Эйлера—Лагранжа (называемое также сопряженным уравнением)

$$d\hat{p}(t)/dt=-\hat{p}(t)\hat{\varphi}_x(t)+\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_{ix}(t);\qquad(8а)$$

(9) и (12)—как *условия трансверсальности*

$$\hat{p}(\hat{t}_1)+\hat{l}_{x_1}=0,\quad -\hat{p}(\hat{t}_0)+\hat{l}_{x_0}=0\qquad(9а)$$

и

$$-\hat{H}(\hat{t}_1)+\hat{l}_{t_1}=0,\quad \hat{H}(\hat{t}_0)+\hat{l}_{t_0}=0.\qquad(12а)$$

Соответствующие преобразования проделываются так же, как в п. 4.1.1, и нет нужды их повторять. В точках разрыва управления в уравнении (8а) надо брать правую производную $\hat{p}'_+(\cdot)$ (левая тоже существует). Непрерывность $\hat{p}(\cdot)$ и $\hat{H}(\cdot)$ в точках разрыва управления (т. е. в точках излома $x(\cdot)$) носит название *условия Вейерштрасса—Эрдмана*.

Можно проделать и здесь анализ «на полноту» набора необходимых условий, даваемых сформулированной теоремой. В сущности он не будет отличаться от того, что было сделано для задачи Лагранжа. Вместо того, чтобы определять $u(\cdot)$ как функцию $x(\cdot)$ и $p(\cdot)$ из условия $L_u=0$, теперь это нужно сделать, исходя из соотношения (10).

4.2.3. Игольчатые вариации. Как и в простейшей ситуации, разобранной нами в п. 1.5.4, основной прием в доказательстве принципа максимума, которое будет изложено далее, состоит в замене рассматриваемой нами задачи оптимального управления некоторой конечномерной экстремальной задачей или, точнее говоря, целым набором таких задач. Для этого мы включим оптимальный процесс $(\hat{x}(\cdot),\ \hat{u}(\cdot),\ \hat{t}_0,\ \hat{t}_1)$ в некоторое специальное семейство управляемых процессов — «пакет» игольчатых вариаций — и рассмотрим ограничение задачи (1)—(3) п. 4.2.2 на это семейство.

Нужное нам семейство процессов зависит от следующих параметров:

начальные данные (t_0, x_0), конечный момент времени t_1;

набор $\bar{\alpha}=(\alpha_1,\ \ldots,\ \alpha_N)$, где все $\alpha_i \in \mathbf{R}$ достаточно малы; для краткости обозначим $\alpha=\sum \alpha_i$;

набор $\bar{\tau}=(\tau_1,\ \ldots,\ \tau_N)$, где $\hat{t}_0<\tau_1\leqslant\tau_2\leqslant\ldots\leqslant\tau_N<\hat{t}_1$, причем среди τ_i содержатся все точки разрыва оптимального управления $\hat{u}(\cdot)$;

набор $\bar{v}=(v_1,\ \ldots,\ v_N)$, где $v_i \in \mathfrak{U}$.

В дальнейшем $\bar{\tau}$ и $\bar{v}$ фиксированы, а $(t_1,\ t_0,\ x_0,\ \bar{\alpha})$ меняются, и по ним мы будем дифференцировать различные функции.

Пусть сначала $N=1$. Элементарная игольчатая вариация управления $\hat{u}(t)$, или просто «элементарная иголка», отвечающая паре $(\tau,\ v)$, определяется равенствами

$$u(t;\ \alpha,\ \tau,\ v)=\begin{cases}\hat{u}(t), & t\notin[\tau,\ \tau+\alpha),\\ v, & t\in[\tau,\ \tau+\alpha).\end{cases}\tag{1}$$

Соответственно игольчатая вариация $x(t;\ \alpha,\ \tau,\ v)$ фазовой траектории $\hat{x}(t)$ определяется как решение задачи Коши

$$\begin{aligned}\dot{x}&=\varphi(t,\ x,\ u(t;\ \alpha,\ \tau,\ v)),\\ x(\tau)&=\hat{x}(\tau).\end{aligned}\tag{2}$$

С точностью до несущественного различия (вместо $[\tau,\ \tau+\alpha)$ был взят полуинтервал $[\tau-\lambda,\ \tau))$ мы рассматривали подобную вариацию в п. 1.5.4. Было обнаружено, что функция $x(t;\ \alpha,\ \tau,\ v)$ дифференцируема по α и при $t\geqslant\tau$ ее производная $y(t)=x_\alpha(t;\ 0,\ \tau,\ v)$ является решением уравнения в вариациях

$$\dot{y}=\hat{\varphi}_x(t)\,y\tag{3}$$

с начальным условием

$$y(\tau)=\varphi(\tau,\ \hat{x}(\tau),\ v)-\varphi(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau)). \qquad (4)$$

Следуя п. 2.5.4, мы будем здесь и в остальной части параграфа обозначать через $\Omega(t,\ \tau)$ фундаментальную матрицу решений уравнения (3). Кроме того, для сокращения записи введем специальное обозначение для правой части равенства (4):

$$\Delta\hat{\varphi}(\tau,\ v)=\varphi(\tau,\ \hat{x}(\tau),\ v)-\varphi(\tau,\ \hat{x}(\tau),\ \hat{u}(\tau)). \qquad (5)$$

В этих обозначениях

$$y(t)=x_\alpha(t;\ 0,\ \tau,\ v)=\Omega(t,\ \tau)\,\Delta\hat{\varphi}(\tau,\ v). \qquad (6)$$

В п. 1.5.4 мы смогли обойтись рассмотрением одной «иголки». Теперь нам понадобится целый их «пакет», поскольку задача стала сложнее, и нам нужно иметь в распоряжении больше свободных параметров. В пакете объединяется любое конечное число иголок с параметрами $(\tau_i,\ v_i,\ \alpha_i)$, $i=1,\ \ldots,\ N$. При этом существенно, что некоторые из этих иголок могут различаться только значениями управления v_i при одних и тех же τ_i. Поэтому задавать иголки формулами вида (1) теперь нельзя: при одинаковых τ_i разные иголки будут накладываться друг на друга. Чтобы избежать этого, мы сдвинем полуинтервал действия i-й иголки на величину $i\alpha=i\sum\limits_{j=1}^{N}\alpha_j$, так что теперь это будет $\Delta_i=[\tau_i+i\alpha,\ \tau_i+i\alpha+\alpha_i]$ (очевидно, что если $\alpha_j>0$ и достаточно малы, то Δ_i не перекрываются).

Согласно сделанным в п. 4.2.1 предположениям оптимальное управление $\hat{u}(\cdot)$ непрерывно слева в точке $\hat{t}_1$ и справа — в точке $\hat{t}_0$. Продолжим $\hat{u}(\cdot)$ вне отрезка $[\hat{t}_0,\ \hat{t}_1]$ с сохранением непрерывности (например, положив $\hat{u}(t)\equiv\equiv\hat{u}(\hat{t}_0)$ при $t<\hat{t}_0$ и $\hat{u}(t)\equiv\hat{u}(\hat{t}_1)$ при $t>\hat{t}_1$) и впредь не будем этого оговаривать.

Пусть все $\alpha_i>0$. Определим игольчатую вариацию управления $\hat{u}(\cdot)$ формулами

$$u(t;\ \bar{\alpha},\ \bar{\tau},\ \bar{v})=\begin{cases}\hat{u}(t), & t\in(\hat{t}_0-\delta,\ \hat{t}_1+\delta)\setminus\bigcup\limits_{j=1}^{N}\Delta_i,\\ v_i, & t\in\Delta_i=\left[\tau_i+i\sum\limits_{j=1}^{N}\alpha_j,\ \tau_i+i\sum\limits_{j=1}^{N}\alpha_j+\alpha_i\right).\end{cases} \qquad (7)$$

Соответствующее семейство фазовых траекторий $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ определим как решение задачи Коши

$$\dot{x} = \varphi(t, x, u(t; \bar{\alpha}, \bar{\tau}, \bar{v})), \quad x(t_0) = x_0. \tag{8}$$

Для краткости далее $\hat{x}_0 = \hat{x}(\hat{t}_0)$.

Лемма о пакете иголок. 1) *При достаточно малых $\varepsilon_0 > 0$, $\delta > 0$ решение задачи Коши* (8) *такое, что*

$$|t_0 - \hat{t}_0| < \varepsilon_0, \ |x_0 - \hat{x}_0| < \varepsilon_0, \quad 0 < \alpha_j < \varepsilon_0, \tag{9}$$

определено для $\hat{t}_0 - \delta \leqslant t \leqslant \hat{t}_1 + \delta$.

2) *Если* $t_0 \to \hat{t}_0$, $x_0 \to \hat{x}_0$ *и* $\alpha_j \downarrow 0$, *то* $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}) \to \hat{x}(t)$ *равномерно на отрезке* $[\hat{t}_0, \hat{t}_1]$.

3) *Отображение* $(t_1, t_0, x_0, \bar{\alpha}) \mapsto x(t_1; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ *может быть продолжено до отображения, определенного и непрерывно дифференцируемого в некоторой окрестности точки* $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$, *и при этом*

$$\partial\hat{x}/\partial t_1 = x_{t_1}(\hat{t}_1; \hat{t}_0, \hat{x}_0, \bar{0}, \bar{\tau}, \bar{v}) = \\ = \varphi(\hat{t}_1, \hat{x}(\hat{t}_1), \hat{u}(\hat{t}_1)) = \hat{\varphi}(\hat{t}_1), \tag{10}$$

$$\partial\hat{x}/\partial t_0 = x_{t_0}(\hat{t}_1; \hat{t}_0, \hat{x}_0, \bar{0}, \bar{\tau}, \bar{v}) = \\ = -\Omega(\hat{t}_1, \hat{t}_0)\varphi(\hat{t}_0, \hat{x}_0, \hat{u}(\hat{t}_0)) = -\Omega(\hat{t}_1, \hat{t}_0)\hat{\varphi}(\hat{t}_0), \tag{11}$$

$$\partial\hat{x}/\partial x_0 = x_{x_0}(\hat{t}_1; \hat{t}_0, \hat{x}_0, \bar{0}, \bar{\tau}, \bar{v}) = \Omega(\hat{t}_1, \hat{t}_0), \tag{12}$$

$$\partial\hat{x}/\partial\alpha_k = x_{\alpha_k}(\hat{t}_1; \hat{t}_0, \hat{x}_0, \bar{0}, \bar{\tau}, \bar{v}) = \\ = \Omega(\hat{t}_1, \hat{t}_0)\Delta\hat{\varphi}(\tau_k, v_k), \tag{13}$$

где $\Omega(t, \tau)$ — фундаментальная матрица решений системы уравнений в вариациях (3) и $\Delta\hat{\varphi}(\tau, v)$ определено формулой (5).

Доказательство этой леммы будет приведено в п. 4.2.6. Ясно, что она обеспечивает нам дифференцируемость терминальных членов, входящих в условия задачи п. 4.2.2. Что же касается интегральных членов, то, как и в п. 1.5.4, мы посвятим им отдельную лемму.

Лемма об интегральных функционалах. *Пусть функции $f_i(t, x, u)$ удовлетворяют тем же условиям, что и в формулировке теоремы* п. 4.2.2.

Если $u(t, \bar{\alpha}, \bar{\tau}, \bar{v})$ определено при $\alpha_i > 0$ формулами (7), *а $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ — решение задачи Коши* (8), *то функции*

$$F_i(t_1, t_0, x_0, \bar{\alpha}) = \int_{t_0}^{t_1} f_i(t, x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}), u(t; \bar{\alpha}, \bar{\tau}, \bar{v}))\, dt, \quad i = 0, 1, \ldots, m,$$

могут быть продолжены до непрерывно дифференцируемых в некоторой окрестности точки $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$ *функций и при этом*

$$\partial \hat{F}_i/\partial t_1 = F_{it_1}(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0}) = f_i(\hat{t}_1, x(\hat{t}_1), \hat{u}(t_1)) = \hat{f}_i(\hat{t}_1), \tag{14}$$

$$\partial \hat{F}_i/\partial t_0 = F_{it_0}(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0}) = -f_i(\hat{t}_0, \hat{x}_0, \hat{u}(\hat{t}_0)) + p_{0i}(\hat{t}_0)\varphi(\hat{t}_0, \hat{x}_0, \hat{u}(\hat{t}_0)) = -\hat{f}_i(\hat{t}_0) + p_{0i}(\hat{t}_0)\hat{\varphi}(\hat{t}_0), \tag{15}$$

$$\partial \hat{F}_i/\partial x_0 = F_{ix_0}(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0}) = -p_{0i}(\hat{t}_0), \tag{16}$$

$$\partial \hat{F}_i/\partial \alpha_k = F_{i\alpha_k}(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0}) = \Delta \hat{f}_i(\tau_k, v_k) - p_{0i}(\tau_k)\Delta\hat{\varphi}(\tau_k, v_k), \tag{17}$$

где $p_{0i}(\tau)$ — *решение задачи Коши*

$$dp_{0i}(\tau)/d\tau = -p_{0i}(\tau)\varphi_x(\tau, \hat{x}(\tau), \hat{u}(\tau)) + f_{ix}(\tau, \hat{x}(\tau), \hat{u}(\tau)), \tag{18}$$

$$p_{0i}(\hat{t}_1) = 0$$

для неоднородной линейной системы, сопряженной к системе уравнений в вариациях (3), *а* $\Delta \hat{f}_i(\tau, v)$ *определяется формулой, аналогичной* (5). Доказательство этой леммы приведено в п. 4.2.7.

4.2.4. Редукция к конечномерной задаче. График оптимальной фазовой траектории $\hat{x}(t)$ лежит в области G и в силу второго утверждения леммы п. 4.2.3. (о пакете иголок) график решения $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ задачи Коши (8) того же пункта также лежит в этой области при $t \in [t_0, t_1]$, если выполнены неравенства (9) п. 4.2.3 с достаточно малым ε_0. Значит, четверка $(x(\cdot\,; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}), u(\cdot\,; \bar{\alpha}, \bar{\tau}, \bar{v}), t_0, t_1)$ является управляемым процессом для задачи (1)—(3) п. 4.2.2.

Положим

$$\Psi_i(t_1, t_0, x_0, \bar{\alpha}) = \psi_i(t_0, x_0, t_1, x(t_1; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})), \quad (1)$$

$$F_i(t_1, t_0, x_0, \bar{\alpha}) = \int_{t_0}^{t_1} f_i(t, x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}), u(t; \bar{\alpha}, \bar{\tau}, \bar{v}))\, dt,$$

$$i = 0, 1, \ldots, m; \quad (2)$$

В силу третьего утверждения леммы о пакете иголок и по теореме о суперпозиции, функции (1) непрерывно дифференцируемы в некоторой окрестности точки $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$. Для функций (2) то же самое следует из леммы об интегральных функционалах п. 4.2.3.

Теперь мы можем рассмотреть конечномерную экстремальную задачу — сужение задачи (1)—(3) п. 4.2.2 на построенное семейство управляемых процессов:

$$I_0(t_1, t_0, x_0, \bar{\alpha}) = F_0(t_1, t_0, x_0, \bar{\alpha}) + \Psi_0(t_1, t_0, x_0, \bar{\alpha}) \to \inf, \quad (3)$$

$$I_i(t_1, t_0, x_0, \bar{\alpha}) = F_i(t_1, t_0, x_0, \bar{\alpha}) + \Psi_i(t_1, t_0, x_0, \bar{\alpha}) \leqslant 0, \quad (4)$$

$$i = 1, 2, \ldots, m,$$

$$\alpha_k \geqslant 0, \quad k = 1, 2, \ldots, N. \quad (5)$$

Для этой задачи точка $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$ является локальным решением. Действительно, если выполнены ограничения (4), то четверка $(x(\cdot\,; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}), u(\cdot\,; \bar{\alpha}, \bar{\tau}, \bar{v}), t_0, t_1)$ является допустимым управляемым процессом задачи (1)—(3) п. 4.2.2. Если ε_0 в неравенствах (9) п. 4.2.3 достаточно мало, то в силу второго утверждения леммы о пакете иголок выполняются неравенства (4) п. 4.2.1. Следовательно, верно неравенство (5) п. 4.2.1, откуда

$$I_0(t_1, t_0, x_0, \bar{\alpha}) =$$

$$= \mathcal{I}_0(x(\cdot\,; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}), u(\cdot\,, \bar{\alpha}, \bar{\tau}, \bar{v}), t_0, t_1) \geqslant$$

$$\geqslant \mathcal{I}_0(\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1) = I_0(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0}),$$

а это и означает, что $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$ — локальное решение задачи (3)—(5).

Применяя к задаче (3)—(5) правило множителей Лагранжа из § 3.2, получаем следующее утверждение.

Лемма (правило множителей Лагранжа для вспомогательной конечномерной задачи). *Существуют такие множители Лагранжа*

$$\hat{\lambda}_0 \geqslant 0,\ \hat{\lambda}=(\hat{\lambda}_1,\ \ldots,\ \hat{\lambda}_m)\in \mathbf{R}^{m*},\quad \hat{\mu}=(\mu_1,\ \ldots,\ \mu_N), \tag{6}$$

не равные нулю одновременно, что для функции Лагранжа

$$\Lambda(t_1,\ t_0,\ x_0,\ \overline{\alpha};\ \lambda,\ \mu,\ \lambda_0)=\sum_{i=0}^{m}\lambda_i(F_i+\Psi_i)+\sum_{k=1}^{N}\mu_k\alpha_k \tag{7}$$

выполняются:

1) *условия стационарности*

$$\hat{\Lambda}_{t_1}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{it_1}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{it_1}=0, \tag{8}$$

$$\hat{\Lambda}_{t_0}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{it_0}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{it_0}=0, \tag{9}$$

$$\hat{\Lambda}_{x_0}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{ix_0}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{ix_0}=0, \tag{10}$$

$$\hat{\Lambda}_{\alpha_k}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{i\alpha_k}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{i\alpha_k}+\mu_k=0, \tag{11}$$

где обозначено $\hat{\Lambda}_{\alpha_k}=\Lambda_{\alpha_k}(\hat{t}_1,\ \hat{t}_0,\ \hat{x}_0,\ \overline{0};\ \hat{\lambda},\ \hat{\mu},\ \hat{\lambda}_0)$ *и т. п.*;

2) *условия согласования знаков*

$$\hat{\lambda}_i \geqq 0,\quad \hat{\mu}_k \leqslant 0;\quad i=1,\ 2,\ \ldots,\ m;\quad k=1,\ 2,\ \ldots,N; \tag{12}$$

3) *условия дополняющей нежесткости*

$$\hat{\lambda}_i[F_i(\hat{t}_1,\ \hat{t}_0,\ \hat{x}_0,\ \overline{0})+\Psi_i(\hat{t}_1,\ \hat{t}_0,\ \hat{x}_0,\ \overline{0})]=0,\quad i=1,\ 2,\ \ldots,\ m. \tag{13}$$

4.2.5. Доказательство принципа максимума. За исключением одного не совсем тривиального места, остальная часть доказательства будет посвящена расшифровке условий (8)—(13) п. 4.2.4. Собственно принцип максимума (или в лагранжевой форме — принцип минимума), т. е. соотношения (11) (или (10)) п. 4.2.2, получается при этом из равенств (11) п. 4.2.4, т. е. условий $\hat{\Lambda}_{\alpha_k}=0$. Каждое такое условие соответствует одной из иголок, входящих в «пакет», и если эта иголка определялась парой $(\tau_k,\ v_k)$, то мы получаем из него неравенство $H(\tau_k,\ \hat{x}(\tau_k),\ v_k,\ \cdot)\leqslant \leqslant H(\tau_k,\ \hat{x}(\tau_k),\ \hat{u}(\tau_k),\ \cdot)$. Таким образом, чтобы получить принцип максимума, мы должны перебрать иголки, отвечающие всем возможным парам (τ, v), тогда как наш «пакет» содержит любое, но конечное их число. Положение спасают простые топологические соображения, основанные на компактности. Применяя «лемму о центрированной системе», мы можем выбрать «универсальные» множители Лагранжа, пригодные для всего множества иголок.

А) Существование «универсальных» множителей Лагранжа. Изучим подробнее условия (8)—(13) п. 4 2.4, внеся в них значения производных функций F_i и Ψ_i, вычисленные при помощи двух лемм п. 4.2.3, и уделяя особое внимание членам, в которые входят параметры иголок (τ_k, v_k).

Производные функций F_i даны прямо в лемме об интегральных функционалах и, обозначив

$$A_i(\tau, v) = f_i(\tau, \hat{x}(\tau), v) - f_i(\tau, \hat{x}(\tau), \hat{u}(\tau)) - \\ - p_{0i}(\tau)[\varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)] = \\ = \Delta\hat{f}(\tau, v) - p_{0i}(\tau)\Delta\hat{\varphi}(\tau, v), \quad (1)$$

мы видим, что

$$\hat{F}_{i\alpha_k} = A_i(\tau_k, v_k),$$

а ни в одну из остальных производных параметры иголок не входят, так что эти производные имеют одно и то же значение для любого «пакета иголок» п. 4.2.3. Аналогично производные функций

$$\Psi_i(t_1, t_0, x_0, \bar{\alpha}) = \psi_i(t_0, x_0, t_1, x(t_1; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}))$$

вычисляются по теореме о суперпозиции и лемме о пакете иголок, причем от параметров иголок зависят лишь $\hat{\Psi}_{i\alpha_k}$. Если обозначить

$$B_i(\tau, v) = \frac{\partial\psi_i}{\partial x_1}(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1))\,\Omega(\hat{t}_1, \tau)[\varphi(\tau, \hat{x}(\tau), v) - \\ - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau))] = \hat{\psi}_{ix_1}\Omega(\hat{t}_1, \tau)\,\Delta\hat{\varphi}(\tau, v), \quad (2)$$

то

$$\hat{\Psi}_{i\alpha_k} = B_i(\tau_k, v_k),$$

а значения производных $\hat{\Psi}_{it_0}$, $\hat{\Psi}_{it_1}$, $\hat{\Psi}_{ix_0}$ одни и те же для любого «пакета иголок».

Теперь имеем

$$\hat{\Lambda}_{\alpha_k} = \sum_{i=0}^{m} \hat{\lambda}_i \hat{F}_{i\alpha_k} + \sum_{i=0}^{m} \hat{\lambda}_i \hat{\Psi}_{i\alpha_k} + \hat{\mu}_k =$$

$$= \sum_{i=0}^{m} \hat{\lambda}_i (A_i(\tau_k, v_k) + B_i(\tau_k, v_k)) + \hat{\mu}_k.$$

Поскольку, согласно условиям (12) и (11) п. 4.2.4, $\hat{\mu}_k \leqslant 0$ и $\hat{\Lambda}_{\alpha_k} = 0$, должны выполняться неравенства

$$\sum_{i=0}^{m} \hat{\lambda}_i [A_i(\tau_k, v_k) + B_i(\tau_k, v_k)] \geqslant 0. \qquad (3)$$

Рассмотрим теперь в пространстве $\mathbf{R}^{(m+1)*}$ следующие множества:

$$S = \{(\lambda, \lambda_0) \mid \sum_{i=0}^{m} \lambda_i^2 = 1\},$$

$$K(\tau, v) = \{(\lambda, \lambda_0) \mid \sum_{i=0}^{m} \lambda_i (A_i(\tau, v) + B_i(\tau, v)) \geqslant 0\},$$

$$T_l = \{(\lambda, \lambda_0) \mid \sum_{i=0}^{m} \lambda_i (\hat{F}_{it_l} + \hat{\Psi}_{it_l}) = 0\}, \quad l = 0, 1,$$

$$X = \{(\lambda, \lambda_0) \mid \sum_{i=0}^{m} \lambda_i (\hat{F}_{ix_0} + \hat{\Psi}_{ix_0}) = 0\},$$

и, наконец, множество Z, состоящее из тех (λ, λ_0), для которых $\lambda_0 \geqslant 0$, а λ_i, $i = 1, \ldots, m$, удовлетворяют условиям согласования знаков и дополняющей нежесткости (12) и (13) п. 4.2.4, тождественных условиям (13) и (14) п. 4.2.2.

Все эти множества замкнуты, а сфера S еще и компактна, как замкнутое ограниченное подмножество конечномерного пространства $\mathbf{R}^{(m+1)*}$.

Условия стационарности (8)—(10) п. 4.2.4 соответствуют включениям $(\hat{\lambda}, \hat{\lambda}_0) \in T_1, T_0, X$, условия согласования знаков и дополняющей нежесткости — включению $(\hat{\lambda}, \hat{\lambda}_0) \in Z$ (все эти условия одинаковы для всех пакетов иголок). Зависящие от иголок условия (11) п. 4.2.4 выполняются, когда $(\hat{\lambda}, \hat{\lambda}_0) \in K(\tau_k, v_k)$, $k = 1, \ldots, N$. Наконец, поскольку множители Лагранжа определены с точностью до пропорциональности и не могут одновременно обращаться в нуль, их можно пронормировать так, чтобы было $(\hat{\lambda}, \hat{\lambda}_0) \in S$. Поэтому утверждение леммы предыдущего пункта означает, что

$$S \cap T_0 \cap T_1 \cap X \cap Z \cap \bigcap_{k=1}^{N} K(\tau_k, v_k) \neq \varnothing. \tag{4}$$

Рассмотрим систему замкнутых подмножеств

$$\tilde{K}(\tau, v) = S \cap T_0 \cap T_1 \cap X \cap Z \cap K(\tau, v),$$
$$\tau \in [\hat{t}_0, \hat{t}_1], \quad v \in \mathfrak{U}, \tag{5}$$

компактного множества S.

Лемма о центрированной системе. *Система множеств $\tilde{K}(\tau, v)$ имеет непустое пересечение.*

Доказательство. Согласно (4) любое конечное число множеств $\tilde{K}(\tau, v)$ имеет непустое пересечение. Такая система множеств называется *центрированной*. По известной теореме [КФ, стр. 99] пересечение центрированной системы подмножеств компакта непусто. ■

Таким образом, существуют

$$(\hat{\lambda}, \hat{\lambda}_0) \in \bigcap_{\substack{\tau \in [\hat{t}_0, \hat{t}_1], \\ v \in \mathfrak{U}}} \tilde{K}(\tau, v) =$$
$$= S \cap T_0 \cap T_1 \cap X \cap Z \cap \bigcap_{\substack{\tau \in [\hat{t}_0, \hat{t}_1] \\ v \in \mathfrak{U}}} K(\tau, v). \tag{6}$$

Это и есть искомые «универсальные» множители Лагранжа. Действительно, $\hat{\lambda}_i$ не равны нулю одновременно ($(\hat{\lambda}, \hat{\lambda}_0) \in S$), удовлетворяют условиям согласования знаков и дополняющей нежесткости ($(\hat{\lambda}, \hat{\lambda}_0) \in Z$), удовлетворяют условиям $\hat{\Lambda}_{t_l} = 0$, $l = 0, 1$, и $\hat{\Lambda}_{x_0} = 0$ ($(\hat{\lambda}, \hat{\lambda}_0) \in T_0 \cap T_1 \cap X$) и, наконец,

$$\sum_{i=0}^{m} \hat{\lambda}_i [A_i(\tau, v) + B_i(\tau, v)] \geqslant 0 \tag{7}$$

для всех $\tau \in [\hat{t}_0, \hat{t}_1]$, $v \in \mathfrak{U}$.

Б) Вывод принципа максимума. Подставив (1) и (2) в неравенство (7), преобразуем его к виду

$$0 \leqslant \sum_{i=0}^{m} \hat{\lambda}_i [\Delta \hat{f}_i(\tau, v) - p_{0i}(\tau)\Delta\hat{\varphi}(\tau, v)] +$$

$$+ \sum_{i=0}^{m} \hat{\lambda}_i \frac{\partial \hat{\psi}_i}{\partial x_1} \Omega(\hat{t}_1, \tau)[\Delta\hat{\varphi}(\tau, v)] =$$

$$= f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau)) -$$

$$- \hat{p}(\tau)[\varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau))] =$$

$$= -H(\tau, \hat{x}(\tau), v, \hat{p}(\tau)) + H(\tau, \hat{x}(\tau), \hat{u}(\tau), \hat{p}(\tau)), \quad (8)$$

где

$$f(\tau, x, u) = \sum_{i=0}^{m} \hat{\lambda}_i f_i(\tau, x, u), \quad (9)$$

$$\hat{p}(\tau) = \sum_{i=0}^{m} \hat{\lambda}_i p_{0i}(\tau) - \sum_{i=0}^{m} \hat{\lambda}_i \frac{\partial \hat{\psi}_i}{\partial x_1} \Omega(\hat{t}_1, \tau) \quad (10)$$

— решение уравнения Эйлера—Лагранжа (8а) п. 4.2.2, как мы увидим далее, и

$$H(\tau, x, u, p) = p\varphi(\tau, x, u) - f(\tau, x, u) \quad (11)$$

— функция Понтрягина (ср. (7) в п. 4 2 2).

Неравенство (8), справедливое при всех $\tau \in [\hat{t}_0, \hat{t}_1]$ и $v \in \mathfrak{U}$, равносильно принципу максимума (11) п. 4.2.2

$$H(\tau, \hat{x}(\tau), \hat{u}(\tau), \hat{p}(\tau)) = \max H(\tau, \hat{x}(\tau), v, \hat{p}(\tau)). \quad (12)$$

В) Вывод условий стационарности по $x(\cdot)$. Дифференцируя (10) и учитывая, что $p_{0i}(\tau)$ удовлетворяют уравнениям (18) п. 4.2.3, а

$$\frac{\partial \Omega(t, \tau)}{\partial \tau} = -\Omega(t, \tau)\hat{\varphi}_x(\tau),$$

(см. теорему п. 2.5.4), имеем

$$d\hat{p}(\tau)/d\tau = \sum_{i=0}^{m} \hat{\lambda}_i \{-p_{0i}(\tau)\hat{\varphi}_x(\tau) + \hat{f}_{ix}(\tau)\} -$$

$$- \sum_{i=0}^{m} \hat{\lambda}_i \frac{\partial \hat{\psi}_i}{\partial x_1} \Omega(\hat{t}_1, \tau)\hat{\varphi}_x(\tau) = -\hat{p}(\tau)\hat{\varphi}_x(\tau) + \sum_{i=0}^{m} \hat{\lambda}_i \hat{f}_{ix}(\tau) =$$

$$= -\hat{p}(\tau)\hat{\varphi}_x(\tau) + \hat{f}_x(\tau),$$

т. е $\hat{p}(\cdot)$ является решением уравнения Эйлера—Лагранжа (8а) п. 4.2.2.

Далее, прямо из (10)

$$\hat{p}(\hat{t}_1)=\sum_{i=0}^{m}\hat{\lambda}_i p_{0i}(\hat{t}_1)-\sum_{i=0}^{m}\hat{\lambda}_i\frac{\partial\hat{\psi}_i}{\partial x_1}\Omega(\hat{t}_1,\ \hat{t}_1)=$$

$$=-\sum_{i=0}^{m}\hat{\lambda}_i\frac{\partial\hat{\psi}_i}{\partial x_1}=-\frac{\partial\hat{l}}{\partial x_1},\quad(13)$$

поскольку $p_{0i}(\hat{t}_1)=0$ в силу (18) п 4.2.3, а $\Omega(t,\ t)\equiv E$. Здесь, как и в (6) п. 4.2.2,

$$l(t_0,\ x_0,\ t_1,\ x_1)=\sum_{i=0}^{m}\hat{\lambda}_i\psi_i(t_0,\ x_0,\ t_1,\ x_1)\quad(14)$$

— терминант, а полученное равенство (13) равносильно первому равенству (9а) из п. 4.2.2 .

Наконец, условие $\hat{\Lambda}_{x_0}=0$, т. е. (10) п. 4.2.4, дает после подстановки значений производных (12) и (16) п. 4.2.3

$$0=\hat{\Lambda}_{x_0}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{ix_0}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{ix_0}=$$

$$=-\sum_{i=0}^{m}\hat{\lambda}_i p_{0i}(\hat{t}_0)+\sum_{i=0}^{m}\hat{\lambda}_i\left(\hat{\psi}_{ix_0}+\hat{\psi}_{ix_1}\frac{\partial\hat{x}}{\partial x_0}\right)=$$

$$=-\sum_{i=0}^{m}\hat{\lambda}_i p_{0i}(\hat{t}_0)+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\psi}_{ix_0}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\psi}_{ix_1}\Omega(\hat{t}_1,\ \hat{t}_0)=$$

$$=-\hat{p}(\hat{t}_0)+\hat{l}_{x_0},$$

так что верно и второе условие трансверсальности (9а) п. 4 2.2.

Г) Вывод условий стационарности по t_i. После подстановки значений производных (10), (11), (14), (15) п. 4.2.3 в условия $\hat{\Lambda}_{t_0}=\hat{\Lambda}_{t_1}=0$ (8) и (9) п. 4.2.4 получаем с учетом обозначений (9)—(11) и уже доказанного равенства (13)

$$0=\hat{\Lambda}_{t_1}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{it_1}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{it_1}=$$

$$=\sum_{i=0}^{m}\hat{\lambda}_i\hat{f}_i(\hat{t}_1)+\sum_{i=0}^{m}\hat{\lambda}_i\left[\hat{\psi}_{it_1}+\hat{\psi}_{ix_1}\frac{\partial\hat{x}}{\partial t_1}\right]=$$

$$=\hat{f}(\hat{t}_1)+\hat{l}_{t_1}+\hat{l}_{x_1}\hat{\varphi}(\hat{t}_1)=\hat{f}(\hat{t}_1)-\hat{p}(\hat{t}_1)\hat{\varphi}(\hat{t}_1)+\hat{l}_{t_1}=-\hat{H}(t_1)+\hat{l}_{t_1}$$

и

$$0=\hat{\Lambda}_{t_0}=\sum_{i=0}^{m}\hat{\lambda}_i\hat{F}_{it_0}+\sum_{i=0}^{m}\hat{\lambda}_i\hat{\Psi}_{it_0}=$$

$$=\sum_{i=0}^{m}\hat{\lambda}_i[-\hat{f}_i(\hat{t}_0)+p_{0i}(\hat{t}_0)\hat{\varphi}(\hat{t}_0)]+\sum_{i=0}^{m}\hat{\lambda}_i\left[\hat{\psi}_{it_0}+\hat{\psi}_{ix_1}\frac{\partial\hat{x}}{\partial t_0}\right]=$$

$$=-\hat{f}(\hat{t}_0)+\sum_{i=0}^{m}\hat{\lambda}_i p_{0i}(\hat{t}_0)\hat{\varphi}(\hat{t}_0)+\hat{l}_{t_0}+\hat{l}_{x_1}[-\Omega(\hat{t}_1,\ \hat{t}_0)\hat{\varphi}(\hat{t}_0)]=$$

$$=-\hat{f}(\hat{t}_0)+\hat{p}(\hat{t}_0)\hat{\varphi}(\hat{t}_0)+\hat{l}_{t_0}=\hat{H}(\hat{t}_0)+\hat{l}_{t_0}.$$

Этим доказаны обе формулы (12а) п. 4.2.2.

Д) Непрерывность функции $\hat{H}(\cdot)$. Вместе с управлением $\hat{u}(t)$ функция

$$\hat{H}(t)=H(t,\ \hat{x}(t),\ \hat{u}(t),\ \hat{p}(t))=$$
$$=-f(t,\ \hat{x}(t),\ \hat{u}(t))+\hat{p}(t)\varphi(t,\ \hat{x}(t),\ \hat{u}(t))=$$
$$=-\sum_{i=0}^{m}\hat{\lambda}_i f_i(t,\ \hat{x}(t),\ \hat{u}(t))+\hat{p}(t)\varphi(t,\ \hat{x}(t),\ \hat{u}(t))$$

непрерывна справа. Кроме того, каждая из функций $H(t,\ \hat{x}(t),\ v,\ \hat{p}(t))$ непрерывна по t и, переходя к пределу в неравенстве

$$H(t,\ \hat{x}(t),\ v,\ \hat{p}(\tau))\leqslant H(t,\ \hat{x}(t),\ \hat{u}(t),\ \hat{p}(t))=\hat{H}(t),\qquad(15)$$

мы получаем

$$H(\tau,\ \hat{x}(\tau),\ v,\ \hat{p}(\tau))\leqslant\lim_{t\to\tau-0}\hat{H}(t)=\hat{H}(\tau-0),$$

откуда

$$\hat{H}(\tau)=\sup_{v\in\mathfrak{U}}H(\tau,\ \hat{x}(t),\ v,\ p(\tau))\leqslant\hat{H}(\tau-0).\qquad(16)$$

С другой стороны,

$$\hat{H}(\tau-0)=\lim_{t\to\tau-0}H(t,\hat{x}(t),\ \hat{u}(t),\hat{p}(t))=$$
$$=H(\tau,\hat{x}(\tau),\hat{u}(\tau-0),\ \hat{p}(\tau))=\lim_{v\to\hat{u}(\tau-0)}H(\tau,\hat{x}(\tau),v,\hat{p}(\tau))\leqslant$$
$$\leqslant H(\tau,\hat{x}(t),\hat{u}(\tau),\hat{p}(\tau))=\hat{H}(\tau)$$

в силу того же неравенства (15). Вместе с (16) это означает, что $\hat{H}(\tau-0)=\hat{H}(\tau)$, т. е. $\hat{H}(t)$ непрерывна и слева. ■

4.2.6. Доказательство леммы о пакете иголок. Напомним, что игольчатая вариация $x(t;\ t_0, x_0, \overline{\alpha}, \overline{\tau}, \overline{v})$ оптимальной фазовой траектории $\hat{x}(\cdot)$—это решение задачи Коши

$$\dot{x}=\varphi(t, x, u(t;\ \overline{\alpha}, \overline{\tau}, \overline{v})), \tag{1}$$

$$x(t_0)=x_0, \tag{2}$$

где в свою очередь вариация $u(t;\ \overline{\alpha}, \overline{\tau}, \overline{v})$ оптимального управления $\hat{u}(t)$ определяется при $\alpha_j>0$ формулами

$$u(t;\ \overline{\alpha}, \overline{\tau}, \overline{v})=\begin{cases} v_i, & t\in\Delta_i=[\tau_i+i\alpha,\ \tau_i+i\alpha+\alpha_i),\\ \hat{u}(t), & t\notin\bigcup\limits_i\Delta_i;\ \alpha=\sum\limits_{j=1}^{N}\alpha_j. \end{cases} \tag{3}$$

Доказательство леммы о пакете иголок (лемма п. 4.2.3) удобно разбить на две части. В первой из них, выделенной в отдельное утверждение, мы доказываем равномерную сходимость проварьированных фазовых траекторий к $\hat{x}(\cdot)$. Существенную роль здесь играет не столько кусочная непрерывность управлений, сколько интегральная сходимость соответствующих правых частей дифференциальной связи (см. ниже формулу (5)). Чтобы оттенить это, мы распространяем утверждение также и на случай измеримых управлений (определив сначала подходящим образом понятие измеримости в случае, когда управления принимают значения в произвольном топологическом пространстве). Вторая часть доказательства леммы о пакете иголок непосредственно связана с кусочной непрерывностью и опирается на классическую теорему о гладкой зависимости решений от начальных условий.

Определение. Пусть $\mathfrak{U}$—произвольное топологическое пространство и I—отрезок числовой прямой. Отображение u: $I\to\mathfrak{U}$ называется *измеримым* (в смысле Лебега), если существует конечная или счетная последовательность непересекающихся измеримых подмножеств A_n такая, что:

а) ограничение $u\,|\,A_n$ продолжается до функции, непрерывной на замыкании $\overline{A}_n$;

б) $I\setminus\bigcup\limits_n A_n$ имеет лебегову меру нуль.

Ясно, что всякое кусочно-непрерывное отображение измеримо в смысле этого определения (в качестве A_n берем интервалы непрерывности). Почти также очевидно, что в случае, когда $\mathfrak{U}\subset\mathbf{R}^r$ функция, измеримая в смысле этого определения, измерима и в обычном смысле [КФ, гл. V, § 4].

Действительно, пусть $u_n(\cdot)$ — функция, непрерывная на $\overline{A}_n$, и совпадающая с $u(\cdot)$ на A_n. Полагая $u_n(t)=0$ вне $\overline{A}_n$, мы получаем измеримую функцию на $\mathbf{R}$. Характеристическая функция $\chi_{A_n}(\cdot)$ множества A_n также измерима, а поскольку $u(t)=\sum\limits_n \chi_{A_n}(t)\,u_n(t)$ почти всюду на I, измерима и $u(t)$.

Упражнение. Докажите, что если $\mathfrak{U}\subset\mathbf{R}^r$ и функция u: $I\to\mathfrak{U}$ измерима в стандартном лебеговом смысле, то она измерима и в смысле приведенного выше определения. Указание: воспользуйтесь C-свойством Лузина [КФ, стр. 291].

В формулировке следующей леммы $\mathfrak{A}$ — некоторое хаусдорфово топологическое пространство [КФ, стр. 95].

Лемма 1. *Пусть функции φ и φ_x непрерывны в $G\times\mathfrak{U}$ $\hat{x}$: $[\hat{t}_0, \hat{t}_1]\to\mathbf{R}^n$ — решение дифференциального уравнения*

$$\dot{x}=\varphi(t, x, \hat{u}(t)) \tag{4}$$

и его график $\Gamma=\{(t, \hat{x}(t))\mid \hat{t}_0\leqslant t\leqslant\hat{t}_1\}\subset G$.

Пусть далее $\delta_0>0$ и u_α: $(\hat{t}_0-\delta_0, \hat{t}_1+\delta_0)\to\mathfrak{U}$ — семейство измеримых управлений такое, что для всех t и всех $\alpha\in\mathfrak{A}$ значения $u_\alpha(t)$ содержатся в некотором компакте $\mathcal{K}_0\subset\mathfrak{U}$

Пусть, наконец, $\hat{u}(t)\equiv u_{\hat{\alpha}}(t)$ для некоторого $\hat{\alpha}\in\mathfrak{A}$ и

$$\lim_{\alpha\to\hat{\alpha}}\int\limits_{\hat{t}_0-\delta_0}^{\hat{t}_1+\delta_0}|\varphi(t, \hat{x}(t), u_\alpha(t))-\varphi(t, \hat{x}(t), \hat{u}(t))|\,dt=0. \tag{5}$$

Тогда решения $X_\alpha(t, t_0, x_0)$ задачи Коши

$$\dot{x}=\varphi(t, x, u_\alpha(t)),\quad x(t_0)=x_0$$

при некотором $\hat{\delta}>0$ определены на отрезке $[\hat{t}_0-\hat{\delta}, \hat{t}_1-\hat{\delta}]$ для всех (t_0, x_0, α) из некоторой окрестности точки $(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{\alpha})$ в $\mathbf{R}\times\mathbf{R}^n\times\mathfrak{A}$ и при $t_0\to\hat{t}_0$, $x_0\to\hat{x}(\hat{t}_0)$, $\alpha\to\hat{\alpha}$

$$X_\alpha(t, t_0, x_0)\to\hat{x}(t)$$

равномерно по $t\in[\hat{t}_0, \hat{t}_1]$.

Доказательство. Применим к рассматриваемому случаю теорему 2 п. 2.5.5, положив

$$F_\alpha(t, x) \doteq \varphi(t, x, u_\alpha(t)). \tag{6}$$

А) Если A_n — те множества, о которых говорится в определении измеримости, примененном к $u_\alpha(\cdot)$, то при фиксированном x функции $t \mapsto \varphi(t, x, u_\alpha(t))$ непрерывны на A_n, а следовательно, $t \mapsto F_\alpha(t, x)$ — измеримые функции и выполнено предположение А) п. 2.5.1. При фиксированном t функции F_α вместе с φ дифференцируемы по x, так что выполнено и предположение Б) п. 2.5.1.

Б) Для любого компакта $\mathscr{K} \subset G$ функции φ и φ_x непрерывны на компакте $\mathscr{K} \times \mathscr{K}_0$, а следовательно, и ограничены. Обозначив

$$M = \max_{\mathscr{K} \times \mathscr{K}_0} |\varphi(t, x, u)|, \quad M_x = \max_{\mathscr{K} \times \mathscr{K}_0} \|\varphi_x(t, x, u)\|,$$

мы видим, что выполняется условие В') теоремы 2 п. 2 5.5 с $\varkappa(t) \equiv M$ и $k(t) \equiv M_x$ поскольку по условию $u_\alpha(t) \in \mathscr{K}_0$. Условие Г) той же теоремы, совпадает с (5).

Таким образом, к рассматриваемой ситуации применима теорема 2 п. 2.5 5, из которой и вытекает доказываемое утверждение. ■

Доказательство леммы о пакете иголок. А) Проверим выполнение условий предыдущей леммы. Часть из них, относящаяся к φ и $\hat{x}(\cdot)$, входит в условия теоремы о принципе максимума.

Выше мы уже предположили, что управление $\hat{u}(\cdot)$ продолжено вне отрезка $[\hat{t}_0, \hat{t}_1]$, так что при $t \leqslant \hat{t}_0$ и при $t \geqslant \hat{t}_1$ это непрерывная функция. Поэтому мы будем считать далее, что $\hat{u}(\cdot)$ определена на отрезке $\Delta = [\hat{t}_0 - \delta_0, \hat{t}_1 + \delta_0]$, $\delta_0 > 0$, а ее точки разрыва лежат в интервале $(\hat{t}_0, \hat{t}_1)$

Пусть $t^{(1)} < t^{(2)} < \ldots < t^{(s)}$ — эти точки, $t^{(0)} = \hat{t}_0 - \delta_0$, $t^{(s+1)} = \hat{t}_1 + \delta_0$. Функция

$$u_\iota(t) = \begin{cases} \hat{u}(t), & t^{(\iota-1)} \leqslant t < t^{(\iota)}, \\ \hat{u}(t^{(\iota)} - 0), & t = t^{(\iota)}, \end{cases}$$

непрерывна на отрезке $I_\iota = [t^{(\iota-1)}, t^{(\iota)}]$ и образ $\mathscr{K}_\iota$ этого отрезка при непрерывном отображении u_ι: $I \to \mathfrak{U}$ компактен в $\mathfrak{U}$, поскольку I_ι — компакт. В соответствии с формулами (3) значения управлений $u(t; \bar{\alpha}, \bar{\tau}, \bar{v})$ лежат в компакте

$$\mathscr{K}_0 = \bigcup_{i=1}^{s+1} \mathscr{K}_i \cup \{v_1, \ldots, v_N\}.$$

Остается проверить условие интегральной сходимости (5). На компакте

$$\mathscr{K} = \{(t, \hat{x}(t), u) \mid t \in \Delta, \ u \in \mathscr{K}_0\} = \Gamma \times \mathscr{K}_0$$

непрерывная функция φ ограничена: $|\varphi| \leqslant M$. При столь малых $\alpha_j > 0$, что полуинтервалы Δ_j не пересекаются, из формул (3) имеем

$$\int_{\hat{t}_0 - \delta_0}^{\hat{t}_1 + \delta_0} |\varphi(t, \hat{x}(t), u(t; \bar{\alpha}, \bar{\tau}, \bar{v})) - \varphi(t, \hat{x}(t), \hat{u}(t))| \, dt =$$

$$= \sum_{j=1}^{N} \int_{\Delta_j} |\varphi(t, \hat{x}(t), v_j) - \varphi(t, \hat{x}(t), \hat{u}(t))| \, dt \leqslant 2M \sum_{j=1}^{N} \alpha_j =$$

$$= 2M\alpha \to 0$$

при $\alpha_j \downarrow 0$, так что условие (5) выполнено.

Применяя лемму 1, убеждаемся в справедливости первых двух утверждений леммы о пакете иголок п. 4.2.3: для достаточно малого $\varepsilon_0 > 0$ решение $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ задачи Коши (1), (2) при (t_0, x_0, α), удовлетворяющих неравенствам (9) п. 4.2.3, определено на $\Delta = [\hat{t}_0 - \hat{\delta}, \hat{t}_1 + \hat{\delta}]$ (первое утверждение), и при $t_0 \to \hat{t}_0$, $x_0 \to \hat{x}_0$, $\alpha_j \downarrow 0$

$$x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}) \to \hat{x}(t)$$

равномерно по $t \in [\hat{t}_0, \hat{t}_1]$ (второе утверждение).

Б) Перейдем к рассмотрению отображения $(t_1, t_0, x_0, \bar{\alpha}) \to x(t_1; t_0, \bar{\alpha}, \bar{\tau}, \bar{v})$. Обозначим, как обычно, через $X(t, t_0, x_0)$ решение задачи Коши

$$\dot{x} = \varphi(t, x, \hat{u}(t)), \quad x(t_0) = x_0. \tag{7}$$

По теореме 1 п. 2.5.5 функция $X(\cdot, \cdot, \cdot)$ определена и непрерывна в области $(\hat{t}_0 - \delta_0, \hat{t}_1 + \delta_0) \times \hat{G}$, где $\hat{G}$ — некоторая окрестность графика Γ решения $\hat{x}(\cdot)$. Если же мы сузим область определения так, чтобы t_0 и t не переходили через точки $t^{(i)}$ разрыва управления $\hat{u}(\cdot)$, то дифференциальное уравнение в (7) удовлетворяет условиям классической теоремы о дифференцируемой зависимости решения от начальных данных (п. 2.5.7), так что $X(t, t_0, x_0)$ будет непрерывно дифференцируемо по совокупности аргументов. В частности, X непрерывно дифференцируемо в некоторых окрестностях точек $(\hat{t}_k, \hat{t}_k, \hat{x}(\hat{t}_k))$, поскольку в некоторых окрестностях точек $\hat{t}_k$ управление $\hat{u}(\cdot)$ по нашим предположениям непрерывно.

Далее, обозначим через $\Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})$ значение при $t=\hat{t}_1$ решения задачи Коши для уравнения (1) с начальным условием $x(\hat{t}_0)=\xi$, т. е.

$$\Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})=x(\hat{t}_1, \hat{t}_0, \xi, \bar{\alpha}, \bar{\tau}, \bar{v}). \tag{8}$$

В соответствии с формулой (13) п. 2.5.5 и формулами (3), задающими управление $u(t, \bar{\alpha}, \bar{\tau}, \bar{v})$, имеем

$$x(t_1, t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})=X(t_1, \hat{t}_1, \Xi(X(\hat{t}_0, t_0, x_0), \bar{\alpha}, \bar{\tau}, \bar{v})) \tag{9}$$

(от t_0 до $\hat{t}_0$ мы решаем уравнение $\dot{x}=\varphi(t, x, u)$ с управлением $u=\hat{u}(t)$, а затем от $\hat{t}_0$ до $\hat{t}_1$—с управлением $u=u(t; \bar{\alpha}, \bar{\tau}, \bar{u})$ и, наконец, от $\hat{t}_1$ до t_1—снова с управлением $u=\hat{u}(t)$. Важно отметить, что, как и формула (13) п. 2.5.5 формула (9) верна при любом расположении точек t_i относительно $\hat{t}_i$.

Если мы теперь сможем продолжить отображение $(\xi, \bar{\alpha}) \mapsto \Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})$ до отображения класса C^1, определенного в некоторой окрестности точки $(\hat{x}(\hat{t}_0), \bar{0})$ (т. е. сможем продолжить его на отрицательные α_j с сохранением непрерывной дифференцируемости), то в соответствии с формулой (9) и отображение $(t_1, t_0, x_0, \bar{\alpha}) \mapsto x(t_1; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ будет продолжено на некоторую окрестность точки $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$, а по теореме п. 2.2.2 оно будет непрерывно дифференцируемым как суперпозиция трех непрерывно дифференцируемых отображений: $(t_0, x_0) \mapsto X(\hat{t}_0, t_0, x_0)$, $(\xi, \bar{\alpha}) \mapsto \Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})$, $(t_1, \eta_0) \mapsto (t_1, \hat{t}_1, \eta_0)$.

В) Пусть $\tau \in (\hat{t}_0, \hat{t}_1)$ и $v \in \mathfrak{U}$ фиксированы. Обозначим через $Y_v(t, t_0, y_0)$ решение задачи Коши

$$\dot{y}=\varphi(t, y, v), \quad y(t_0)=y_0. \tag{10}$$

В силу локальной теоремы существования (п. 2.5.2) и классической теоремы о дифференцируемости решения по начальным данным (п. 2.5.7) $Y_v(t, t_0, y_0)$ определено и непрерывно дифференцируемо в некоторой окрестности точки $(\tau, \tau, \hat{x}(\tau))$.

Если τ — точка разрыва управления $\hat{u}(\cdot)$, т. е. $\hat{u}(\tau-0) \neq \hat{u}(\tau)$, то наряду с $\hat{u}(\cdot)$ рассмотрим управление

$$\tilde{u}(t) = \begin{cases} \hat{u}(t), & t \geqslant \tau, \\ \hat{u}(\tau), & t < \tau. \end{cases}$$

Поскольку $\tilde{u}(\cdot)$ непрерывно в некоторой окрестности точки τ, решение $\tilde{X}(t, t_0, x_0)$ задачи Коши

$$\dot{x} = \varphi(t, x, \tilde{u}(t)), \quad x(t_0) = x_0, \tag{11}$$

также определено и непрерывно дифференцируемо в некоторой окрестности точки $(\tau, \tau, \hat{x}(\tau))$. Если же τ — точка непрерывности $\hat{u}(\cdot)$, то само $X(t, t_0, x_0)$ обладает тем же свойством, и в следующих далее формулах можно считать $\tilde{X} \equiv X$.

Дополним набор $\tau = (\tau_1, \ldots, \tau_N)$ точками $\tau_0 = \hat{t}_0$ и $\tau_{N+1} = \hat{t}_1$, так что $\hat{t}_0 = \tau_0 < \tau_1 \leqslant \ldots \leqslant \tau_N < \tau_{N+1} = \hat{t}_1$, и рассмотрим суперпозицию отображений

$$\tilde{\Xi}(\xi, \bar{\alpha}, \bar{\tau}, \bar{v}) = P \circ X_N \circ Z_N \circ \ldots \circ X_k \circ Z_k \circ X_{k-1} \circ \ldots \circ Z_1 \circ X_0, \tag{12}$$

где $P(\xi, \bar{\alpha}) = \xi$

$$X_k(\xi, \bar{\alpha}) = (X(\tau_{k+1}, \tau_k, \xi), \bar{\alpha}) \tag{13}$$

и

$$Z_k(\xi, \bar{\alpha}) = (\tilde{X}(\tau_k, \tau_k + k\alpha + \alpha_k, Y_{v_k}(\tau_k + k\alpha + \alpha_k, \tau_k + k\alpha, \tilde{X}(\tau_k + k\alpha, \tau_k, \xi))), \bar{\alpha}). \tag{14}$$

Поскольку X_k и Z_k непрерывно дифференцируемы в окрестности точки $p_k = (\hat{x}(\tau_k), \bar{0})$, причем $X_k(p_k) = p_{k+1}$, $Z_k(p_k) = p_k$, а P линейно, суперпозиция (12) определена и непрерывно дифференцируема в некоторой окрестности точки $p_0 = (\hat{x}(\hat{t}_0), \bar{0}) = (\hat{x}_0, 0)$.

Лемма 2. *Если все $\alpha_k > 0$ и достаточно малы, то*

$$\tilde{\Xi}(\xi, \bar{\alpha}, \bar{\tau}, \bar{v}) = \Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v}) = x(\hat{t}_1, \hat{t}_0, \xi, \bar{\alpha}, \bar{\tau}, \bar{v}). \tag{15}$$

Доказательство. Обозначим для краткости $x(t) = x(t, \hat{t}_0, \xi, \bar{\alpha}, \bar{\tau}, \bar{v})$. Если все $\alpha_k > 0$ и достаточно малы, то полуинтервалы Δ_i, входящие в определение (3), попарно не пересекаются и расположены между $\hat{t}_0$ и $\hat{t}_1$ в порядке возрастания номеров.

Пусть $s_i = \tau_i + i\alpha = \tau_i + i\sum\limits_{j=1}^{N} \alpha_j$ — левый конец полуинтервала Δ_i, $s_0 = \tau_0 = \hat{t}_0$, $s_{N+1} = \tau_{N+1} = \hat{t}_1$. Докажем по индукции, что

$$x(s_k) = X(s_k, \tau_k, \xi_k), \tag{16}$$

где $\xi_0 = \xi$ и

$$\xi_k = P \circ X_{k-1} \circ Z_{k-1} \circ X_{k-2} \circ \ldots \circ Z_1 \circ X_0(\xi, \overline{\alpha}), \quad k \geqslant 1; \tag{17}$$

в частности, $\xi_1 = P \circ X_0(\xi, \overline{\alpha}) = X(\tau_1, \tau_0, \xi)$. Поскольку при $\hat{t}_0 = \tau_0 \leqslant t < s_1$ дифференциальные уравнения в (7) и (1) совпадают и $x(\tau_0) = \xi = X(\tau_0, \tau_0, \xi_0)$, то в этом полуинтервале $x(t) \equiv X(t, \tau_0, \xi_0)$, и, переходя к пределу при $t \to s_1$ и учитывая тождество $X(s_1, \tau_0, \xi_0) = X(s_1, \tau_1, X(\tau_1, \tau_0, \xi_0))$, получаем (16) при $k = 1$.

Предположим, что (16) верно для некоторого k. Чтобы перейти по решению $x(\cdot)$ от s_k до s_{k+1}, надо в силу (3) проделать следующее. От s_k до $s_k + \alpha_k$ решаем уравнение (10) с начальным условием $y_0 = x(s_k)$, в результате чего получаем $x(s_k + \alpha_k) = Y_{v_k}(s_k + \alpha_k, s_k, x(s_k))$. Далее от $s_k + \alpha_k$, т. е. от конца Δ_k, до s_{k+1} — начала Δ_{k+1}, решаем уравнение (7) с начальным условием $x_0 = x(s_k + \alpha_k)$ и получаем равенство

$$\begin{aligned} x(s_{k+1}) = X(s_{k+1}, s_k + \alpha_k, x(s_k + \alpha_k)) = \\ = X(s_{k+1}, s_k + \alpha_k, Y_{v_k}(s_k + \alpha_k, s_k, x(s_k)). \end{aligned}$$

Согласно (13) п. 2.5 5 справедливо тождество

$$X(t, \tau, X(\tau, t_0, x_0)) \equiv X(t, t_0, x_0).$$

Воспользовавшись им и равенством (16), верным по индукционной гипотезе, получаем последовательно

$$\begin{aligned} x(s_{k+1}) = X(s_{k+1}, \tau_{k+1}, X(\tau_{k+1}, \tau_k, X(\tau_k, s_k + \alpha_k, x(s_k + \alpha_k)))) = \\ = X(s_{k+1}, \tau_{k+1}, X(\tau_{k+1}, \tau_k, X(\tau_k, s_k + \alpha_k, \\ Y_{v_k}(s_k + \alpha_k, s_k, X(s_k, \tau_k, \xi_k)))))). \end{aligned} \tag{18}$$

Теперь заметим, что если все $\alpha_k > 0$, то $\tau_k < s_k < s_k + \alpha_k$, а так как при $t \geqslant \tau_k$ уравнения (7) и (11) совпадают, то $X(\tau_k, s_k + \alpha_k, \xi) = \tilde{X}(\tau_k, s_k + \alpha_k, \xi)$ и $X(s_k, \tau_k, \xi_k) = \tilde{X}(s_k, \tau_k, \xi_k)$. Следовательно, используя сначала определения (13) и (14), а затем (17), мы получаем

$$X(\tau_{k+1}, \tau_k, X(\tau_k, s_k+\alpha_k, Y_{v_k}(s_k+\alpha_k, s_k, X(s_k, \tau_k, \xi_k)))) = \\ = P\circ X_k\circ Z_k(\xi_k, \bar{\alpha}) = P\circ X_k\circ Z_k\circ X_{k-1}\circ\ldots\circ X_0(\xi, \bar{\alpha}) = \xi_{k+1}.$$

Поэтому (18) совпадает с (16), в котором k заменено на $k+1$, и, таким образом, (16), верно при всех $k \geqslant 1$. Остается заметить, что при $k=N+1$ (16) превращается в равенство

$$x(\hat{t}_1) = X(\hat{t}_1, \hat{t}_1, \xi_{N+1}) = \xi_{N+1} = \\ = P\circ X_N\circ Z_N\circ\ldots\circ X_0(\xi, \bar{\alpha}) = \tilde{\Xi}(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})$$

(последнее согласно (12)) и, так как по определению $x(\hat{t}_1) = \Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})$, (15) доказано. ■

Как уже было сказано, суперпозиция (12) непрерывно дифференцируема в некоторой окрестности точки $(\hat{x}_0, \bar{0})$. Следовательно, $\Xi(\xi, \bar{\alpha}, \bar{\tau}, \bar{v})$ (которая раньше была определена только при $\alpha_j > 0$) допускает непрерывно дифференцируемое продолжение на эту окрестность. Согласно п. Б) отсюда вытекает, что $x(t_1, t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ допускает (при фиксированных $\bar{\tau}$, $\bar{v}$) непрерывно дифференцируемое продолжение на окрестность точки $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$.

Остается доказать лишь формулы (10)—(13) п. 4.2.3 для частных производных этой функции в точке $(\hat{t}_1, \hat{t}_0, \hat{x}_0, \bar{0})$.

Г) Прежде всего заметим, что

$$x(t, t_0, x_0, \bar{0}, \bar{\tau}, \bar{v}) \equiv X(t, t_0, x_0), \tag{19}$$

$$x(t, t_0, \hat{x}(t_0), \bar{0}, \bar{\tau}, \bar{v}) \equiv X(t, t_0, \hat{x}(t_0)) = \hat{x}(t), \tag{20}$$

$$x(\hat{t}_1, \hat{t}_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}) \equiv \Xi(x_0, \bar{\alpha}, \bar{\tau}, \bar{v}). \tag{21}$$

Согласно (19) и (21) частную производную по x_0 мы можем найти, применяя теорему п. 2.5.6

$$\frac{\partial\hat{x}}{\partial x_0} = \frac{\partial x}{\partial x_0}(\hat{t}_1, \hat{t}_0, x_0, \bar{0}, \bar{\tau}, \bar{v})\Big|_{x_0=\hat{x}(\hat{t}_0)} = \frac{\partial\Xi(x_0, \bar{0}, \bar{\tau}, \bar{v})}{\partial x_0}\Big|_{x_0=\hat{x}(\hat{t}_0)} = \\ = \frac{\partial X(\hat{t}_1, \hat{t}_0, x_0)}{\partial x_0}\Big|_{x_0=\hat{x}(\hat{t}_0)} = \Omega(\hat{t}_1, \hat{t}_0), \tag{22}$$

где $\Omega(t, \tau)$—фундаментальная матрица решений уравнения в вариациях

$$\dot{z} = \varphi_x(t, \hat{x}(t), \hat{u}(t))\, z$$

(в формуле (2) п. 2.5.6 надо положить $F(t, x) = \varphi(t, x, \hat{u}(t))$ и учесть (20)). Этим доказано (12) п. 4.2.3.

Теперь воспользуемся формулой (9), заметив, что, как было сказано в п. Б), в окрестности точек $(\hat{t}_k, \hat{t}_k, \hat{x}(\hat{t}_k))$, $k=0, 1$, производные функции $X(t, t_0, x_0)$ можно вычислять, применяя классическую теорему о дифференцируемости по начальным данным (формулы (1)—(4) п. 2.5.7, где снова $F(t, x)=\varphi(t, x, \hat{u}(t))$. Дифференцируя (9) и учитывая уже найденную формулу (22), а также, что, согласно (20) и (21),

$$\Xi(X(\hat{t}_0, \hat{t}_0, \hat{x}(\hat{t}_0)), \overline{0}, \overline{\tau}, \overline{v})=\Xi(\hat{x}(\hat{t}_0), \overline{0}, \overline{\tau}, \overline{v})=$$
$$=x(\hat{t}_1, \hat{t}_0, \hat{x}(\hat{t}_0), \overline{0}, \overline{\tau}, \overline{v})=\hat{x}(\hat{t}_1),$$

имеем

$$\partial\hat{x}/\partial t_1=\partial X(\hat{t}_1, \hat{t}_1, \hat{x}(\hat{t}_1))/\partial t_1=\varphi(\hat{t}_1, \hat{x}(\hat{t}_1), \hat{u}(\hat{t}_1)),$$
$$\partial\hat{x}/\partial t_0=\partial\Xi(\xi, \overline{0}, \overline{\tau}, \overline{v})/\partial\xi\,|_{\xi=\hat{x}(\hat{t}_0)}\,\partial X(\hat{t}_0, t_0, \hat{x}(\hat{t}_0)/\partial t_0\,|_{t_0=\hat{t}_0}=$$
$$=\Omega(\hat{t}_1, \hat{t}_0)[-\varphi(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{u}(\hat{t}_0))].$$

Этим доказаны (10) и (11) п. 4.2.3.

Пусть теперь $\alpha_i=0$, $i\neq k$ и $\alpha_k>0$. Обозначим для краткости

$$x(t, \alpha_k)=x(t, \hat{t}_0, \hat{x}(\hat{t}_0), (0, \ldots, 0, \alpha_k, \ldots, 0), \overline{\tau}, \overline{v}).$$

При таком выборе $\overline{\alpha}$, согласно (3), в полуинтервале $\hat{t}_0\leqslant t<\tau_k+k\alpha_k$ дифференциальные уравнения (1) и (7) совпадают и $x(\hat{t}_0, \alpha_k)=\hat{x}(t_0)$, а потому в этом полуинтервале $x(t, \alpha_k)\equiv\hat{x}(t)$ и, переходя к пределу, получаем

$$x(\tau_k+k\alpha_k, \alpha_k)=\hat{x}(\tau_k+k\alpha_k). \tag{23}$$

Далее, при $\tau_k+k\alpha_k\leqslant t<\tau_k+k\alpha_k+\alpha_k$ мы решаем дифференциальное уравнение (10) с начальным условием (23) и с $v=v_k$, откуда

$$x(\tau_k+k\alpha_k+\alpha_k)=$$
$$=Y_{v_k}(\tau_k+k\alpha_k+\alpha_k, \tau_k+k\alpha_k, \hat{x}(\tau_k+k\alpha_k)). \tag{24}$$

В оставшемся промежутке $\tau_k+(k+1)\alpha_k\leqslant t\leqslant\hat{t}_1$ мы снова возвращаемся к уравнению (7), теперь уже с начальным условием (24). Этот промежуток мы разобьем на два, фиксировав s так, что в интервале (τ_k, s) не было точек разрыва $\hat{u}(\cdot)$. Величину $\alpha_k>0$ будем считать столь малой, чтобы было $\tau_k < \tau_k + (k+1)\alpha_k < s$. Имеем

$$x(\hat{t}_1, \alpha_k) = X(\hat{t}_1, s, x(s, \alpha_k)) =$$
$$= X(\hat{t}_1, s, X(s, \tau_k + (k+1)\alpha_k, x(\tau_k + (k+1)\alpha_k, \alpha_k))) =$$
$$= X(\hat{t}_1, s, X(s, \tau_k + (k+1)\alpha_k, Y_{v_k}(\tau_k + (k+1)\alpha_k, \tau_k + k\alpha_k, \hat{x}(\tau_k + k\alpha_k))).$$

При дифференцировании этой формулы мы должны учесть следующие равенства:

$$\frac{\partial X}{\partial x_0}(\hat{t}_1, s, x_0)\Big|_{x_0 = x(s,\,0) = \hat{x}(s)} = \Omega(\hat{t}_1, s)$$

((1) п. 2.5.6; $\hat{t}_1$ и s фиксированы);

$$\frac{\partial X}{\partial t_0}(s, t_0, x(\tau_k, 0))\Big|_{t_0 = \tau_k} = -\Omega(s, \tau_k)\,\varphi(\tau_k, \hat{x}(\tau_k), \hat{u}(\tau_k)),$$
$$\frac{\partial X}{\partial x_0}(s, \tau_k, x_0)\Big|_{x_0 = x(\tau_k,\,0) = \hat{x}(\tau_k)} = \Omega(s, \tau_k)$$

((5) и (6) п. 2.5.7; классическая теорема применима, поскольку на отрезке $[\tau_k, s]$ управление $\hat{u}(\cdot)$ непрерывно и $\tau_k < \tau_k + (k+1)\alpha_k < s$);

$$\frac{\partial Y_{v_k}}{\partial t}(t, \tau_k, \hat{x}(\tau_k))\Big|_{t = \tau_k} = \varphi(\tau_k, \hat{x}(\tau_k), v_k),$$
$$\frac{\partial Y_{v_k}}{\partial t_0}(\tau_k, t_0, \hat{x}(\tau_k))\Big|_{t_0 = \tau_k} = -\varphi(\tau_k, \hat{x}(\tau_k), v_k),$$
$$\frac{\partial Y_{v_k}}{\partial y_0}(\tau_k, \tau_k, y_0) = E$$

(формулы (4)—(6) п. 2.5.7, где следует положить $F(t, x) = \varphi(t, x, v_k)$ и учесть, что всегда $\Omega(t, t) = E$); и, наконец, $\dot{\hat{x}}(\tau_k) = \varphi(\tau_k, \hat{x}(\tau_k), \hat{u}(\tau_k))$ из (7).

Таким образом,

$$\frac{\partial \hat{x}}{\partial \alpha_k} = \frac{\partial x}{\partial \alpha_k}(\hat{t}_1, \alpha_k)\Big|_{\alpha_k = 0} =$$
$$= \frac{\partial X}{\partial x_0}(\hat{t}_1, s, x_0)\Big|_{x_0 = \hat{x}(s)} \Big\{ \frac{\partial X(s, \tau_k, \hat{x}(\tau_k))}{\partial t_0}(k+1) +$$
$$+ \frac{\partial X}{\partial x_0}(s, \tau_k, \hat{x}(\tau_k)) \Big[\frac{\partial Y}{\partial t}(\tau_k, \tau_k, \hat{x}(\tau_k))(k+1) +$$
$$+ \frac{\partial Y}{\partial t_0}(\tau_k, \tau_k, \hat{x}(\tau_k))k + \frac{\partial Y}{\partial y_0}(\tau_k, \tau_k, \hat{x}(\tau_k))\,\dot{\hat{x}}(\tau_k)k\Big]\Big\} =$$
$$= \Omega(\hat{t}_1, s)\{-\Omega(s, \tau_k)\,\varphi(\tau_k, \hat{x}(\tau_k), \hat{u}(\tau_k))(k+1) +$$
$$+ \Omega(s, \tau_k)[\varphi(\tau_k, \hat{x}(\tau_k), v_k)(k+1) - \varphi(\tau_k, \hat{x}(\tau_k), v_k)k +$$
$$+ \varphi(\tau_k, \hat{x}(\tau_k), \hat{u}(\tau_k)\ k]\} =$$
$$= \Omega(\hat{t}_1, \tau_k)[\varphi(\tau_k, \hat{x}(\tau_k), v_k) - \varphi(\tau_k, \hat{x}(\tau_k), \hat{u}(\tau_k))],$$

чем доказано (13) п. 4.2.3 (равенство $\Omega(\hat{t}_1, s)\Omega(s, \tau_k)=\Omega(\hat{t}_1, \tau_k)$ имеет место по основному свойству фундаментальной матрицы (11) п. 2.5.4).

4.2.7. Доказательство леммы об интегральных функционалах. Пусть $\xi=(\xi_0, \xi_1, \ldots, \xi_m)$, $f=(f_0, f_1, \ldots, f_m)$, $F=(F_0, F_1, \ldots, F_m)$ (функции F_i определены в формулировке леммы в п. 4.2.3), $u(t, \bar{\alpha}, \bar{\tau}, \bar{v})$ и $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$ имеют тот же смысл, что и в пп. 4.2.3 и 4.2.6.

Рассмотрим при $\alpha_k > 0$ задачу Коши для системы $(n+m+1)$-го порядка

$$\dot{\tilde{x}}=\begin{pmatrix}\dot{x}\\ \dot{\xi}\end{pmatrix}=\tilde{\varphi}(t, x, \xi, u(t, \bar{\alpha}, \bar{\tau}, \bar{v}))=\begin{pmatrix}\varphi(t, x, u(t, \bar{\alpha}, \bar{\tau}, \bar{v}))\\ f(t, x, u(t, \bar{\alpha}, \bar{\tau}, \bar{v}))\end{pmatrix}, \tag{1}$$

$$x(t_0)=x_0, \quad \xi(t_0)=0.$$

Уравнения для x здесь не содержат ξ и совпадают с (1), (2) п. 4.2.6. Поэтому их решением является $x(t; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v})$, а тогда, как легко видеть,

$$F(t_1, t_0, x_0, \bar{\alpha}) \equiv \xi(t_1; t_0, x_0, \bar{\alpha}, \bar{\tau}, \bar{v}).$$

Следовательно, возможность продолжения этой функции на неположительные α_k и непрерывная дифференцируемость в некоторой окрестности $(\hat{t}_1, \hat{t}_0, \hat{x}_0, 0)$ вытекают из леммы о пакете иголок, примененной к системе (1), а значения производных находятся по формулам, аналогичным (10)—(13) п. 4.2.3:

$$\begin{gathered}\frac{\partial \hat{\tilde{x}}}{\partial t_1}=\hat{\tilde{\varphi}}(\hat{t}_1), \qquad \frac{\partial \hat{\tilde{x}}}{\partial t_0}=-\tilde{\Omega}(\hat{t}_1, \hat{t}_0)\hat{\tilde{\varphi}}(\hat{t}_0),\\ \frac{\partial \hat{\tilde{x}}}{\partial \tilde{x}_0}=\tilde{\Omega}(\hat{t}_1, \hat{t}_0), \quad \frac{\partial \hat{\tilde{x}}}{\partial \alpha_k}=\tilde{\Omega}(\hat{t}_1, \tau_k)\Delta\hat{\tilde{\varphi}}(\tau_k, v_k),\end{gathered} \tag{2}$$

где

$$\tilde{\Omega}(t, \tau)=\begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix}$$

— фундаментальная матрица решений системы уравнений в вариациях

$$\dot{\tilde{z}}=\begin{pmatrix}\dot{z}\\ \dot{\zeta}\end{pmatrix}=\hat{\tilde{\varphi}}_{\tilde{x}}(t)\begin{pmatrix}z\\ \zeta\end{pmatrix} \tag{3}$$

(Ω_{ij} — прямоугольные блоки в матрице $\tilde{\Omega}$ размерами $n\times n$, $n\times m$, $m\times n$ и $m\times m$ соответственно). Учитывая, что $\tilde{\varphi}$ не зависит от ξ, мы получаем из (3), что $\tilde{\Omega}$ является решением задачи Коши для матричного дифференциального уравнения

$$\frac{d}{dt}\begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix}=\begin{pmatrix}\hat{\varphi}_x(t) & 0\\ \hat{f}_x(t) & 0\end{pmatrix}\begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix},$$
$$\begin{pmatrix}\Omega_{11}(\tau,\tau) & \Omega_{12}(\tau,\tau)\\ \Omega_{21}(\tau,\tau) & \Omega_{22}(\tau,\tau)\end{pmatrix}=\begin{pmatrix}E_n & 0\\ 0 & E_m\end{pmatrix}. \tag{4}$$

Решая поочередно четыре матричных уравнения, составляющие (4), имеем

$$d\Omega_{11}(t,\tau)/dt=\hat{\varphi}_x(t)\,\Omega_{11}(t,\tau),\quad \Omega_{11}(\tau,\tau)=E_n,$$

откуда $\Omega_{11}(t,\tau)=\Omega(t,\tau)$ — фундаментальная матрица системы (3) п. 4.2.3. Уравнениям

$$d\Omega_{12}(t,\tau)/dt=\hat{\varphi}_x(t)\,\Omega_{12}(t,\tau),\quad \Omega_{12}(\tau,\tau)=0,$$

удовлетворяет $\Omega_{12}(t,\tau)\equiv 0$, и по теореме единственности другого решения не может быть. Далее

$$d\Omega_{21}(t,\tau)/dt=\hat{f}_x(t)\,\Omega_{11}=\hat{f}_x(t)\,\Omega(t,\tau),\quad \Omega_{21}(t,\tau)=0,$$

откуда

$$\Omega_{21}(t,\tau)=\int_{\tau}^{t}\hat{f}_x(s)\,\Omega(s,\tau)\,ds.$$

В частности,

$$-\Omega_{21}(\hat{t}_1,\tau)=p_0(\tau)=(p_{00}(\tau),\ \ldots,\ p_{0m}(\tau))$$

является решением задачи Коши (18) п. 4.2.3 (см. (14) в п. 2.5.4; можно убедиться в этом также и непосредственно, дифференцируя по τ и учитывая свойства фундаментальной матрицы, описанные в той же теореме п. 2.5.4). Наконец,

$$d\Omega_{22}(t,\tau)/dt=\hat{f}_x(t)\,\Omega_{12}(t,\tau)=0,\quad \Omega_{22}(\tau,\tau)=E_m,$$

откуда $\Omega_{22}(t,\tau)\equiv E_m$. Таким образом,

$$\tilde{\Omega}(\hat{t}_1,\tau)=\begin{pmatrix}\Omega(\hat{t}_1,\tau) & 0\\ -p_0(\tau) & E_m\end{pmatrix}.$$

Подставляя в (2) и отделяя то, что относится к

$\xi(t_1; t_0, x_0, \overline{\alpha}, \overline{\tau}, \overline{v}) = F(t_1, t_0, \overline{x}_0, \overline{\alpha})$, мы получаем

$$\partial\hat{F}/\partial t_1 = \partial\hat{\xi}/\partial t_1 = \hat{f}(\hat{t}_1),$$

$$\partial\hat{F}/\partial t_0 = \partial\hat{\xi}/\partial t_0 = -\Omega_{21}(\hat{t}_1, \hat{t}_0)\hat{\varphi}(\hat{t}_0) - \Omega_{22}(\hat{t}_1, \hat{t}_0)\hat{f}(\hat{t}_0) = = p_0(\hat{t}_0)\hat{\varphi}(\hat{t}_0) - \hat{f}(\hat{t}_0),$$

$$\partial\hat{F}/\partial x_0 = \partial\hat{\xi}/\partial x_0 = \Omega_{21}(\hat{t}_1, \hat{t}_0) = -p_0(\hat{t}_0),$$

$$\partial\hat{F}/\partial\alpha_k = \partial\hat{\xi}/\partial\alpha_k = \Omega_{21}(\hat{t}_1, \tau_k)\Delta\hat{\varphi}(\tau_k, v_k) + \Omega_{22}(\hat{t}_1, \tau_k)\Delta\hat{f}(\tau_k, v_k) = -p_0(\tau_k)\Delta\hat{\varphi}(\tau_k, v_k) + \Delta\hat{f}(\tau_k, v_k),$$

что совпадает с (14)—(17) п. 4.2.3. ■

§ 4.3*. Задачи оптимального управления, линейные по фазовым переменным

Этот параграф посвящен одному специальному классу задач оптимального управления. Этот класс достаточно важен и с прикладной точки зрения, но для нас он интересен не только этим. На примере задач с линейной структурой по фазовым переменным можно наиболее отчетливо продемонстрировать одну из самых принципиальных идей всей теории. Речь идет о «скрытой выпуклости», всегда присутствующей в задачах оптимального управления. Именно она в конечном счете дает возможность записать необходимое условие в виде «принципа максимума», т. е. в форме, характерной для задач выпуклого программирования. И если в предыдущем параграфе мы старались подчеркнуть связь задач оптимального управления с общей теорией гладких экстремальных задач, то здесь на первый план выдвигается их связь с выпуклым анализом.

Отметим еще, что (как это и характерно для задач выпуклого программирования) необходимые условия экстремума почти смыкаются здесь с достаточными. Наконец, и это тоже важно с точки зрения выявления скрытой выпуклости, мы будем в этом параграфе рассматривать измеримые управления, а не только кусочно-непрерывные.

4.3.1. Редукция задачи оптимального управления, линейной по фазовым переменным, к задаче ляпуновского типа. Пусть $\Delta = [t_0, t_1]$ — фиксированный отрезок числовой прямой, $a_i: \Delta \to \mathbf{R}^{n*}$, $i = 0, 1, \ldots, m$, и $A: \Delta \to \mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$ — интегрируемые векторные и матричная функции; $\mathfrak{U}$ — топологическое пространство, $f_i: \Delta \times \mathfrak{U} \to \mathbf{R}$,

$i=0, 1, \ldots, m$, и $F(\cdot, \cdot)$: $\Delta\times\mathfrak{U}\to\mathbf{R}^n$ — непрерывные функции; γ_{0i}, γ_{1i}, $i=0, 1, \ldots, m$, — элементы $\mathbf{R}^{n*}$, c_i, $i=1, 2, \ldots, m$, — числа.

Экстремальная задача

$$J_0(x(\cdot), u(\cdot))=\int_\Delta (a_0(t)x(t)+f_0(t, u(t)))\,dt+$$
$$+\gamma_{00}x(t_0)+\gamma_{10}x(t_1)\to\inf,$$
$$\dot x=A(t)x+F(t, u(t)),\quad u(t)\in\mathfrak{U}, \tag{1}$$
$$J_i(x(\cdot), u(\cdot))=\int_\Delta (a_i(t)x(t)+f_i(t, u(t)))\,dt+$$
$$+\gamma_{0i}x(t_0)+\gamma_{1i}x(t_1)\leqslant c_i,$$

в которую величины, связанные с фазовой траекторией x, входят линейно, называется *задачей оптимального управления, линейной по фазовым переменным*. Она, разумеется, является частным случаем общей задачи оптимального управления из § 4.2. Однако здесь мы несколько расширим совокупность допустимых процессов.

Обозначим через $\mathcal{U}$ множество измеримых (в смысле определения п. 4.2.6) отображений u: $\Delta\to\mathfrak{U}$ таких, что функции $t\mapsto f_i(t, u(t))$ и $t\mapsto F(t, u(t))$ интегрируемы. Пара $(x(\cdot), u(\cdot))$ называется *допустимым процессом*, если $x(\cdot)$: $\Delta\to\mathbf{R}^n$ абсолютно непрерывная вектор-функция (см. п. 2.1.8), $u(\cdot)\in\mathcal{U}$,

$$\dot x(t)=A(t)x(t)+F(t, u(t))$$

почти всюду и выполняются неравенства $J_i(x(\cdot), u(\cdot))\leqslant c_i$. Допустимый процесс $(\hat x(\cdot), \hat u(\cdot))$ называется *оптимальным*, если существует такое $\varepsilon>0$, что для любого допустимого процесса $(x(\cdot), u(\cdot))$, удовлетворяющего условию $\|x(\cdot)-\hat x(\cdot)\|_{C(\Delta, \mathbf{R}^n)}<\varepsilon$ выполнено неравенство

$$J_0(x(\cdot), u(\cdot))\geqslant J_0(\hat x(\cdot), \hat u(\cdot)).$$

Мы покажем сейчас, что линейная структура позволяет преобразовать задачу (1) к такому виду, где $x(\cdot)$ и $u(\cdot)$ в некотором смысле разделены (и, в частности, отсутствует дифференциальная связь).

Пусть $\Omega(\cdot, \cdot)$ — фундаментальная матрица (п. 2.5.4) решений однородной линейной системы

$$\dot x=A(t)x. \tag{2}$$

В соответствии с формулой (13) п. 2.5.4

$$x(t)=\Omega(t, t_0)x(t_0)+\int_{t_0}^{t}\Omega(t, \tau)F(\tau, u(\tau))\,d\tau. \qquad (\tilde{3})$$

Подставляя (3), преобразуем функционалы задачи (1): $J_i(x(\cdot), u(\cdot))=$

$$=\int_{t_0}^{t_1}\left\{a_i(t)\left[\Omega(t, t_0)x(t_0)+\int_{t_0}^{t}\Omega(t, \tau)F(\tau, u(\tau))\,d\tau\right]\right\}dt+$$

$$+\int_{t_0}^{t_1}f_i(t, u(t))\,dt+\gamma_{0i}x(t_0)+$$

$$+\gamma_{1i}\left[\Omega(t_1, t_0)x(t_0)+\int_{t_0}^{t_1}\Omega(t_1, \tau)F(\tau, u(\tau))\,d\tau\right]=$$

$$=\int_{t_0}^{t_1}\int_{t_0}^{t}a_i(t)\Omega(t, \tau)F(\tau, u(\tau))\,d\tau\,dt+\int_{t_0}^{t_1}f_i(\tau, u(\tau))\,d\tau+$$

$$+\left\{\int_{t_0}^{t_1}a_i(t)\Omega(t, t_0)\,dt+\gamma_{0i}+\gamma_{1i}\Omega(t_1, t_0)\right\}x(t_0)+$$

$$+\gamma_{1i}\int_{t_0}^{t_1}\Omega(t_1, \tau)F(\tau, u(\tau))\,d\tau=$$

$$=\int_{t_0}^{t_1}\left\{\int_{\tau}^{t_1}a_i(t)\Omega(t, \tau)\,dt\,F(\tau, u(\tau))+\right.$$

$$\left.+f_i(\tau, u(\tau))+\gamma_{1i}\Omega(t_1, \tau)F(\tau, u(\tau))\right\}d\tau+$$

$$+\left\{\int_{t_0}^{t_1}a_i(t)\Omega(t, t_0)\,dt+\gamma_{0i}+\gamma_{1i}\Omega(t_1, t_0)\right\}x(t_0)=$$

$$=-\int_{t_0}^{t_1}G_i(\tau, u(\tau))\,d\tau+\beta_i x(t_0), \qquad (4)$$

где

$$G_i(\tau, u)=p_i(\tau)F(\tau, u)-f_i(\tau, u), \qquad (5)$$

$$\beta_i=\gamma_{0i}-p_i(t_0), \qquad (6)$$

$$p_i(\tau)=-\gamma_{1i}\Omega(t_1, \tau)-\int_{\tau}^{t_1}a_i(t)\Omega(t, \tau)\,dt, \qquad (7)$$

$$i=0, 1, \ldots, m.$$

Теперь задача (1) приобретает вид

$$J_0(x(\cdot),\ u(\cdot)) = -\int_\Delta G_0(t, u(t))\,dt + \beta_0 x(t_0) \to \inf,$$
$$J_i(x(\cdot),\ u(\cdot)) = -\int_\Delta G_i(t, u(t))\,dt + \beta_i x(t_0) \leqslant c_i. \tag{8}$$

Такую задачу мы будем называть задачей ляпуновского типа или просто *ляпуновской задачей*.

4.3.2. Теорема Ляпунова. В установлении связи между задачами оптимального управления и выпуклого программирования теорема Ляпунова играет ключевую роль. Она будет сформулирована здесь для случая векторной функции p: $\mathbf{R} \to \mathbf{R}^n$ и меры Лебега m на прямой $\mathbf{R}$. Незначительное усовершенствование доказательств позволяет распространить эту теорему на случай произвольного пространства с заданной в нем σ-алгеброй $\mathfrak{S}$ и произвольной непрерывной (неатомической) конечной векторной мерой μ: $\mathfrak{S} \to \mathbf{R}^n$. В приведенной ниже формулировке $\mathfrak{S}$ — это σ-алгебра измеримых по Лебегу подмножеств в $\mathbf{R}$, а $\mu(A) = \int_A p(t)\,dt$.

В отличие от других мест нашей книги, здесь и в следующем пункте нам придется использовать некоторые факты из функционального анализа и теории меры, которые, несмотря на их большое принципиальное значение, обычно не входят в основные университетские курсы. Начнем с определений. Термины «измеримость», «интегрируемость», «почти всюду» и т. п. далее понимаются в смысле Лебега (для измеримости удобно пользоваться определением, приведенным в п. 4.2.6). Две функции называются *эквивалентными*, если они совпадают почти всюду.

Определение 1. Пусть Δ — произвольный промежуток числовой прямой. Пространствами $L_1(\Delta)$ и $L_\infty(\Delta)$ называются нормированные пространства, элементами которых являются классы эквивалентных между собой функций x: $\Delta \to \mathbf{R}$, обладающих конечной нормой, которая задается равенствами:

в пространстве $L_1(\Delta)$:

$$\|x(\cdot)\| = \int_\Delta |x(t)|\,dt; \tag{1}$$

в пространстве $L_\infty(\Delta)$:

$$\|x(\cdot)\| = \sup_{\Delta} \operatorname{vrai} |x(t)| = \inf_{N,\, m(N)=0} \sup_{t \in \Delta \setminus N} |x(t)|. \qquad (2)$$

Предложение 1. *Пространства $L_1(\Delta)$ и $L_\infty(\Delta)$ банаховы, причем $L_1(\Delta)$ сепарабельно. Пространство $L_\infty(\Delta)$ изометрически изоморфно сопряженному пространству $L_1(\Delta)^*$, так что замкнутые шары в нем $*$-слабо компактны.*

Доказательство этих утверждений мы заменим следующими ссылками: полнота и сепарабельность пространства $L_1(\Delta)$ — [КФ, гл. VII, § 1] (существование счетного базиса у меры Лебега на бесконечном промежутке сразу следует из его существования на конечных отрезках, поскольку каждый такой промежуток представим в виде объединения счетного числа отрезков); компактность в $*$-слабой топологии замкнутых шаров в пространстве, сопряженном к сепарабельному [КФ, стр. 202], изометрический изоморфизм между $L_1(\Delta)^*$ и $L_\infty(\Delta)$ (а, значит, и полнота последнего) [4, стр. 314].

Определение 2. Пусть X — линейное пространство, $A \subset X$. Точка x множества A называется его *крайней точкой*, если

$$x = \alpha x_1 + (1-\alpha) x_2, \quad 0 < \alpha < 1, \quad x_1, x_2 \in A \Rightarrow x_1 = x_2.$$

Упражнение 1. Докажите, что по крайней мере одна из точек $x_1, \ldots, x_n$ является крайней для их выпуклой оболочки $\operatorname{conv}\{x_1, \ldots, x_N\}$.

Теорема Крейна — Мильмана. *Компактное выпуклое подмножество локально выпуклого линейного топологического пространства X содержит по крайней мере одну крайнюю точку и совпадает с выпуклой замкнутой оболочкой своих крайних точек.*

Доказательство этой теоремы см. [4, стр. 477].

Перейдем теперь к основной теореме этого пункта.

Теорема А. А. Ляпунова. *Если $p\colon \Delta \to \mathbf{R}^n$, $p(\cdot) = (p_1(\cdot), \ldots, p_n(\cdot))$ — интегрируемая векторная функция, то множество*

$$M = \left\{ x \in \mathbf{R}^n \mid x = \int_A p(t)\, dt,\ A \in \mathfrak{S} \right\}$$

является выпуклым компактом в $\mathbf{R}^n$. (Напомним, что $\mathfrak{S}$ — это σ-алгебра множеств, измеримых по Лебегу.)

Доказательство А) Рассмотрим отображение Λ: $L_\infty(\Delta) \longrightarrow \mathbf{R}^n$, определяемое равенством

$$\Lambda(\psi(\cdot)) = \int_\Delta \psi(t)p(t)\,dt = \left(\int_\Delta \psi(t)p_1(t)\,dt, \ldots, \int_\Delta \psi(t)p_n(t)\,dt\right).$$

Оно линейно, и, кроме того, оно непрерывно относительно $*$-слабой топологии в $L_\infty(\Delta)$ (по определению эта топология слабейшая из тех, в которых все отображения $\psi(\cdot) \mapsto \langle\psi(\cdot), p(\cdot)\rangle$, $p(\cdot) \in L_1(\Delta)$ непрерывны).

Множество

$$W = \{\psi(\cdot) \in L_\infty(\Delta) \mid 0 \leqslant \psi(t) \leqslant 1,\ t \in \Delta\}$$

является замкнутым шаром в $L_\infty(\Delta)$ (радиуса 1/2 и с центром в точке $\hat{\psi}(\cdot)$, $\hat{\psi}(t) \equiv 1/2$). Следовательно, оно выпукло и, согласно предложению 1, компактно в $*$-слабой топологии. Поэтому его линейный и непрерывный образ ΛW также является выпуклым и компактным.

Характеристическая функция

$$\chi_A(t) = \begin{cases} 1, & t \in A, \\ 0, & t \notin A \end{cases}$$

любого множества $A \in \mathfrak{S}$, очевидно, принадлежит W, а потому

$$\int_A p(t)\,dt = \int_\Delta \chi_A(t)p(t)\,dt = \Lambda(\chi_A(\cdot)) \in \Lambda W,$$

и, следовательно, $M \subset \Lambda W$. Остается доказать, что $M = \Lambda W$.

Б) Пусть точка $\xi \in \Lambda W$. Ее прообраз $W_\xi = W \cap \Lambda^{-1}(\xi)$ является пересечением выпуклого и компактного в $*$-слабой топологии множества W и замкнутого в той же топологии (поскольку Λ относительно нее непрерывно) выпуклого множества (аффинного многообразия) $\Lambda^{-1}(\xi)$. Следовательно, W_ξ выпукло и компактно и по теореме Крейна — Мильмана имеет крайнюю точку $x(\cdot) \in W_\xi$.

Если мы докажем, что $x(\cdot)$ является характеристической функцией некоторого множества $A \in \mathfrak{S}$, то тогда

$$\xi = \int_\Delta x(t)p(t)\,dt = \int_A p(t)\,dt \in M$$

и теорема будет доказана.

В) Если $0 \leqslant x(t) \leqslant 1$ и $x(\cdot)$ не является характеристической функцией, то для некоторого $\varepsilon > 0$ множество $B_\varepsilon = \{t \mid \varepsilon \leqslant x(t) \leqslant 1-\varepsilon\}$ имеет положительную меру. Функция $\alpha \mapsto m(\alpha) = m[(t_0, \alpha) \cap B_\varepsilon]$ непрерывна и изменяется от $0 = m(t_0)$ до $m(B_\varepsilon) = m(t_1)$ ($m(B)$ — мера Лебега множества B). Задав произвольно N, выберем α_k, $k = 1, \ldots, N-1$ так, чтобы $m(\alpha_k) = k/NB_\varepsilon$. Этим определяется разбиение $B_\varepsilon = B_1 \cup \ldots \cup B_N$ множества B_ε на N попарно непересекающихся подмножеств

$$B_1 = (-\infty, \alpha_1] \cap B_\varepsilon, \ \ldots, \ B_k = (\alpha_{k-1}, \alpha_k] \cap B_\varepsilon, \ \ldots \\ \ldots, \ B_N = (\alpha_{N-1}, \infty) \cap B_\varepsilon$$

положительной меры.

Положим теперь

$$y(t) = \begin{cases} y_k, & t \in B_k, \\ 0, & t \notin B_\varepsilon. \end{cases} \tag{3}$$

При $N > n$ однородная линейная система

$$\int_\Delta y(t)\, p(t)\, dt = \sum_k y_k \int_{B_k} p_i(t)\, dt = 0$$

имеет ненулевое решение $(\hat{y}_1, \ldots, \hat{y}_N)$. Соответствующую функцию (3) обозначим $\hat{y}(\cdot)$. Если $0 < \lambda < \varepsilon \max\{|\hat{y}_k|\}^{-1}$, то

$$|\lambda y(t)| \begin{cases} \leqslant \varepsilon, & t \in B_\varepsilon, \\ = 0, & t \notin B_\varepsilon, \end{cases}$$

и потому $x(\cdot) \pm \lambda \hat{y}(\cdot) \in W$. При этом

$$\Lambda(x(\cdot) \pm \lambda \hat{y}(\cdot)) = \Lambda(x(\cdot)) \pm \lambda \Lambda \hat{y}(\cdot) = \\ = \xi \pm \lambda \int_\Delta \hat{y}(t)\, p(t)\, dt = \xi.$$

Следовательно, $x(\cdot) \pm \lambda \hat{y}(\cdot) \in W_\xi$, и так как $\lambda \neq 0$, то $x(\cdot)$ — не крайняя точка.

Таким образом, $x(\cdot)$ — характеристическая функция и теорема доказана. ■

4.3.3. Принцип Лагранжа для ляпуновских задач. В этом пункте Δ — фиксированный промежуток числовой прямой (конечный или бесконечный), $\mathfrak{U}$ — некоторое топологическое пространство.

Лемма о суперпозиционной измеримости. *Если функция $f\colon \Delta \times \mathfrak{U} \to \mathbf{R}$ непрерывна, а $u\colon \Delta \to \mathfrak{U}$*

измерима (в смысле определения п. 4.2.6), *то и функция* $t \mapsto f(t, u(t))$ *измерима.*

Доказательство. Пусть $A_n \subset \Delta$ выбраны так, что $m\left(\Delta \setminus \bigcup_n A_n\right) = 0$ (m — мера Лебега) и ограничения $u(\cdot)|_{A_n}$ могут быть продолжены до непрерывных функций на $\overline{A}_n$. Тогда и функции $t \mapsto f(t, u(t))$, $t \in A_n$, могут быть продолжены до непрерывных на $\overline{A}_n$. ■

Пусть теперь заданы непрерывные функции f_i: $\Delta \times \mathfrak{U} \to \mathbb{R}$. Обозначим через $\mathcal{U}$ совокупность измеримых отображений u: $\Delta \to \mathfrak{U}$, для которых суперпозиции $t \mapsto f_i(t, u(t))$ не только измеримы, но и интегрируемы на Δ, так что на $\mathcal{U}$ определены функции $u(\cdot) \to \int\limits_\Delta f_i(t, u(t))\,dt$.

С другой стороны, пусть X — линейное пространство, функции g_i: $X \to \overline{\mathbb{R}}$ — выпуклые для $i = 0, 1, \ldots, m'$ и аффинные с конечными значениями для $i = m'+1, \ldots, m$, $A \subset X$ — выпуклое подмножество.

Экстремальную задачу

$$\mathcal{F}_0(u(\cdot)) + g_0(x) = \int\limits_\Delta f_0(t, u(t))\,dt + g_0(x) \to \inf,$$

$$\mathcal{F}_i(u(\cdot)) + g_i(x) = \qquad\qquad (1)$$
$$= \int\limits_\Delta f_i(t, u(t))\,dt + g_i(x) \begin{cases} \leqslant 0, & i = 1, \ldots, m', \\ = 0, & i = m'+1, \ldots, m, \end{cases}$$
$$x \in A, \quad u(\cdot) \in \mathcal{U}$$

мы будем называть *ляпуновской задачей*. В качестве частного случая ляпуновская задача включает в себя стандартную задачу выпуклого программирования, изученную нами в п. 1.3.3 (надо положить $m' = m$ и $f_i = 0$), и мы увидим далее, что теорема Ляпунова позволяет распространить на рассматриваемую здесь более общую ситуацию рассуждения, при помощи которых была доказана теорема Куна—Таккера.

Функция

$$\mathcal{L}(u(\cdot), x, \lambda, \lambda_0) = \sum_{i=0}^{n} \lambda_i (\mathcal{F}_i(u(\cdot)) + g_i(x)) \qquad (2)$$

называется *функцией Лагранжа* задачи (1).

Теорема (принцип Лагранжа для ляпуновской задачи). *Пусть функции* f_i: $\Delta \times \mathcal{U} \to \mathbb{R}$

непрерывны, g_i: $X \longrightarrow \overline{\mathbb{R}}$ *выпуклые для* $i=0, 1, \ldots, m'$ *и аффинные конечные для* $i=m'+1, \ldots, m$, $A \subset X$ *выпукло.*

1. *Если пара* $(\hat{u}(\cdot), \hat{x})$ *является решением задачи* (1), *то найдутся вектор* $\hat{\lambda}=(\hat{\lambda}_1, \ldots, \hat{\lambda}_m) \in \mathbb{R}^{m*}$ *и число* $\hat{\lambda}_0$, *не равные одновременно нулю и такие, что:*

$$\text{а)}\ \min_{u \in U} \sum_{i=0}^{m} \hat{\lambda}_i f_i(t, u) = \sum_{i=0}^{m} \hat{\lambda}_i (t, \hat{u}(t)) \text{ почти всюду}, \tag{3}$$

$$\min_{x \in A} \sum_{i=0}^{m} \hat{\lambda}_i g_i(x) = \sum_{i=0}^{m} \hat{\lambda}_i g_i(\hat{x}) \tag{4}$$

(*принцип минимума*);

$$\text{б)} \qquad \hat{\lambda}_i \geqslant 0, \quad i=0, 1, \ldots, m' \tag{5}$$

(*условие согласования знаков*);

$$\text{в)} \qquad \hat{\lambda}_j \{\mathcal{F}_j(\hat{u}(\cdot)) + g_j(\hat{x})\} = 0, \quad j=1, \ldots, m \tag{6}$$

(*условия дополняющей нежесткости*).

2. *Если* $(\hat{u}(\cdot), \hat{x}) \in \mathcal{U} \times A$ *и существуют такие* $\hat{\lambda} \in \mathbb{R}^{m*}$ *и* $\hat{\lambda}_0 > 0$, *что выполняются условия* (3)—(6), *то* $(\hat{u}(\cdot), \hat{x})$ — *решение задачи* (1).

Доказательство. А) Лемма о выпуклости образа.

Образ

$$\operatorname{im} \mathcal{F} = \{\xi = (\xi_0, \ldots, \xi_m) \mid \xi_i = \mathcal{F}_i(u(\cdot)),\ u(\cdot) \in \mathcal{U}\}$$

отображения $u(\cdot) \mapsto \mathcal{F}(u(\cdot)) = (\mathcal{F}_0(u(\cdot)), \ldots, \mathcal{F}_m(u(\cdot)))$ *является выпуклым множеством в* $\mathbb{R}^{m+1}$.

Доказательство. Пусть

$$\xi^k = (\xi_0^k, \xi_1^k, \ldots, \xi_m^k) = \mathcal{F}(u^{(k)}(\cdot)), \quad k=1, 2.$$

Функция $p(\cdot) = (p_0(\cdot), \ldots, p_m(\cdot))$: $\mathbb{R} \to \mathbb{R}^{m+1}$, определяемая равенствами

$$p_i(t) = \begin{cases} f_i(t, u^{(1)}(t)) - f_i(t, u^{(2)}(t)), & t \in \Delta, \\ 0, & t \notin \Delta \end{cases}$$

интегрируема, поскольку $u^{(k)}(\cdot) \in \mathcal{U}$. По теореме Ляпунова множество

$$M = \left\{\xi \mid \xi = \int_A p(t)\,dt,\ A \in \mathfrak{S}\right\}$$

выпукло, и так как этому множеству принадлежат точки 0 $(A = \varnothing)$ и $\xi^1 - \xi^2$ $(A = \Delta)$, то для любого $\alpha \in [0, 1]$

существуют такие $A_\alpha \in \mathfrak{S}$ (т. е. измеримые по Лебегу), что

$$\alpha(\xi_i^1 - \xi_i^2) =$$

$$= \int_{A_\alpha} p_i(t)\,dt = \int_{A_\alpha \cap \Delta} [f_i(t, u^{(1)}(t)) - f_i(t, u^{(2)}(t))]\,dt =$$

$$= \int_{A_\alpha \cap \Delta} f_i(t, u^{(1)}(t))\,dt + \int_{\Delta \setminus A_\alpha} f_i(t, u^{(2)}(t))\,dt - \xi_i^2,$$

откуда

$$\int_{A_\alpha \cap \Delta} f_i(t, u^{(1)}(t))\,dt + \int_{\Delta \setminus A_\alpha} f_i(t, u^{(2)}(t))\,dt =$$

$$= \alpha\xi_i^1 + (1-\alpha)\,\xi_i^2, \quad i = 0, 1, \ldots, m. \tag{7}$$

Положим теперь

$$u_\alpha(t) = \begin{cases} u^{(1)}(t), & t \in A_\alpha \cap \Delta, \\ u^{(2)}(t), & t \in \Delta \setminus A_\alpha. \end{cases}$$

Тогда в силу (7)

$$\mathcal{F}_i(u_\alpha(\cdot)) = \alpha\xi_i^1 + (1-\alpha)\,\xi_i^2,$$

так что $\mathcal{F}(u_\alpha(\cdot)) = \alpha\xi^1 + (1-\alpha)\,\xi^2$. ■

Б) Существование множителей Лагранжа. Как уже было сказано, наши рассуждения будут параллельны доказательству теоремы Куна—Таккера из п 1.3.3.

По условию пара $(\hat{u}(\cdot), \hat{x})$ — решение задачи (1). Не ограничивая общности, можно считать, что $\mathcal{F}_0(\hat{u}(\cdot)) + g_0(\hat{x}) = 0$ (в противном случае можно вычесть из $\mathcal{F}_0 + g_0$ эту константу). Докажем, что множество

$$C = \{\alpha = (\alpha_0, \ldots, \alpha_m) \mid \exists\, (\bar{u}(\cdot), \bar{x}) \in \mathcal{U} \times A,$$
$$\mathcal{F}_0(\bar{u}(\cdot)) + g_0(\bar{x}) < \alpha_0,\ \mathcal{F}_i(\bar{u}(\cdot)) + g_i(\bar{x}) \leqslant \alpha_i,$$
$$i = 1, \ldots, m',\ \mathcal{F}_i(\bar{u}(\cdot)) + g_i(\bar{x}) = \alpha_i, \quad i = m'+1, \ldots, m\} \tag{8}$$

выпукло. Действительно, пусть $\alpha^k = (\alpha_0^k, \ldots, \alpha_m^k) \in C$, $k = 1, 2$, $0 < \theta < 1$, $(u^{(k)}(\cdot), x^{(k)})$ — такие элементы из $\mathcal{U} \times A$, что

$$\mathcal{F}_0(u^{(k)}(\cdot)) + g_0(x^{(k)}) < \alpha_0^k,$$

$$\mathcal{F}_i(u^{(k)}(\cdot)) + g_i(x^{(k)}) \begin{cases} \leqslant \alpha_i^k, & i = 1, \ldots, m', \\ = \alpha_i^k, & i = m'+1, \ldots, m. \end{cases}$$

По лемме о выпуклости образа существует функция $u_\theta(\cdot)\in\mathcal{U}$ такая, что

$$\mathcal{F}_i(u_\theta(\cdot))=\theta\mathcal{F}_i(u^{(1)}(\cdot))+(1-\theta)\mathcal{F}_i(u^{(2)}(\cdot)),\\ i=0,\ 1,\ \ldots,\ m.$$

Кроме того, $x_\theta=\theta x^{(1)}+(1-\theta)x^{(2)}\in A$ ввиду выпуклости A и

$$g_i(x_\theta)=g_i(\theta x^{(1)}+(1-\theta)x^{(2)})=\theta g_i(x^{(1)})+(1-\theta)g_i(x^{(2)})$$

для $i=m'+1,\ \ldots,\ m$, поскольку эти функции аффинные. Следовательно,

$$\mathcal{F}_i(u_\theta(\cdot))+g_i(x_\theta)=\\ =\theta[\mathcal{F}_i(u^{(1)}(\cdot))+g_i(x^{(1)})]+(1-\theta)[\mathcal{F}_i(u^{(2)}(\cdot))+g_i(x^{(2)})]=\\ =\theta\alpha_i^1+(1-\theta)\alpha_i^2,\quad i=m'+1,\ \ldots,\ m.$$

Далее, g_0 — выпуклая функция, и потому

$$\mathcal{F}_0(u_\theta(\cdot))+g_0(x_\theta)=\\ =\theta\mathcal{F}_0(u^{(1)}(\cdot))+(1-\theta)\mathcal{F}_0(u^{(2)}(\cdot))+g_0(\theta x^{(1)}+(1-\theta)x^{(2)})\leqslant\\ \leqslant\theta(\mathcal{F}_0(u^{(1)}(\cdot))+g_0(x^{(1)}))+(1-\theta)(\mathcal{F}_0(u^{(2)}(\cdot))+g_0(x^{(2)}))<\\ <\theta\alpha_0^1+(1-\theta)\alpha_0^2.$$

Аналогично доказывается, что

$$\mathcal{F}_j(u_\theta(\cdot))+g_j(x_\theta)\leqslant\theta\alpha_j^1+(1-\theta)\alpha_j^2,\quad j=1,\ \ldots,\ m'.$$

Но тогда из определения множества C видно, что $\theta\alpha^1+(1-\theta)\alpha^2\in C$, т. е. C выпукло.

Полагая в (8) $(\bar{u}(\cdot),\ \bar{x})=(\hat{u}(\cdot),\ \hat{x})$, получаем включение

$$C\supset\{\alpha=(\alpha_0,\ \ldots,\ \alpha_m)\,|\,\alpha_0>0,\ \alpha_i\geqslant 0,\ i=1,\ \ldots,\ m',\\ \alpha_i=0,\ i=m'+1,\ \ldots,\ m\}.\quad (9)$$

В частности, C непусто. Кроме того, $0\notin C$, так как в противном случае, согласно (8),

$$\mathcal{F}_0(\bar{u}(\cdot))+g_0(\bar{x})<0,\quad \mathcal{F}_i(\bar{u}(\cdot))+g_i(\bar{x})\leqslant 0,\\ i=1,\ \ldots,\ m',\\ \mathcal{F}_i(\bar{u}(\cdot))+g_i(\bar{x})=0,\quad i=m'+1,\ \ldots,\ m,$$

для некоторой пары $(\bar{u}(\cdot),\ \bar{x})$ и, следовательно, $(\hat{u}(\cdot),\ \hat{x})$ не является решением задачи (1).

Применяя конечномерную теорему отделимости (п. 1.3.3), находим такие $\hat{\lambda}_i$, $i=0,\ 1,\ \ldots,\ m$, не равные одновременно нулю, что для всех $\alpha \in C$ выполняется неравенство $\sum_{i=0}^{m} \hat{\lambda}_i \alpha_i \geqslant 0$.

Поскольку $(\varepsilon, 0, \ldots, 0, \alpha_j = 1, 0, \ldots, 0) \in C$, где $1 \leqslant j \leqslant m'$ и $\alpha_i = 0$ для $i \neq 0$, j, то выполняется неравенство $\varepsilon\hat{\lambda}_0 + \hat{\lambda}_j \geqslant 0$. Переходя к пределу при $\varepsilon \downarrow 0$, получаем $\hat{\lambda}_j \geqslant 0$. Точно так же

$$(\varepsilon, 0, \ldots, 0) \in C \Rightarrow \hat{\lambda}_0 \geqslant 0,$$

и таким образом, верно (5).

При любом $\varepsilon > 0$

$$(\varepsilon, 0, \ldots, 0, \alpha_j = \mathscr{F}_j(\hat{u}(\cdot)) + g_j(\hat{x}), 0, \ldots, 0) \in C \Rightarrow \hat{\lambda}_0 \varepsilon + \hat{\lambda}_j [\mathscr{F}_j(\hat{u}(\cdot)) + g_j(\hat{x})] \geqslant 0 \Rightarrow \hat{\lambda}_j [\mathscr{F}_j(\hat{u}(\cdot)) + g_j(\hat{x})] \geqslant 0. \quad (10)$$

Поскольку $(\hat{u}(\cdot), \hat{x})$ — допустимая пара,

$$\mathscr{F}_j(\hat{u}(\cdot)) + g_j(\hat{x}) \begin{cases} \leqslant 0, & 1 \leqslant j \leqslant m', \\ 0, & j = m'+1, \ldots, m. \end{cases} \quad (11)$$

Вместе с (10) и уже доказанными неравенствами $\hat{\lambda}_j \geqslant 0$, $1 \leqslant j \leqslant m'$, это дает (6).

Далее, для любых $\varepsilon > 0$, $(u(\cdot), x) \in \mathscr{U} \times A$, согласно (8),

$$(\mathscr{F}_0(u(\cdot)) + g_0(x) + \varepsilon, \mathscr{F}_1(x(\cdot)) + g_1(x), \ldots, \mathscr{F}_m(u(\cdot)) + g_m(x)) \in C \Rightarrow \sum_{i=0}^{m} \hat{\lambda}_i (\mathscr{F}_i(u(\cdot)) + g_i(x)) + \varepsilon\hat{\lambda}_0 = \mathscr{L}(u(\cdot), x, \hat{\lambda}, \hat{\lambda}_0) + \varepsilon\hat{\lambda}_0 \geqslant 0 \Rightarrow \mathscr{L}(u(\cdot), x, \hat{\lambda}, \hat{\lambda}_0) \geqslant 0.$$

Но, согласно (6), (11), $j \geqslant m'+1$ и сделанному выше предположению о $\mathscr{F}_0 + g_0$, $\mathscr{L}(\hat{u}(\cdot), \hat{x}, \hat{\lambda}, \hat{\lambda}_0) = 0$. Поэтому для функции Лагранжа (2) выполнен принцип минимума

$$\min_{\mathscr{U} \times A} \mathscr{L}(u(\cdot), x, \hat{\lambda}, \hat{\lambda}_0) = \mathscr{L}(\hat{u}(\cdot), \hat{x}, \hat{\lambda}, \hat{\lambda}_0). \quad (12)$$

Если же $\hat{\lambda}_0 > 0$, имеет место (12) и выполняются условия (5) и (6), то для всякой допустимой пары $(u(\cdot), x)$

$$\hat{\lambda}_0(\mathcal{F}_0(u(\cdot))+g_0(x)) \overset{(5)}{\geqslant} \sum_{i=0}^{m'} \hat{\lambda}_i(\mathcal{F}_i(u(\cdot))+g_i(x)) =$$
$$= \sum_{i=0}^{m} \hat{\lambda}_i(\mathcal{F}_i(u(\cdot))+g_i(x)) = \mathcal{L}(u(\cdot), x, \hat{\lambda}, \hat{\lambda}_0) \overset{(12)}{\geqslant}$$
$$\geqslant \mathcal{L}(\hat{u}(\cdot), \hat{x}, \hat{\lambda}, \hat{\lambda}_0) = \sum_{i=0}^{m} \hat{\lambda}_i(\mathcal{F}_i(\hat{u}(\cdot))+g_i(\hat{x})) \overset{(6),\,(11)}{=}$$
$$= \hat{\lambda}_0(\mathcal{F}_0(\hat{u}(\cdot))+g_0(\hat{x})) \Rightarrow \mathcal{F}_0(u(\cdot))+g_0(x) \geqslant$$
$$\geqslant \mathcal{F}_0(\hat{u}(\cdot)+g_0(\hat{x})),$$

так что $(\hat{u}(\cdot), \hat{x})$ — решение задачи (1).

В) Редукция к элементарной задаче. Легко видеть, что соотношение (12) эквивалентно одновременному выполнению соотношений (4) и

$$\min_{u(\cdot)\in\mathcal{U}} \sum_{i=0}^{m} \hat{\lambda}_i \mathcal{F}_i(u(\cdot)) =$$
$$= \min_{u(\cdot)\in\mathcal{U}} \int_\Delta \sum_{i=0}^{m} \hat{\lambda}_i f_i(t, u(t))\,dt = \int_\Delta \sum_{i=0}^{m} \hat{\lambda}_i f_i(t, \hat{u}(t))\,dt =$$
$$= \sum_{i=0}^{m} \hat{\lambda}_i \mathcal{F}_i(\hat{u}(\cdot)). \qquad (13)$$

Действительно, если выполняются (4) и (13), то

$$\mathcal{L}(u(\cdot), x, \hat{\lambda}, \hat{\lambda}_0) = \sum_{i=0}^{m} \hat{\lambda}_i \mathcal{F}_i(u(\cdot)) + \sum_{i=0}^{m} \hat{\lambda}_i g_i(x) \geqslant$$
$$\geqslant \sum_{i=0}^{m} \hat{\lambda}_i \mathcal{F}_i(\hat{u}(\cdot)) + \sum_{i=0}^{m} \hat{\lambda}_i g_i(\hat{x}) = \mathcal{L}(\hat{u}(\cdot), \hat{x}, \hat{\lambda}, \hat{\lambda}_0)$$

для всех $u(\cdot)\in\mathcal{U}$ и $x\in A$, т. е. имеет место (12). Если же, например, $\sum_{i=0}^{m} \hat{\lambda}_i g_i(\bar{x}) < \sum_{i=0}^{m} \hat{\lambda}_i g_i(\hat{x})$, для некоторого $\bar{x}\in A$, то

$$\mathcal{L}(\hat{u}(\cdot), \bar{x}, \hat{\lambda}, \hat{\lambda}_0) = \sum_{i=0}^{m} \hat{\lambda}_i \mathcal{F}_i(\hat{u}(\cdot)) +$$
$$+ \sum_{i=0}^{m} \hat{\lambda}_i g_i(\bar{x}) < \sum_{i=0}^{m} \hat{\lambda}_i \mathcal{F}_i(\hat{u}(\cdot)) + \sum_{i=0}^{m} \hat{\lambda}_i g_i(\hat{x}),$$

так что (12) не может иметь места.

Чтобы перейти теперь от (4) и (13) к принципу минимума, т. е. к (3) и (4), нам нужно показать, что в (13) можно «внести min под знак интеграла», т. е. показать что функция $\hat{u}(\cdot)$ тогда и только тогда является решением вспомогательной задачи оптимального управления

$$\int_{\Delta} \sum_{i=0}^{m} \hat{\lambda}_i f_i(t, u(\cdot))\, dt \to \inf, \quad u(\cdot) \in \mathcal{U}, \tag{14}$$

когда имеет место соотношение (3).

Подобное утверждение уже было доказано нами в п. 3.1.4. Однако там предполагалось, что функция $\hat{u}(\cdot)$ кусочно-непрерывна, а здесь она только измерима, так что нам придется воспользоваться более тонкими рассуждениями.

Г) Элементарная задача с измеримыми управлениями.

Лемма об условиях минимума в элементарной задаче оптимального управления. *Пусть Δ — промежуток в $\mathbf{R}$, $\mathfrak{U}$ — топологическое пространство, функция $f\colon \Delta \times \mathfrak{U} \to \mathbf{R}$ непрерывна и $\mathcal{U}$ — совокупность измеримых отображений $u\colon \Delta \to \mathfrak{U}$ таких, что функции $t \mapsto f(t, u(t))$ интегрируемы на Δ.*

Для того чтобы $\hat{u}(\cdot) \in \mathcal{U}$ было решением задачи

$$\mathcal{F}(u(\cdot)) = \int_{\Delta} f(t, u(t))\, dt \to \inf, \quad u(\cdot) \in \mathcal{U}, \tag{15}$$

необходимо и достаточно, чтобы равенство

$$f(t, \hat{u}(t)) = \min_{u \in \mathfrak{U}} f(t, u) \tag{16}$$

было верно почти всюду в Δ.

Доказательство. «Достаточно». Согласно (16) $f(t, u(t)) \geqslant f(t, \hat{u}(t))$ почти всюду для любой $u(\cdot) \in \mathcal{U}$, откуда $\mathcal{F}(u(\cdot)) \geqslant \mathcal{F}(\hat{u}(\cdot))$.

«Необходимо». Пусть $\hat{u}(\cdot)$ — решение задачи (15). Поскольку $\hat{u}(\cdot)$ измеримо, существуют (п. 4.2.6.) измеримые множества $A_n \subset \Delta$ такие, что $m\left(\Delta \setminus \bigcup_n A_n\right) = 0$ и $\hat{u}|_{A_n}$ продолжается до непрерывной на $\bar{A}_n$ функции. Назовем точку $\tau \in A_n$ *существенной*, если $m((\tau - \delta, \tau + \delta) \cap \cap A_n) > 0$ для любого $\delta > 0$, и покажем, что равенство (16) имеет место во всех таких точках.

В самом деле, если бы для некоторой существенной $\tau \in A_n$ и некоторого $v \in \mathfrak{U}$ выполнялось неравенство $f(\tau, v) < f(\tau, \hat{u}(\tau))$, то ввиду непрерывности функций $\hat{u}(\cdot)|_{A_n}$ и $t \mapsto f(t, v)$ такое же неравенство сохранилось бы и в некоторой δ-окрестности точки τ в множестве A_n, т. е.

$$f(t, v) < f(t, \hat{u}(t)), \quad \forall t \in (\tau-\delta, \tau+\delta) \cap A_n.$$

Функция

$$u(t) = \begin{cases} \hat{u}(t), & t \notin (\tau-\delta, \tau+\delta) \cap A_n, \\ v, & t \in (\tau-\delta, \tau+\delta) \cap A_n, \end{cases}$$

очевидно, принадлежит $\mathcal{U}$ и

$$\mathcal{F}(u(\cdot)) - \mathcal{F}(\hat{u}(\cdot)) = \int\limits_{(\tau-\delta,\, \tau+\delta) \cap A_n} [f(t, v) - f(t, \hat{u}(t))]\,dt < 0,$$

поскольку $f(t, v) < f(t, \hat{u}(t))$ и $m((\tau-\delta, \tau+\delta) \cap A_n) > 0$.

Остается доказать, что множество несущественных точек имеет меру нуль. Если $\tau \in A_n$ несущественна, то по определению найдется такое δ_τ, что $m((\tau-\delta_\tau, \tau+\delta_\tau) \cap A_n) = 0$. Сузив интервал, найдем такие рациональные $(\alpha_\tau, \beta_\tau)$, что $m((\alpha_\tau, \beta_\tau) \cap A_n) = 0$. Множество всех интервалов с рациональными концами счетно, поэтому не более чем счетно и множество тех из них, которые имеют вид $(\alpha_\tau, \beta_\tau)$ для некоторой несущественной точки $\tau \in A_n$. Пусть это интервалы $I_1, I_2, \ldots$ Тогда $A_n \cap \bigcup\limits_k I_k$ содержит все несущественные точки A_n и $m\left(A_n \cap \bigcup\limits_k I_k\right) \leqslant$ $\leqslant \sum\limits_k m(A_n \cap I_k) = 0$.

Поскольку множеств A_n не более чем счетное число, то объединение множеств несущественных точек каждого из них также имеет меру нуль. По доказанному (16) может нарушаться лишь на $\Delta \setminus \bigcup\limits_n A_n$ и в несущественных точках, т. е. оно верно почти всюду.

Д) Завершение доказательства. Согласно В) и Г) для выполнения условия (12) необходимо и достаточно, чтобы одновременно имели место (3) и (4). Заменив в последнем абзаце пункта Б) (12) на (3), (4), получаем все утверждения доказываемой теоремы. ■

4.3.4. Теорема двойственности. Обозначения в этом пункте те же, что и в предыдущем. Напомним, что понятие двойственности для задач выпуклого программирования было введено в п. 3.3.2. Оно было связано с методом возмущений, т. е. с включением индивидуальной экстремальной задачи в семейство задач того же типа. Здесь мы поступим точно так же, ограничившись для простоты одними интегральными функционалами. Кроме того, мы предположим (и это будет существенно использовано в доказательстве), что Δ — это отрезок числовой прямой и топологическое пространство $\mathfrak{U}$, в котором принимают значения различные управления $u(\cdot)$, сепарабельно, т. е. содержит счетное всюду плотное подмножество $\{u_n\}$. Например, в качестве $\mathfrak{U}$ можно взять любое подмножество в $\mathbf{R}^r$, $C(\Delta, \mathbf{R}^n)$, $C^1(\Delta, \mathbf{R}^n)$, $L_1(\Delta)$, но пространство $L_\infty(\Delta)$ (см. п 4.3.2) уже несепарабельно.

Упражнение 1. Докажите эти утверждения.

Итак, рассмотрим ляпуновскую задачу

$$\begin{gathered}\mathscr{F}_0(u(\cdot)) = \int_\Delta f_0(t, u(t))\,dt \to \inf,\\ \mathscr{F}_i(u(\cdot)) = \int_\Delta f_i(t, u(t))\,dt \leqslant 0, \quad u(t) \in \mathfrak{U},\\ i = 1, 2, \ldots, m.\end{gathered} \tag{$\mathfrak{z}$}$$

Фиксировав f_i, включим ее в семейство

$$\mathscr{F}_0(u(\cdot)) \to \inf, \quad \mathscr{F}_i(u(\cdot)) + \alpha_i \leqslant 0. \tag{$\mathfrak{z}(\alpha)$}$$

Как и в п. 3.3.2, S-функция семейства $\mathfrak{z}(\alpha)$ определяется равенством

$$S(\alpha) = \\ = \inf\{\mathscr{F}_0(u(\cdot)) \mid \mathscr{F}_i(u(\cdot)) + \alpha_i \leqslant 0, \ u(\cdot) \in \mathscr{U}\} \quad \alpha \in \mathbf{R}^m. \tag{1}$$

Аналогия между ляпуновскими задачами и задачами выпуклого программирования сохраняется и здесь и, хотя функции $\mathscr{F}_i$ не являются выпуклыми, справедлива

Лемма 1. *Функция $\alpha \mapsto S(\alpha)$ выпукла.*

Доказательство. Согласно лемме из п. 4.3.3 множество

$$\operatorname{im}\mathscr{F} = \{\xi = (\xi_0, \ldots, \xi_m) \mid \xi_i = \mathscr{F}_i(u(\cdot)), \ u(\cdot) \in \mathscr{U}\}$$

выпукло в $\mathbf{R}^{m+1}$. Переписав (1) в виде

$$S(\alpha) = \inf\{\xi_0 \mid \xi_i + \alpha_i \leqslant 0, \ \xi \in \operatorname{im}\mathscr{F}\}$$

и, полагая $g_i(\xi) = g_i(\xi_0, \ldots, \xi_m) = \xi_i$, мы обнаруживаем, что $S(\alpha)$ является значением задачи выпуклого программирования

$$g_0(\xi)\to\inf,\quad g_i(\xi)+\alpha_i\leqslant 0,\quad \xi\in\operatorname{im}\mathcal{F}.$$

Эта последняя является частным случаем задачи, рассмотренной в п. 3.3.2 (отсутствуют ограничения типа равенства, в связи с чем Y, Λ и т. п. следует опустить), так что выпуклость S обеспечивается следствием 1 п 3.3.2.

Теорема двойственности для ляпуновских задач. *Пусть Δ—промежуток в* $\mathbf{R}$, $\mathfrak{U}$—*сепарабельное топологическое пространство, функции* $f_i\colon \Delta\times\mathfrak{U}\to\mathbf{R}$, $i=0, 1, \ldots, m$, *непрерывны*, $\mathcal{U}$—*совокупность измеримых отображений* $u\colon \Delta\to\mathfrak{U}$ *таких, что функции* $t\mapsto f_i(t, u(t))$ *интегрируемы на* Δ.

Если S-функция семейства $\mathfrak{z}(\alpha)$ *непрерывна в точке* $\alpha=0$, *то для любого* $\alpha\in\operatorname{int}(\operatorname{dom}S)$

$$S(\alpha)=\sup_{\lambda\geqslant 0}\left(\lambda\alpha+\int_\Delta\Phi(t, \lambda)\,dt\right), \tag{2}$$

где

$$\Phi(t, \lambda)=\inf_{u\in\mathfrak{U}}\left(f_0(t, u)+\sum_{i=1}^m\lambda_i f_i(t, u)\right). \tag{3}$$

Доказательство. А) По лемме функция $\alpha\mapsto S(\alpha)$ выпукла, а из непрерывности ее в одной точке вытекает (предложение 3 п. 2.6.2) непрерывность и равенство $S(\alpha)=(\overline{\operatorname{conv}} S)(\alpha)$ для всех $\alpha\in\operatorname{int}(\operatorname{dom}S)$. По теореме Фенхеля—Моро (п. 2.6.3) $\overline{\operatorname{conv}}\, S=S^{**}$, так что для тех же α

$$S(\alpha)=S^{**}(\alpha)=\sup_{\lambda\in\mathbf{R}^{m*}}\{\lambda\alpha-S^*(\lambda)\}. \tag{4}$$

Теперь вычислим $S^*(\lambda)$. По определению

$$S^*(\lambda)=\sup_\alpha(\lambda\alpha-S(\alpha))=$$

$$=\sup_\alpha\left\{\lambda\alpha-\inf\int_\Delta f_0(t, u(t))\,dt\,\middle|\,\int_\Delta f_i(t, u(t))\,dt+\alpha_i\leqslant 0,\ u(\cdot)\in\mathcal{U}\right\}=$$

$$=\sup_{u(\cdot)\in\mathcal{U}}\ \sup_{\alpha_i\leqslant-\int_\Delta f_i(t, u(t))\,dt}\left\{\lambda\alpha-\int_\Delta f_0(t, u(t))\,dt\right\}=$$

$$=\begin{cases}+\infty, & \lambda\notin\mathbf{R}^m_+,\\ -\inf\limits_{u(\cdot)\in\mathcal{U}}\displaystyle\int_\Delta\left[f_0(t, u(t))\,dt+\sum_{i=1}^m\lambda_i f_i(t, u(t))\right]dt.\end{cases}$$

Подставляя найденное выражение в (4), мы видим, что равенство (2), а с ним и вся теорема, будут доказаны, если мы проверим, что

$$\inf_{u(\cdot)\in\mathcal{U}}\int\limits_{\Delta}\left[f_0(t,\ u(t))+\sum_{i=1}^{m}\lambda_i f_i(t,\ u(t))\right]dt=\int\limits_{\Delta}\Phi(t,\ \lambda)\,dt, \quad (5)$$

где $\Phi(t,\ \lambda)$ определено равенством (3).

Б) Лемма 2. *Пусть Δ — промежуток в $\mathbf{R}$, $\mathfrak{U}$ — сепарабельное топологическое пространство, функция f: $\Delta\times\mathfrak{U}\to\mathbf{R}$ непрерывна, $\mathcal{U}$ — совокупность измеримых отображений u: $\Delta\to\mathfrak{U}$ таких, что функция $t\mapsto f(t,\ u(t))$ интегрируема на Δ.*

Тогда функция

$$\varphi(t)=\inf_{v\in\mathfrak{U}}\{f(t,\ v)\} \quad (6)$$

измерима на Δ, интеграл $\int\limits_{\Delta}\varphi(t)\,dt$ (конечный или равный $-\infty$) существует и

$$\int\limits_{\Delta}\varphi(t)\,dt=\int\limits_{\Delta}\inf_{\mathfrak{U}}f(t,\ u)\,dt=\inf_{u(\cdot)\in\mathcal{U}}\int\limits_{\Delta}f(t,\ u(t))\,dt. \quad (7)$$

Доказательство. Пусть $\{v_k|k\geqslant 1\}$ — счетное всюду плотное подмножество в $\mathfrak{U}$. Ввиду непрерывности f функции

$$\varphi_n(t)=\min_{1\leqslant k\leqslant n}f(t,\ v_k) \quad (8)$$

непрерывны и при $n\to\infty$

$$\varphi_n(t)\downarrow\lim_{n\to\infty}\inf_{1\leqslant k\leqslant n}f(t,\ v_k)=\inf_k f(t,\ v_k)=$$
$$=\inf_{v\in\mathfrak{U}}f(t,\ v)=\varphi(t). \quad (9)$$

Кроме того, существуют функции $u_n(\cdot)\in\mathcal{U}$ такие, что $\varphi_n(t)=f(t,\ u_n(t))$. Действительно, при $n=1$ это верно, и если $\varphi_{n-1}(t)=f(t,\ u_{n-1}(t))$, то

$$\varphi_n(t)=\min_{1\leqslant k\leqslant n}[f(t,\ v_k)]=\min[\varphi_{n-1}(t),\ f(t,\ v_n)]=$$
$$=\min[f(t,\ u_{n-1}(t)),\ f(t,\ v_n)]=f(t,\ \tilde{u}_n(t)),$$

где функция

$$u_n(t)=\begin{cases} u_{n-1}(t), & \text{если}\quad \varphi_{n-1}(t)\leqslant f(t,\ v_n),\\ v_n, & \text{если}\quad \varphi_{n-1}(t)>f(t,\ v_n),\end{cases}$$

принадлежит $\mathcal{U}$ (ввиду непрерывности функций неравенство $\varphi_{n-1}(t) > f(t, v_n)$ выполняется на открытом множестве, и потому $u_n(\cdot)$ измерима; интегрируемость функции $t \mapsto f(t, u_n(t)) = \varphi_n(t)$ следует из непрерывности $\varphi_n(t)$).

Будучи в силу (9) поточечным пределом непрерывных (а значит, и измеримых) функций, функция $\varphi(\cdot)$ измерима [КФ, стр. 284]. Представив ее в виде разности

$$\varphi(t) = f(t, v) - [\varphi(t) - f(t, v)]$$

непрерывной и неотрицательной функций, мы убеждаемся, что интеграл $\int\limits_\Delta \varphi(t)\,dt$ существует, конечный или равный $(-\infty)$ (непрерывная на отрезке Δ функция $f(t, v)$ имеет конечный интеграл, а интеграл неотрицательной измеримой функции $\varphi(t) - f(t, v)$ существует, но может равняться $+\infty$).

Поскольку $f(t, u(t)) \geqslant \varphi(t)$ для любой $u(\cdot) \in \mathcal{U}$, неравенство

$$\inf_{u(\cdot)\in\mathcal{U}} \int\limits_\Delta f(t, u(t))\,dt \geqslant \int\limits_\Delta \varphi(t)\,dt \tag{10}$$

очевидно. Если левая часть равна $-\infty$, то равенство (7) верно. Если же она конечна, то ввиду неравенств

$$\int\limits_\Delta \varphi_1(t)\,dt \geqslant \int\limits_\Delta \varphi_n(t)\,dt = \int\limits_\Delta f(t, u_n(t))\,dt \geqslant \inf_{u(\cdot)\in\mathcal{U}} \int\limits_\Delta f(t, u(t))\,dt$$

последовательность интегралов $\int\limits_\Delta \varphi_n(t)dt$ ограничена. Применяя теорему Б. Леви [КФ, стр. 303] к неубывающей последовательности $-\varphi_n(\cdot)$, мы находим, что функция $\varphi(\cdot)$ интегрируема и

$$\int\limits_\Delta \varphi(t)\,dt = \lim_{n\to\infty} \int\limits_\Delta \varphi_n(t)\,dt \geqslant \inf_n \int\limits_\Delta \varphi_n(t)\,dt = \inf_n \int\limits_\Delta f(t, u_n(t))\,dt \geqslant \inf_{u(\cdot)\in\mathcal{U}} \int\limits_\Delta f(t, u(t))\,dt.$$

Вместе с (10) это дает (7). ■

По лемме 2 верно (5), что, как мы видели, доказывает теорему.

Упражнение 2. Сформулируйте и докажите теорему двойственности для ляпуновской задачи общего вида (см. (1) в п. 4.3 3).

4.3.5. Принцип максимума для задач оптимального управления, линейных по фазовым переменным. Вернемся к поставленной в п. 4.3.1 задаче

$$J_0(x(\cdot),\ u(\cdot))=\int_\Delta [a_0(t)x(t)+f_0(t,\ u(t))]\,dt+$$
$$+\gamma_{00}x(t_0)+\gamma_{10}x(t_1)\to\inf, \qquad (1)$$
$$\dot x=A(t)x+F(t,\ u(t)),\quad u(t)\in\mathfrak{U}, \qquad (2)$$
$$J_i(x(\cdot),\ u(\cdot))=\int_\Delta [a_i(t)x(t)+f_i(t,\ u(t))]\,dt+$$
$$+\gamma_{0i}x(t_0)+\gamma_{1i}x(t_1)\leqslant c_i,\quad i=1,\ 2,\ \ldots,\ m. \qquad (3)$$

Теорема (принцип максимума для задач, линейных по фазовым переменным). *Пусть функции a_i: $\Delta\to\mathbf{R}^{n*}$, $i=0,\ 1,\ \ldots,\ m$, и A: $\Delta\to\mathscr{L}(\mathbf{R}^n,\ \mathbf{R}^n)$ интегрируемы на отрезке $\Delta=[t_0,\ t_1]\subset\mathbf{R}$; $\mathfrak{U}$ — топологическое пространство; f_i: $\Delta\times\mathfrak{U}\to\mathbf{R}$, $i=0,\ 1,\ \ldots,\ m$, F: $\Delta\times\mathfrak{U}\to\mathbf{R}^n$ — непрерывные функции, $\mathscr{U}$ — совокупность измеримых отображений u: $\Delta\to\mathfrak{U}$ таких, что функции $t\mapsto f_i(t,\ u(t))$ и $t\mapsto F(t,\ u(t))$ интегрируемы.*

1. *Если $(\hat x(\cdot),\ \hat u(\cdot))$ — оптимальный процесс для задачи* (1)—(3), *то найдутся такие не равные одновременно нулю число $\hat\lambda_0\geqslant 0$, вектор $\hat\lambda=(\hat\lambda_1,\ \ldots,\ \hat\lambda_m)\in\mathbf{R}^{m*}$ и абсолютно непрерывная функция $\hat p$: $\Delta\to\mathbf{R}^{m*}$, что выполняются:*

а) *уравнение Эйлера—Лагранжа (сопряженное уравнение)*

$$\dot{\hat p}(t)=-\hat p(t)A(t)+\sum_{i=0}^m\hat\lambda_i a_i(t); \qquad (4)$$

б) *условия трансверсальности*

$$\hat p(t_k)=(-1)^k\sum_{i=0}^m\hat\lambda_i\gamma_{ki};\quad k=0,\ 1 \qquad (5)$$

в) *принцип максимума*

$$\max_{v\in\mathfrak{U}}G(t,\ v)=G(t,\ \hat u(t)) \qquad (6)$$

почти всюду, где

$$G(t,\ v)=\hat p(t)F(t,\ v)-\sum_{i=0}^m\hat\lambda_i f_i(t,\ v); \qquad (7)$$

г) *условие согласования знаков*

$$\hat{\lambda}_i \gtrless 0, \quad i=1, \ldots, m; \tag{8}$$

д) *условия дополняющей нежесткости*

$$\hat{\lambda}_i [J_i(\hat{x}(\cdot), \hat{u}(\cdot)) - c_i] = 0, \quad i=1, \ldots, m. \tag{9}$$

2. *Если для допустимого процесса* $(\hat{x}(\cdot), \hat{u}(\cdot))$ *существуют такие* $\hat{\lambda}_0 > 0$, $\hat{\lambda} \in \mathbf{R}^{m*}$ *и абсолютно непрерывная* $\hat{p}(\cdot): \Delta \to \mathbf{R}^{m*}$, *что выполняются условия* (4)—(9), *то* $(\hat{x}(\cdot), \hat{u}(\cdot))$*—оптимальный процесс для задачи* (1)—(3).

Формально, условия (4)—(9) здесь те же, что и соответствующие условия в п. 4.2.2 (читатель легко проведет необходимые преобразования от одних условий к другим), но напомним, что содержание их несколько иное. Управление $\hat{u}(\cdot)$ предполагается здесь всего лишь измеримым, а следовательно, необходимые условия (4) — (9) из результатов § 4.2 не вытекают.

Кроме того, вторая часть теоремы дает нам достаточные условия оптимальности, чего не было в § 4.2.

Доказательство. А) Редукция к ляпуновской задаче. В п. 4.3.1 задача (1)—(3) уже была сведена к экстремальной задаче ляпуновского типа (см. (8) в п. 4.3.1), которая отличается от задач п. 4.3.3 возможностью присутствия в ограничениях неравенств $\geqslant$. Это, однако, не страшно, поскольку терминальные члены в (8) п. 4.3.1 линейны и, поменяв знаки там, где нужно, мы снова получим выпуклые функции.

Перенумеруем индексы так, чтобы в (3) ограничения типа равенства стояли в конце, а затем, как и в § 3.2, положим $\varepsilon_i = 1$ для $i=0$ и тех i, для которых в (3) стоят знаки $\leqslant$ и $=$, и $\varepsilon_i = -1$ для тех i, для которых в (3) стоит знак $\geqslant$. Далее положим $g_0(x(\cdot)) = \beta_0 x(t_0)$ и $g_i(x(\cdot)) = \varepsilon_i(\beta_i x(t_0) - c_i)$, $i \geqslant 1$.

Теперь задача (8) п. 4.3.1 приобретает стандартный ляпуновский вид

$$\tilde{J}_0(x(\cdot), u(\cdot)) = -\varepsilon_0 \int_\Delta G_0(t, u(t))\,dt + g_0(x(\cdot)) \to \inf.$$

$$\tilde{J}_i(x(\cdot), u(\cdot)) = -\varepsilon_i \int_\Delta G_i(t, u(t))\,dt + g_i(x(\cdot)) \begin{cases} \leqslant 0, & i=1, \ldots, m', \\ =0, & i=m'+1, \ldots, m, \end{cases} \tag{10}$$

где G_i определены формулами (5) п. 4.3.1.

Б) «Необходимо». Если $(\hat{x}(\cdot),\ \hat{u}(\cdot))$—оптимальный процесс, то по теореме п. 4.3.3 существуют $\tilde{\lambda}_i$, неотрицательные при $0 \leqslant i \leqslant m'$, не равные одновременно нулю и такие, что

$$-\sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i G_i(t,\ \hat{u}(t)) = \min_v \left(-\sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i G_i(t,\ v)\right) \quad (11)$$

почти всюду,

$$\sum_{i=0}^{m} \tilde{\lambda}_i g_i(\hat{x}(\cdot)) = \min_{x(\cdot)} \sum_{i=0}^{m} \hat{\lambda}_i g_i(x(\cdot)) \quad (12)$$

и

$$\tilde{\lambda}_i \tilde{J}_i(\hat{x}(\cdot),\ \hat{u}(\cdot)) = 0, \quad 1 \leqslant i \leqslant m. \quad (13)$$

Легко видеть, что, полагая $\hat{\lambda}_i = \varepsilon\tilde{\lambda}_i$, мы удовлетворим условию (8), а из (13) следует (9).

Далее, положим

$$\hat{p}(t) = \sum_{i=0}^{m} \hat{\lambda}_i p_i(t), \quad (14)$$

где $p_i(\cdot)$ определены формулами (7) п. 4.3.1, откуда и из (6) п. 4.3.1

$$p_i(t_1) = -\gamma_{1i}, \quad p_i(t_0) = \gamma_{0i} - \beta_i. \quad (15)$$

Непосредственным дифференцированием (или используя формулу (14) п. 2.5.4) проверяем, что $\hat{p}(\cdot)$ является решением уравнения (4) и (5) верно при $k=1$.

Из (7), (14) и из (5) п. 4.3.1 получаем

$$G(t,\ v) = \hat{p}(t) F(t,\ v) - \sum_{i=0}^{m} \hat{\lambda}_i f_i(t,\ v) =$$

$$= \sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i [p_i(t) F(t,\ v) - f_i(t,\ v)] = \sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i G_i(t,\ v),$$

после чего (11) превращается в (6).

Наконец, подставляя в (12) определения g_i, получаем

$$\left(\sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i \beta_i\right) \hat{x}(t_0) = \min_x \left(\sum \varepsilon_i \tilde{\lambda}_i \beta_i\right) x.$$

Такое равенство возможно, только если $\sum_{i=0}^{m} \hat{\lambda}_i \beta_i =$

$= \sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i \beta_i = 0$, после чего ((5), $k=0$) следует из (14) и (15).

В) «Достаточно». Пусть для допустимого процесса $(\hat{x}(\cdot), \hat{u}(\cdot))$ нашлись такие $\hat{\lambda}_0 > 0$, $\hat{\lambda}_i$, $\hat{p}(\cdot)$, что выполняются условия (4)—(9). Пусть, далее, ε_i — те же, что и раньше, $\tilde{\lambda}_i = \varepsilon_i \hat{\lambda}_i$ и $p_i(\cdot)$ по-прежнему определены равенствами (7) п. 4.3.1, так что имеет место (15). Функции $p(\cdot) = \sum_{i=0}^{m} \hat{\lambda}_i p_i(\cdot)$ и $\hat{p}(\cdot)$ удовлетворяют уравнению (4) и, согласно ((5), $k=1$), совпадают при $t = t_1$. По теореме единственности $p(t) \equiv \hat{p}(t)$ и, следовательно,

$$G(t, v) \overset{(7)}{=} \hat{p}(t) F(t, v) - \sum_{i=0}^{m} \hat{\lambda}_i f_i(t, v) =$$

$$= \sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i [p_i(t) F(t, v) - f_i(t, v)] = \sum_{i=0}^{m} \varepsilon_i \tilde{\lambda}_i G_i(t, v) \qquad (16)$$

и

$$\sum_{i=0}^{m} \hat{\lambda}_i \beta_i \overset{(15)}{=} \sum_{i=0}^{m} \hat{\lambda}_i (\gamma_{0i} - p_i(t_0)) =$$

$$= \sum_{i=0}^{m} \hat{\lambda}_i \gamma_{0i} - \sum_{i=0}^{m} \hat{\lambda}_i p_i(t_0) \overset{(5)}{=} \hat{p}(t_0) - p(t_0) = 0. \qquad (17)$$

Из (16) и (6) следует (11), из (9) следует (13), а (8) с учетом равенств $\hat{\lambda}_i = \varepsilon_i \tilde{\lambda}_i$ означает, что $\tilde{\lambda}_i \geqslant 0$, $i = 1, \ldots, m$. Кроме того, из (17) и из определения $g_i(\cdot)$ вытекает, что

$$\sum_{i=0}^{m} \tilde{\lambda}_i g_i(x(\cdot)) = \sum_{i=0}^{m} \tilde{\lambda}_i \varepsilon_i \beta_i x(t_0) - \sum_{i=1}^{m} \tilde{\lambda}_i \varepsilon_i c_i =$$

$$= \left(\sum_{i=0}^{m} \hat{\lambda}_i \beta_i \right) x(t_0) - \sum_{i=1}^{m} \hat{\lambda}_i c_i = - \sum_{i=1}^{m} \hat{\lambda}_i c_i,$$

и потому (12) также имеет место. По теореме п. 4.3.3 условия (11)—(13) вместе с неравенствами $\tilde{\lambda}_0 > 0$, $\tilde{\lambda}_0 \geqslant 0$ достаточны для оптимальности $(\hat{x}(\cdot), \hat{u}(\cdot))$. ■

§ 4.4. Применение общей теории к простейшей задаче классического вариационного исчисления

Общую теорию мы применяем здесь к простейшей задаче классического вариационного исчисления, которой уделено такое большое место в учебной литературе. Подобным же образом можно получить необходимые и достаточные условия слабого экстремума и необходимые условия сильного экстремума и в общей задаче Лагранжа. Это, однако, заняло бы слишком много места, не внеся почти ничего принципиально нового. Поэтому мы ограничились частным случаем, в котором можно довести исследование до конца и продемонстрировать взаимосвязь старой и новой теории.

4.4.1. Уравнение Эйлера. Условие Вейерштрасса. Условие Лежандра. *Простейшей (векторной) задачей классического вариационного исчисления* мы называем (ср. п. 1.2.6) следующую экстремальную задачу:

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t,\ x(t),\ \dot{x}(t))\,dt \to \mathrm{extr}, \tag{1}$$

$$x(t_0) = x_0, \quad x(t_1) = x_1. \tag{2}$$

Здесь функция $L\colon V \to \mathbf{R}$ по крайней мере непрерывна в открытом множестве $V \subset \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$; x_0, x_1, t_0 и t_1 фиксированы.

Задача (1), (2) исследуется в двух вариантах. В первом из них $x(\cdot)$ принадлежит пространству $C^1([t_0,\ t_1],\ \mathbf{R}^n)$ и локальный экстремум в этом случае называется *слабым* (ср. п. 1.4.1). Во втором $x(\cdot)$ принадлежит пространству $KC^1([t_0,\ t_1],\ \mathbf{R}^n)$ функций с кусочно-непрерывной производной, наделенному нормой $\|x(\cdot)\| = \sup\limits_{t_0 \leqslant t \leqslant t_1} |x(t)|$, и локальный экстремум в этом случае называется *сильным*.

Упражнение 1. Проверьте, что пространство $KC^1([t_0, t_1],\ \mathbf{R}^n)$ нормированно, но не банахово (т. е. не является полным).

В обоих вариантах $x(\cdot)$ называется *допустимой*, если ее *расширенный график* $\{(t,\ x(t),\ \dot{x}(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset V$ и удовлетворяются краевые условия (2). Множество допустимых функций $x(\cdot)$ будем обозначать $\mathscr{V}$. Вводимые здесь и далее обозначения и терминология сохраняются на протяжении всего этого параграфа. Для краткости задачу (1), (2) мы будем называть *простейшей задачей*, опуская остальные подробности.

Упражнение 2. Докажите, что

$$\inf_{\mathscr{V}^{\partial} \cap C^1([t_0, t_1], \mathbb{R}^n)} \mathscr{J}(x(\cdot)) = \inf_{\mathscr{V}^{\partial} \cap KC^1([t_0, t_1], \mathbb{R}^n)} \mathscr{J}(x(\cdot)).$$

Указание. Воспользуйтесь леммой о скруглении углов (п. 1.4.3).

Если $x(\cdot)$ доставляет сильный экстремум в задаче (1) и при этом $\dot{x}(\cdot)$ непрерывна, то, очевидно, $x(\cdot)$ доставляет и слабый экстремум. Поэтому необходимые условия слабого экстремума являются (в случае непрерывности $\dot{x}(\cdot)$) необходимыми и для сильного экстремума. Некоторые из необходимых условий мы уже знаем. В § 1.4 элементарными средствами были выведены: уравнение Эйлера—необходимое условие слабого экстремума и (для $n=1$) условие Вейерштрасса—необходимое условие сильного минимума. Посмотрим теперь на простейшую задачу с общих позиций. С одной стороны, (1), (2) сводится к такой задаче Лагранжа:

$$\mathscr{J}(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), u(t))\, dt \to \text{extr}, \tag{3}$$

$$\dot{x} = u, \quad x(t_0) = x_0, \quad x(t_1) = x_1.$$

Если предположить, что не только L, но и ее производные по x, u непрерывны в V, то к (3) можно применить теорему Эйлера—Лагранжа § 4.1. При этом слабая оптимальность процесса $(\hat{x}(\cdot), \hat{u}(\cdot))$ в задаче (3) означает в точности то же самое, что слабая экстремальность $\hat{x}(\cdot)$ в задаче (1), (2).

Если $\hat{x}(\cdot)$ доставляет слабый экстремум в задаче (1), (2), то по теореме Эйлера—Лагранжа существуют такие $\hat{\lambda}_0$ и функция $\hat{p}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^{n*})$, что

$$\dot{\hat{p}}(t) = \hat{\lambda}_0 L_x(t, \hat{x}(t), \dot{\hat{x}}(t)), \quad \hat{p}(t) = \hat{\lambda}_0 L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)).$$

Если $\hat{\lambda}_0 = 0$, то все множители Лагранжа оказываются нулями; поэтому можно положить $\hat{\lambda}_0 = 1$. Исключая $\hat{p}(\cdot)$, получаем систему уравнений Эйлера

$$\frac{d}{dt} \hat{L}_{\dot{x}}(t) = \hat{L}_x(t) \tag{4}$$

(как всегда, здесь $\hat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))$ и т. д.).

С другой стороны, задачу (3) (предполагая для определенности, что это задача на минимум; в задаче на максимум следует заменить L на $-L$) можно рассматривать как задачу оптимального управления с $\mathfrak{U}=\mathbb{R}^n$. Если предположить, что L и L_x непрерывны по совокупности переменных, то можно применить к (1), (2) принцип максимума из § 4.2. При этом оптимальность процесса $(\hat{x}(\cdot), \hat{u}(\cdot))$ означает, что

$$x(\cdot)\in KC^1([t_0, t_1], \mathbb{R}^n)$$

доставляет в задаче (1), (2) сильный минимум.

Из принципа максимума следует, что существуют $\hat{\lambda}_0\geqslant 0$ и кусочно-дифференцируемая функция $\hat{p}(\cdot)$: $[t_0, t_1]\to\mathbb{R}^{n*}$ такие, что

$$\dot{\hat{p}}(t)=\hat{\lambda}_0\hat{L}_x(t), \tag{5}$$

$$\min_{u\in\mathbb{R}^n}\{\hat{\lambda}_0 L(t, \hat{x}(t), u)-\hat{p}(t)u\}=$$
$$=\hat{\lambda}_0 L(t, \hat{x}(t), \dot{\hat{x}}(t))-\hat{p}(t)\dot{\hat{x}}(t). \tag{6}$$

Равенство $\hat{\lambda}_0=0$ здесь невозможно, так как из (5) $\hat{\lambda}_0=0\Rightarrow\hat{p}(t)\equiv\text{const}$, причем $\hat{p}(t)\neq 0$ (иначе все множители Лагранжа равны нулю). Но если $\hat{\lambda}_0=0$ и $\hat{p}\neq 0$, то (6) не может выполняться (нетривиальная линейная функция не имеет минимума на $\mathbb{R}^n$). Таким образом, $\hat{\lambda}_0\neq 0$ и можно считать, что $\hat{\lambda}_0=1$.

Условие (6) можно переписать в субдифференциальной форме

$$\hat{p}(t)\in(\partial_{\dot{x}}L)(t, \hat{x}(t), \dot{\hat{x}}(t)). \tag{7}$$

Если L дифференцируема также и по $\dot{x}$, то из (6) следует равенство $\hat{p}(t)=\hat{L}_{\dot{x}}(t)$, и мы, с одной стороны, снова приходим к уравнению Эйлера (4), а с другой — из (6) получаем *необходимое условие Вейерштрасса*

$$L(t, \hat{x}(t), u)-L(t, \hat{x}(t), \dot{\hat{x}}(t))-$$
$$-L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))(u-\dot{\hat{x}}(t))\geqslant 0, \tag{8}$$
$$\forall u\in\mathbb{R}^n, \quad \forall t\in[t_0, t_1].$$

Введем в рассмотрение (ср. п. 1.4.4) $\mathscr{E}$—*функцию Вейерштрасса* (латинское excessus—излишек)

$$\mathscr{E}(t, x, \xi, u) = L(t, x, u) - L(t, x, \xi) - L_{\dot{x}}(t, x, \xi)(u-\xi). \quad (9)$$

Тогда условие Вейерштрасса (8) можно записать в виде

$$\mathscr{E}(t, \hat{x}(t), \dot{\hat{x}}(t), u) \geqslant 0, \quad \forall u \in \mathbf{R}^n, \ \forall t \in [t_0, t_1]. \quad (10)$$

Если допустить, что не только L и $L_{\dot{x}}$, но и $L_{\dot{x}\dot{x}}$ непрерывиа в V, то из (8) вытекает неравеиство

$$\hat{L}_{\dot{x}\dot{x}}(t) \geqslant 0, \quad t \in [t_0, t_1]. \quad (11)$$

Условие (11) называется *условием Лежандра*.

Итак, уравнение Эйлера является необходимым условием как слабого, так и сильиого экстремума. Кроме того, из принципа максимума мы вывели еще два необходимых условия сильного минимума—условие Вейерштрасса и условие Лежандра (в следующем пункте мы увидим, что последнее должно выполняться и для слабого минимума).

4.4.2. Условия второго порядка для слабого экстремума. Условия Лежандра и Якоби.

Лемма. *Пусть функция L непрерывна в V, а, кроме того, в V существуют и непрерывны ее производные L_{x_i}, $L_{\dot{x}_i}$, $L_{x_i x_j}$, $L_{\dot{x}_i x_j}$, $L_{\dot{x}_i \dot{x}_j}$.*

Если $\hat{x}(\cdot) \in C^1([t_0, t_1], \mathbf{R}^n)$ удовлетворяет уравнению Эйлера, то при $\|h(\cdot)\|_{C^1} \to 0$ справедлива следующая асимптотическая формула:

$$\mathscr{J}(\hat{x}(\cdot)+h(\cdot)) = \mathscr{J}(\hat{x}(\cdot)) + \int_{t_0}^{t_1} \{h(t)^{\mathrm{T}} \hat{L}_{xx}(t) h(t) + 2h^{\mathrm{T}}(t) \hat{L}_{x\dot{x}}(t) \dot{h}(t) + \dot{h}(t)^{\mathrm{T}} \hat{L}_{\dot{x}\dot{x}}(t) \dot{h}(t)\} dt + o\left(\int_{t_0}^{t_1} [|h(t)|^2 + |\dot{h}(t)|^2] dt\right), \quad (1)$$

$$\forall h(\cdot) \in C^1([t_0, t_1], \mathbf{R}^n), \quad h(t_0) = h(t_1) = 0.$$

Доказательство. Как и в п. 2.4.2, окружив график $\hat{x}(\cdot)$ компактом, целиком содержащимся в V, и воспользовавшись затем равномерной непрерывностью вторых производных функции L на нем для оценки остаточного члена в формуле Тейлора, получаем разложение

$$L(t,\ \hat{x}(t)+h(t),\ \dot{\hat{x}}(t)+\dot{h}(t)) =$$
$$=\hat{L}(t)+\hat{L}_x(t)\,h(t)+\hat{L}_{\dot{x}}(t)\,\dot{h}(t)+$$
$$+\frac{1}{2}\{h^{\mathrm{T}}(t)\,\hat{L}_{xx}(t)\,h(t)+2h^{\mathrm{T}}(t)\,\hat{L}_{x\dot{x}}(t)\,\dot{h}(t)+$$
$$+\dot{h}^{\mathrm{T}}(t)\,\hat{L}_{\dot{x}\dot{x}}(t)\,\dot{h}(t)\}+\varepsilon(t,\ h(\cdot))\,[|h(t)|^2+|\dot{h}(t)|^2], \qquad (2)$$

где $\varepsilon(t;\ h(\cdot))\to 0$ равномерно на Δ при $\|h(\cdot)\|_{C^1}\to 0$.

Интегрируя это соотношение от t_0 до t_1, а затем еще, как всегда, интегрируя по частям линейный член с $\dot{h}(t)$ и учитывая уравнение Эйлера, получаем (1). ■

Если $\hat{x}(\cdot)$ доставляет слабый минимум в задаче (1), (2) п. 4.4.1, то, заменяя в (1) $h(\cdot)$ на $\alpha h(\cdot)$ и вычисляя при $\alpha\downarrow 0$ предел $\lim[\mathcal{J}(\hat{x}(\cdot)+\alpha h(\cdot))-\mathcal{J}(\hat{x}(\cdot))]/\alpha^2$, мы получаем, что квадратичный функционал

$$Q(h(\cdot))=\int_{t_0}^{t_1}\{h^{\mathrm{T}}(t)\,\hat{L}_{xx}(t)\,h(t)+2h^{\mathrm{T}}(t)\,\hat{L}_{x\dot{x}}(t)\,\dot{h}(t)+$$
$$+\dot{h}^{\mathrm{T}}(t)\,\hat{L}_{\dot{x}\dot{x}}(t)\,\dot{h}(t)\}\,dt\geqslant 0, \qquad (3)$$
$$\forall h(\cdot)\in C^1([t_0,\ t_1],\ \mathbf{R}^n), \qquad h(t_0)=h(t_1)=0.$$

Это означает, что функция $\hat{h}(\cdot)\equiv 0$ доставляет слабый минимум во вторичной экстремальной задаче

$$Q(h(\cdot))\to\inf, \quad h(t_0)=h(t_1)=0. \qquad (4)$$

По лемме о скруглении углов (п. 1.4.3) нижние грани функционала Q в множествах допустимых функций из пространств $C^1([t_0,\ t_1],\ \mathbf{R}^n)$ и $KC^1([t_0,\ t_1],\ \mathbf{R}^n)$ совпадают. Поэтому $\hat{h}(\cdot)\equiv 0$ доставляет также и сильный минимум в задаче (4). (Заметим, что ввиду однородности Q локальные минимумы оказываются и глобальными и различие между слабым и сильным минимумами сводится к изменению запаса допустимых функций.)

По доказанному в п. 4.4.1 для задачи (4) должно выполняться условие Лежандра, которое, очевидно, имеет тот же самый вид

$$\hat{L}_{\dot{x}\dot{x}}(t)\geqslant 0, \quad \forall t\in[t_0,\ t_1], \qquad (5)$$

что и для (1), (2). Таким образом, условие Лежандра оказалось необходимым не только для сильного минимума, как в п. 4.4.1, но и для слабого.

Уравнение Эйлера для вторичной задачи (4)

$$\frac{d}{dt}[\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t)+\hat{L}_{\dot{x}x}(t)h(t)]=\hat{L}_{x\dot{x}}(t)\dot{h}(t)+\hat{L}_{xx}(t)h(t), \tag{6}$$

называется *уравнением Якоби* исходной задачи (1), (2) п. 4.4.1.

Определение. Точка τ называется *сопряженной к точке* t_0 (относительно функционала Q), если уравнение Якоби (6) имеет решение $\bar{h}(\cdot)$, для которого $\bar{h}(t_0)=\bar{h}(\tau)=0$, но[1])

$$\hat{L}_{\dot{x}\dot{x}}(\tau)\dot{\bar{h}}(\tau)\neq 0. \tag{7}$$

Теорема 1 (необходимые условия слабого экстремума в простейшей задаче). *Пусть функция L удовлетворяет условиям леммы.*

Если $\hat{x}(\cdot)$ *доставляет слабый минимум в задаче* (1), (2) п. 4.4.1, *то*

1) $\hat{x}(\cdot)$ *удовлетворяет уравнению Эйлера* ((4) п. 4.4.1);

2) *выполнено условие Лежандра* (5);

3) *выполнено условие Якоби: в интервале* (t_0, t_1) *нет точек, сопряженных с* t_0;

4) *выполняется неравенство* (3).

В случае, когда $\hat{x}(\cdot)$ *доставляет слабый максимум, в* (3) *и* (5) *следует заменить знак* $\geqslant$ *на* $\leqslant$.

Доказательство. Утверждения 1), 2) и 4) мы уже доказали (1)—в предыдущем пункте). Поэтому нужно доказать лишь 3). Предположим противное, и пусть точка $\tau\in(t_0, t_1)$—сопряженная к t_0, а $\bar{h}(\cdot)$ — соответствующее решение уравнения Якоби: $\bar{h}(t_0)=\bar{h}(\tau)=0$ и имеет место (7). Положим

$$\tilde{h}(t)=\begin{cases}\bar{h}(t), & t_0\leqslant t\leqslant\tau,\\ 0, & \tau\leqslant t\leqslant t_1.\end{cases} \tag{8}$$

---

[1]) Обычно в учебниках по вариационному исчислению сопряженные точки определяют в предположении, что выполняется усиленное условие Лежандра $\hat{L}_{\dot{x}\dot{x}}(t)>0$. При этом предположении $\hat{L}_{\dot{x}\dot{x}}(\tau)\dot{\bar{h}}(\tau)=0\Rightarrow\dot{\bar{h}}(\tau)=0$, и так как $\bar{h}(\tau)=0$, то по теореме единственности $\bar{h}(\cdot)\equiv 0$. Поэтому условие (7) эквивалентно нетривиальности $\bar{h}(\cdot)$ и обычно заменяется этим предположением.

Тогда

$$Q(\tilde{h}(\cdot)) =$$

$$= \int_{t_0}^{\tau} \{\bar{h}(t)^{\mathrm{T}} \hat{L}_{xx}(t) \bar{h}(t) + 2\bar{h}(t)^{\mathrm{T}} \hat{L}_{x\dot{x}}(t) \dot{\bar{h}}(t) +$$

$$+ \dot{\bar{h}}(t)^{\mathrm{T}} \hat{L}_{\dot{x}\dot{x}}(t) \dot{\bar{h}}(t)\} \, dt =$$

$$= \int_{t_0}^{\tau} \bar{h}(t)^{\mathrm{T}} [\hat{L}_{xx}(t) \bar{h}(t) + \hat{L}_{x\dot{x}}(t) \dot{\bar{h}}(t)] \, dt +$$

$$+ \int_{t_0}^{\tau} \dot{\bar{h}}(t)^{\mathrm{T}} [\hat{L}_{\dot{x}\dot{x}}(t) \dot{\bar{h}}(t) + \hat{L}_{\dot{x}x}(t) \bar{h}(t)] \, dt$$

(здесь использовано, что $\bar{h}(t)^{\mathrm{T}} \hat{L}_{x\dot{x}}(t) \dot{\bar{h}}(t) = \dot{\bar{h}}(t)^{\mathrm{T}} \hat{L}_{\dot{x}x}(t) \bar{h}(t)$). Вспоминая, что $\bar{h}(\cdot)$ удовлетворяет уравнению (6), проинтегрируем (с учетом равенств $\bar{h}(t_0) = \bar{h}(\tau) = 0$) по частям в первом интеграле, после чего интегралы сократятся. Таким образом, $Q(\tilde{h}(\cdot)) = 0$, а это означает, что $\tilde{h}(\cdot)$ доставляет (наряду с функцией $\hat{h}(\cdot) \equiv 0$) сильный минимум в задаче (4). Снова применяем принцип максимума из § 4.2. В соответствии с ним найдутся $\hat{\lambda}_0 \geqslant 0$ и такая кусочно-дифференцируемая функция $\tilde{p}(\cdot) \colon [t_0, t_1] \to \mathbb{R}^{n*}$, что

$$\dot{\tilde{p}}(t) = 2\hat{\lambda}_0 \hat{L}_{x\dot{x}}(t) \dot{\tilde{h}}(t) + 2\hat{\lambda}_0 \hat{L}_{xx}(t) \tilde{h}(t),$$

$$\min_{u \in \mathbb{R}^n} \{\hat{\lambda}_0 u^{\mathrm{T}} \hat{L}_{\dot{x}\dot{x}}(t) u + 2\hat{\lambda}_0 \tilde{h}(t)^{\mathrm{T}} \hat{L}_{x\dot{x}}(t) u - \tilde{p}(t) u\} =$$

$$= \hat{\lambda}_0 \dot{\tilde{h}}^{\mathrm{T}}(t) \hat{L}_{\dot{x}\dot{x}}(t) \dot{\tilde{h}}(t) + 2\hat{\lambda}_0 \tilde{h}(t)^{\mathrm{T}} \hat{L}_{x\dot{x}}(t) \dot{\tilde{h}}(t) - \tilde{p}(t) \dot{\tilde{h}}(t). \quad (9)$$

Тем же рассуждением, что и выше, доказывается, что $\hat{\lambda}_0 \neq 0$, и мы можем считать, что $\hat{\lambda}_0 = 1/2$.

Из соотношений (9) следует, что

$$\tilde{p}(t) = \hat{L}_{\dot{x}\dot{x}}(t) \dot{\tilde{h}}(t) + \hat{L}_{\dot{x}x}(t) \tilde{h}(t). \quad (10)$$

Но для $t \geqslant \tau$ функция $\tilde{h}(\cdot) \equiv 0$, так что из (10) вытекает, что $\tilde{p}(\tau + 0) = 0$. С другой стороны, $\tilde{p}(\tau - 0) = \hat{L}_{\dot{x}\dot{x}}(\tau) \dot{\tilde{h}}(\tau - 0) = \hat{L}_{\dot{x}\dot{x}}(\tau) \dot{\bar{h}}(\tau) \neq 0$ в силу (7). Поэтому $\tilde{p}$ — разрывна, но она в силу принципа максимума должна быть непрерывной. Противоречие доказывает выполнение условия Якоби. ■

Итак, мы установили, что все необходимые условия экстремума в простейшей задаче, которые обычно рассматриваются в курсах вариационного исчисления: Эйлера, Лежандра, Якоби и Вейерштрасса — являются следствиями принципа максимума. Перейдем теперь к достаточным условиям слабого экстремума.

Теорема 2 (достаточные условия слабого экстремума в простейшей задаче). *Пусть функция L удовлетворяет условиям леммы, функция* $\hat{x}(\cdot)\in C^1([t_0, t_1], \mathbf{R}^n)$ *и ее расширенный график* $\{t, \hat{x}(t), \dot{\hat{x}}(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset V$.

Если:

1) *функция* $\hat{x}(\cdot)$ *удовлетворяет уравнению Эйлера* ((4) п. 4.4.1);

2) *удовлетворяются краевые условия* $\hat{x}(t_0)=x_0$, $\hat{x}(t_1)=x_1$;

3) *выполнено усиленное условие Лежандра*

$$L_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) > 0, \quad \forall t\in[t_0, t_1]; \qquad (11)$$

4) *выполнено усиленное условие Якоби: в полуинтервале* $(t_0, t_1]$ *нет точек, сопряженных с* t_0, *то* $\hat{x}(\cdot)$ *доставляет строгий слабый минимум (при замене в* (11) *знака* $>$ *на* $<$ *— максимум) в простейшей задаче* (1), (2) п. 4.4.1.

Доказательство этой теоремы будет дано в п. 4.4.5.

4.4.3. Гамильтонов формализм. Теорема об интегральном инвариаите. В п. 4.1.1 мы уже упомянули, что уравнения Эйлера — Лагранжа вместе с уравнениями дифференциальной связи можно привести к *гамильтоновой* (или *канонической*) *системе* вида

$$\dot{x}=\mathscr{H}_p(t, x, p), \quad \dot{p}=-\mathscr{H}_x(t, x, p), \qquad (1)$$

где функция $\mathscr{H}$ называется *гамильтонианом* (или *функцией Гамильтона*) исходной задачи. Здесь мы рассмотрим это преобразование более подробно, оставаясь в рамках простейшей задачи (1), (2) п. 4.4.1. Переход от уравнения Эйлера этой задачи ((4) п. 4.4.1) к канонической системе (1) осуществляется при помощи преобразования Лежандра. Им мы сейчас и займемся.

Определение 1. Пусть функция l: $v \to \mathbf{R}$ принадлежит классу $C^1(v)$ в открытом множестве v нормированного пространства X и ее производная взаимно однозначно отображает v на $d=\{p \mid p=l'(\xi),\ \xi \in v\} \subset X^*$; Ξ: $d \longrightarrow v$ — обратное к l' отображение. Функция h: $d \longrightarrow \mathbf{R}$, определяемая равенством

$$h(p)=\langle p,\ \Xi(p)\rangle - l(\Xi(p)), \tag{2}$$

называется *классическим преобразованием Лежандра* функции l.

Предложение 1. *Пусть функция l: $v \longrightarrow \mathbf{R}$ принадлежит классу $C^2(v)$ в открытом выпуклом множестве v нормированного пространства X. Если l'' положительно определена, т. е.*

$$d^2 l(\xi;\ \eta)=l''(\xi)[\eta,\ \eta]>0, \quad \forall \xi \in v, \quad \forall \eta \in X, \quad \eta \neq 0, \tag{3}$$

то преобразование Лежандра $h(\cdot)$ функции $l(\cdot)$ определено и совпадает в d с преобразованием Юнга—Фенхеля (п. 2.6.3) *функции*

$$\tilde{l}=\begin{cases} l(\xi), & \xi \in v, \\ +\infty, & \xi \notin v, \end{cases} \tag{4}$$

которая выпукла.

Доказательство. Выпуклость $\tilde{l}$ доказана в п. 2.6.1 (пример 2), взаимная однозначность отображения $\xi \longmapsto l'(\xi)$ доказывается аналогично пункту Б) доказательства предложения 2 ниже.

По определению сопряженной функции (п. 2.6.3)

$$\tilde{l}^*(p)=\sup_{\xi}\{\langle p,\xi\rangle - \tilde{l}(\xi)\} = -\inf\{\tilde{l}(\xi)-\langle p,\ \xi\rangle\}.$$

Если $p \in d$, то $p=l'(\bar{\xi})=\tilde{l}'(\bar{\xi})$ для некоторого $\bar{\xi} \in v$. Следовательно, $0=(\tilde{l}'(\bar{\xi})-p) \in \partial[\tilde{l}(\cdot)-\langle p,\ \cdot\rangle](\bar{\xi})$, откуда $\bar{\xi}$ — точка минимума функции $\tilde{l}(\cdot)-\langle p,\ \cdot\rangle$ и

$$\tilde{l}^*(p)=-(\tilde{l}(\bar{\xi})-\langle p,\ \bar{\xi}\rangle)=\langle p,\ \bar{\xi}\rangle - \tilde{l}(\bar{\xi}).$$

Поскольку $p=l'(\bar{\xi})$, то $\bar{\xi}=\Xi(p)$ и по определению $\tilde{l}^*(p)=h(p)$. ■

Вспоминая неравенство Юнга (п. 2.6.3), получаем

Следствие 1. *В условиях предложения 1*

$$l(\xi)+h(p) \geqslant \langle p,\ \xi\rangle, \quad \forall p \in d, \quad \forall \xi \in v. \tag{5}$$

Чтобы получить из лагранжиана $L(t,\ x,\ \dot{x})$ простейшей задачи ее гамильтониан $\mathscr{H}(t,\ x,\ p)$, нужно произвести преобразование Лежандра по аргументу $\dot{x}$. Таким образом,

$$\mathcal{H}(t, x, p) = p\dot{x} - L(t, x, \dot{x}), \tag{6}$$

где $\dot{x}$ нужно исключить, используя соотношение

$$p = L_{\dot{x}}(t, x, \dot{x}), \tag{7}$$

так что формулой (6) $\mathcal{H}$ определено корректно, если отображение $\dot{x} \mapsto L_{\dot{x}}(t, x, \dot{x})$ взаимно однозначно. (В формулах (6) и (7) $\dot{x}$ выступает как самостоятельный аргумент; чтобы не путать его с dx/dt, мы будем далее обозначать $\dot{x} = \xi$, хотя для соответствующих производных L сохраним обозначения $L_{\dot{x}}$, $\partial L/\partial \dot{x}_i$ и т. д.).

Предложение 2. *Пусть функция* $(t, x, \xi) \mapsto L(t, x, \xi)$ *принадлежит классу* $C^r(V)$, $r \geqslant 2$ *в открытом множестве* $V \subset \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$.

1) *Если*:

а) *вторая производная* $L_{\dot{x}\dot{x}}$ *в* V *не вырождается, т. е.*

$$\det\left(\frac{\partial^2 L}{\partial \dot{x}_i \partial \dot{x}_j}(t, x, \xi)\right) \neq 0, \quad \forall (t, x, \xi) \in V;$$

б) *отображение* P, *определяемое равенством* $P(t, x, \xi) = (t, x, L_{\dot{x}}(t, x, \xi))$ *взаимно однозначно отображает* V *на* $D \subset \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^{n*}$, *то* D *открыто*, $P^{-1} \in C^{r-1}(D)$ *и функция* $\mathcal{H}$: $D \to \mathbf{R}$, *определяемая равенствами* (6) *и* (7), *принадлежит классу* $C^r(D)$.

2) *Если при любых* (t, x) *сечение* $V_{t,x} = \{\xi \mid (t, x, \xi) \in V\}$ *выпукло и при любых* $(t, x, \xi) \in V$ *матрица* $L_{\dot{x}\dot{x}}(t, x, \xi)$ *положительно определена, то условия* а) *и* б) *выполняются и заключение п. 1) справедливо.*

Доказательство. А) Обратное к P отображение, существующее согласно б), очевидно, имеет вид $(t, x, p) \mapsto (t, x, \Xi(t, x, p))$, так что

$$p = L_{\dot{x}}(t, x, \xi) \Leftrightarrow \xi = \Xi(t, x, p). \tag{8}$$

Поскольку $L \in C^r(V)$, $r \geqslant 2$ отображение $P \in C^{r-1}(V) \subset C^1(V)$. Его матрица Якоби имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & E & 0 \\ L_{\dot{x}t} & L_{\dot{x}x} & L_{\dot{x}\dot{x}} \end{pmatrix},$$

и ее определитель, равный $\det L_{\dot{x}\dot{x}}$, отличен от нуля согласно а).

Пусть $(t_0, x_0, p_0)=P(t_0, x_0, \xi_0)$ — любая точка из D. По теореме об обратном отображении существует окрестность $U \ni (t_0, x_0, \xi_0)$, образ которой $P(U) \subset D$ является окрестностью (t_0, x_0, p_0). Следовательно, D открыто. По той же теореме P^{-1} (которое определено однозначно) совпадает на $P(U)$ с отображением класса C^1. Поскольку точка $(t_0, x_0, p_0) \in D$ — любая, $P^{-1} \in C^1(D)$ Далее, из (8) видно, что, как неявная функция, $\Xi(t, x, p)$ определяется уравнением $F(t, x, p, \xi)=p-L_{\dot{x}}(t, x, \xi)=0$, и так как оно образовано функциями класса C^{r-1}, то Ξ, а вместе с ней и P^{-1} принадлежит классу $C^{r-1}(D)$ (см. в п. 2.3.4 замечание после теоремы о неявной функции).

Из (6), (7) и (8) выводим тождество

$$\mathscr{H}(t, x, p)=p\Xi(t, x, p)-L(t, x, \Xi(t, x, p)). \tag{9}$$

Дифференцируя его, получаем

$$\mathscr{H}_t=p\Xi_t-L_t-L_{\dot{x}}\Xi_t=[p-L_{\dot{x}}(t, x, \Xi)]\,\Xi_t-L_t= \\ =-L_t(t, x, \Xi(t, x, p)), \tag{10}$$

$$\mathscr{H}_x=p\Xi_x-L_x-L_{\dot{x}}\Xi_x=-L_x(t, x, \Xi(t, x, p)), \tag{11}$$

$$\mathscr{H}_p=\Xi+p\Xi_p-L_{\dot{x}}\Xi_p=\Xi(t, x, p). \tag{12}$$

Из формул (10)—(12) видно, что первые производные функции $\mathscr{H}$ являются суперпозициями функций класса C^{r-1}, следовательно, сами принадлежат этому классу. Поэтому $\mathscr{H} \in C^r(D)$.

Б) Если матрица $L_{\dot{x}\dot{x}}$ положительно определена, то по критерию Сильвестра (см. [7], стр. 181—182) $\det L_{\dot{x}\dot{x}}>0$, так что верно а). Пусть сечения $V_{t,x}$ выпуклы. Тогда для любой пары точек $(t, x, \xi_k) \in V$, $k=1, 2$, точки $(t, x, \alpha\xi_1+(1-\alpha)\xi_2) \in V$, $\alpha \in [0, 1]$. Обозначив $p_k=L_{\dot{x}}(t, x, \xi_k)$, имеем

$$p_1-p_2=L_{\dot{x}}(t, x, \xi_1)-L_{\dot{x}}(t, x, \xi_2)= \\ =\int_0^1 \frac{d}{d\alpha} L_{\dot{x}}(t, x, \alpha\xi_1+(1-\alpha)\xi_2)\,d\alpha= \\ =\int_0^1 L_{\dot{x}\dot{x}}(t, x, \alpha\xi_1+(1-\alpha)\xi_2)(\xi_1-\xi_2)\,d\alpha.$$

Если $\xi_1 \neq \xi_2$, но $P(t, x, \xi_1) = (t, x, p_1) = (t, x, p_2) = P(t, \hat{x}, \xi_2)$, то $p_1 = p_2$ и

$$0 = \langle p_1 - p_2, \ \xi_1 - \xi_2 \rangle = \sum_{i=1}^{n} (p_{1i} - p_{2i})(\xi_{1i} - \xi_{2i}) =$$
$$= \int_0^1 \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial \dot{x}_i \partial \dot{x}_j}(t, x, \alpha\xi_1 + (1-\alpha)\xi_2)(\xi_{1i} - \xi_{2i})(\xi_{1j} - \xi_{2j})\, d\alpha > 0,$$

так как по условию подынтегральная функция положительна ($L_{\dot{x}\dot{x}}$ положительно определена). Противоречие показывает, что $p_1 = p_2 \Rightarrow \xi_1 = \xi_2$, т. е. верно б). ■

Вспоминая следствие 1, получаем

Следствие 2. *В условиях второй части предложения 2 для любых* $(t, x, \xi) \in V$, $(t, x, p) \in D$ *выполняется неравенство*

$$\mathscr{H}(t, x, p) + L(t, x, \xi) \geqslant \langle p, \xi \rangle.$$

Перейдем теперь к связи между уравнением Эйлера

$$\frac{d}{dt} L_{\dot{x}}(t, x, \dot{x}) = L_x(t, x, \dot{x}) \tag{13}$$

и системой Гамильтона (1). Называя далее $x\colon \Delta \to \mathbf{R}^n$ решением уравнения (13) (или пару $(x, p)\colon \Delta \to \mathbf{R}^n \times \mathbf{R}^{n*}$ — решением системы (1)), мы подразумеваем, не оговаривая этого особо, что $(t, x(t), \dot{x}(t)) \in V$ при всех $t \in \Delta$ (соответственно $(t, x(t), p(t)) \in D$). Для краткости решение $x\colon \Delta \to \mathbf{R}^n$ уравнения Эйлера (13), а также его график $\{(t, x(t)) \mid t \in \Delta\}$ мы будем называть *экстремалью*, а решение $(x, p)\colon \Delta \to \mathbf{R}^n \times \mathbf{R}^{n*}$ системы Гамильтона (1) и его график $\{(t, x(t), p(t)) \mid t \in \Delta\}$ — *канонической экстремалью* (из контекста всегда ясно, идет ли речь о функциях или об их графиках).

Предложение 3. *Пусть функция* $L \in C^2(V)$ *удовлетворяет условиям* а) *и* б) *предложения 2.*

Пара $(x(\cdot), p(\cdot))$ *тогда и только тогда является канонической экстремалью, когда* $x(\cdot)$ — *экстремаль и*

$$p(t) \equiv L_{\dot{x}}(t, x(t), \dot{x}(t)). \tag{14}$$

Доказательство. А) Пусть $x(\cdot)$ — экстремаль. Если $p(\cdot)$ определено равенством (14), то, согласно (8), $\dot{x}(t) = \Xi(t, x(t), p(t))$. Теперь первое из уравнений (1) следует из (12), а второе — из (11) и (13).

Б) Пусть $(x(\cdot), p(\cdot))$ — каноническая экстремаль. Тогда из (12), (8) и первого из уравнений (1) следует (14), а тогда второе из уравнений (1) превращается в (13) согласно (11). ■

Замечание. Предположение $L \in C^2(V)$ является здесь более сильным, чем это действительно необходимо: как и в п. 4.4.2, достаточно было бы требовать существование и непрерывность первых и вторых производных только по x_i и x_j. В справедливости этого замечания читатель может убедиться сам, внеся необходимые изменения в формулировки теорем о неявной функции и об обратной функции из п. 2.3.4.

Гамильтоновы системы (1) обладают многими замечательными свойствами (см., например, [1]). Мы ограничимся только одним из них, имеющим отношение к выводу достаточных условий экстремума.

Теорема об интегральном инварианте. *Пусть функция $\mathscr{H}$ принадлежит классу C^2 в открытом множестве $D \subset \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^{n*}$, дифференциальная форма ω определяется в D равенством*

$$\omega = p\,dx - \mathscr{H}\,dt = \sum_{i=1}^{n} p_i\,dx_i - \mathscr{H}(t, x, p)\,dt, \tag{15}$$

$\Gamma_k = \{(t_k(\alpha), x_k(\alpha), p_k(\alpha)) \mid a \leqslant \alpha \leqslant b\}$, $k = 0, 1$ — *два кусочно-гладких замкнутых контура в D (так что $t_k(a) = t_k(b)$, $x_k(a) = x_k(b)$, $p_k(a) = p_k(b)$).*

Если Γ_1 получается из Γ_2 сдвигом по каноническим экстремалям, т. е. если для каждого α точки $\mathscr{P}_0(\alpha) = (t_0(\alpha), x_0(\alpha), p_0(\alpha))$ и $\mathscr{P}_1(\alpha) = (t_1(\alpha), x_1(\alpha), p_1(\alpha))$ лежат на одной канонической экстремали, то

$$\oint_{\Gamma_0} \omega = \oint_{\Gamma_1} \omega. \tag{16}$$

Доказательство. А) Согласно классической теореме о дифференцируемости решений по начальным условиям (п. 2.5.7) и теореме о суперпозиции (п. 2.2.2) решение задачи Коши для системы (1) с начальными условиями $(t_0(\alpha), x_0(\alpha), p_0(\alpha))$ — обозначим это решение $(x(t, \alpha), p(t, \alpha))$ — будет кусочно-непрерывно дифференцируемо по α и непрерывно дифференцируемо по t. Отображение σ: $\{(t, \alpha) \mid a \leqslant \alpha \leqslant b,\ t \in [t_0(\alpha), t_1(\alpha)]\} \to D$ (при $t_1(\alpha) < t_0(\alpha)$ надо здесь написать $[t_1(\alpha), t_0(\alpha)]$), определяемое равенством $\sigma(t, \alpha) = (t, x(t, \alpha), p(t, \alpha))$, задает в D двумерную поверхность Σ (рис. 38). Ориентировав Σ надлежащим образом, будем иметь $\partial\Sigma = \Gamma_1 - \Gamma_0$ (две дуги, отвечающие $\alpha = a$ и $\alpha = b$, геометрически совпадают, но

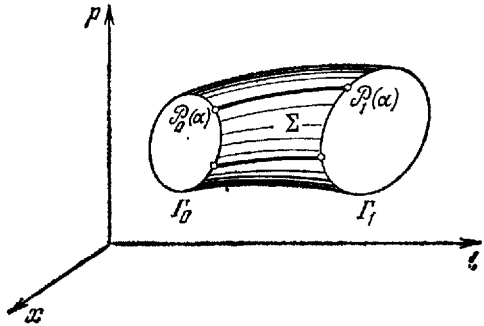


Рис. 38.

имеют противоположную ориентацию, так что интегралы по этим дугам сокращаются). Согласно формуле Стокса

$$\oint_{\Gamma_1} \omega - \oint_{\Gamma_0} \omega = \int_{\partial\Sigma} \omega = \int_{\Sigma} d\omega = \int_a^b \int_{t_0(\alpha)}^{t_1(\alpha)} d\omega[\sigma_t, \sigma_\alpha]\, dt\, d\alpha$$

и равенство (16) будет доказано, если мы проверим, что $d\omega[\sigma_t, \sigma_\alpha] \equiv 0$.

Б) Дифференцируя форму (15), получаем

$$d\omega = \sum_{i=1}^n dp_i \wedge dx_i - \left[\mathcal{H}_t\, dt + \sum_{i=1}^n (\mathcal{H}_{x_i} dx_i + \mathcal{H}_{p_i} dp_i)\right] \wedge dt =$$

$$= \sum_{i=1}^n [dp_i \wedge dx_i - \mathcal{H}_{x_i} dx_i \wedge dt - \mathcal{H}_{p_i} dp_i \wedge dt] \qquad (17)$$

(напомним, что при внешнем умножении $dt \wedge dt = 0$).

Если $\xi = (\xi_t, \xi_x, \xi_p)$ и $\eta = (\eta_t, \eta_x, \eta_p)$ — два касательных вектора, то из (17) получим:

$$d\omega[\xi, \eta] = \sum_{i=1}^n \left\{ \begin{vmatrix} \xi_{p_i} & \xi_{x_i} \\ \eta_{p_i} & \eta_{x_i} \end{vmatrix} - \mathcal{H}_{x_i} \begin{vmatrix} \xi_{x_i} & \xi_t \\ \eta_{x_i} & \eta_t \end{vmatrix} - \mathcal{H}_{p_i} \begin{vmatrix} \xi_{p_i} & \xi_t \\ \eta_{p_i} & \eta_t \end{vmatrix} \right\} =$$

$$= -\sum_i \begin{vmatrix} 1 & \mathcal{H}_{p_i} & -\mathcal{H}_{x_i} \\ \xi_t & \xi_{x_i} & \xi_{p_i} \\ \eta_t & \eta_{x_i} & \eta_{p_i} \end{vmatrix}.$$

Поэтому

$$d\omega[\sigma_t, \sigma_\alpha] = -\sum_{i=1}^{n}\begin{vmatrix} 1 & \mathscr{H}_{p_i} & -\mathscr{H}_{x_i} \\ 1 & x_{it}(t, \alpha) & p_{it}(t, \alpha) \\ 0 & x_{i\alpha}(t, \alpha) & p_{i\alpha}(t, \alpha) \end{vmatrix} = 0,$$

поскольку $\tilde{\sigma}(t, \alpha) = (x(t, \alpha), p(t, \alpha))$ по аргументу t является решением системы (1), а значит, первая и вторая строки определителя совпадают. ■

В связи с равенством (16) дифференциальную форму (15) называют *интегральным инвариантом Пуанкаре — Картана*.

Определение 2. Множество $A \subset D$ называется *лежандровым*, если

$$\oint_\gamma \omega = 0 \tag{18}$$

для любого кусочно-гладкого замкнутого контура $\gamma \subset A$. Другими словами, ограничение $\omega \mid A$ — замкнутая форма.

(Понятие кусочной гладкости нуждается здесь в уточнении. Мы будем предполагать, что A является образом открытого множества $G \subset \mathbf{R}^k$ при отображении φ: $G \to D$ класса $C^1(\bar{G})$, и γ есть образ при этом отображении обычного кусочно-гладкого замкнутого контура в G.)

Следующие три примера лежандровых множеств играют далее существенную роль.

1) Фиксируем некоторую точку $(t_0, x_0, p_0) \in D$ и рассмотрим семейство канонических экстремалей $\{(x(t, \lambda), p(t, \lambda)) \mid \alpha < t < \beta\}$ с начальными условиями

$$x(t_0, \lambda) = x_0, \quad p(t_0, \lambda) = \lambda. \tag{19}$$

Множество

$$\Sigma = \{(t, x, p) \mid x = x(t, \lambda), \ p = p(t, \lambda), \ |p_0 - \lambda| < \varepsilon, \\ \alpha < t < \beta\}$$

— искомое; $\varepsilon > 0$ и интервал $(\alpha, \beta) \ni t_0$ подбираются так, чтобы $\Sigma \subset D$.

Действительно, как было сказано выше, замкнутый кусочно-гладкий контур $\gamma \subset \Sigma$ — это образ кусочно-гладкого контура

$$\Gamma = \{(t(s), \lambda(s)) \mid a \leqslant s \leqslant b, \ t(a) = t(b), \ \lambda(a) = \lambda(b)\} \subset \\ \subset (\alpha, \beta) \times (p_0 - \varepsilon, p_0 + \varepsilon)$$

при отображении $(t, \lambda) \mapsto (t, x(t, \lambda), p(t, \lambda))$. Спроектировав Γ на плоскость $t=t_0$, получим контур $\Gamma_0 = \{(t_0, \lambda(s)) \mid a \leqslant s \leqslant b, \lambda(a) = \lambda(b)\}$, образ которого

$$\gamma_0 = \{(t_0, x(t_0, \lambda(s)), p(t_0, \lambda(s)) \mid a \leqslant s \leqslant b\} = \\ = \{(t_0, x_0, \lambda(s)) \mid a \leqslant s \leqslant b\} \quad (20)$$

в силу (19). Поскольку γ_0 получается из γ сдвигом по каноническим экстремалям, то по теореме об интегральном инварианте

$$\oint_\gamma \omega = \oint_{\gamma_0} \omega = \oint_{\gamma_0} p\,dx - \mathscr{H}\,dt = 0,$$

ибо на γ_0 в силу (20) $dx=0$, $dt=0$. Поскольку (18) верно для любого $\gamma \subset \Sigma$, это множество лежандрово.

2) Пусть в задаче (1), (2) $x(\cdot)$—числовая функция, и пусть задано семейство экстремалей $x(\cdot, \lambda)$: $(\alpha, \beta) \mapsto \mathbf{R}$, $\lambda \in (\alpha_1, \beta_1)$, причем $x(\cdot, \cdot) \in C^1((\alpha, \beta) \times (\alpha_1, \beta_1))$.

Множество

$$\Sigma = \{(t, x, p) \mid x = x(t, \lambda), \quad p = p(t, \lambda) = \\ = L_{\dot x}(t, x(t, \lambda), \dot x(t, \lambda)), \ t \in (\alpha, \beta), \ \lambda \in (\alpha_1, \beta_1)\}$$

— искомое. Действительно, пусть γ, γ_0, Γ и Γ_0 определены, как выше.

Тогда

$$\oint_\gamma \omega = \oint_{\gamma_0} \omega = \oint_{\gamma_0} p\,dx = \\ = \int_a^b p(t_0, \lambda(s)) \frac{\partial x}{\partial \lambda}(t_0, \lambda(s))\, \lambda'(s)\, ds. \quad (21)$$

Обозначив через $\Phi(\lambda)$ первообразную функции $\lambda \mapsto p(t_0, \lambda)\, x_\lambda(t_0, \lambda)$, продолжаем равенства (21):

$$\oint_\gamma \omega = \int_a^b \Phi'(\lambda(s))\, \lambda'(s)\, ds = \Phi(\lambda(s)) \Big|_{s=a}^{s=b} = 0,$$

так как $\lambda(a) = \lambda(b)$ (контур замкнут).

3) Предложение 4. *Пусть в открытом множестве $G \subset \mathbf{R} \times \mathbf{R}^n$ задана функция p: $G \to \mathbf{R}^{n*}$, $p = (p_1, \ldots, p_n)$, график которой $\Sigma = \{(t, x, p(t, x)) \mid (t, x) \in G\} \subset D$. Для того чтобы Σ было лежандровым множеством:*

а) *предполагая, что p непрерывна, необходимо и достаточно, чтобы существовала функция $S \in C^1(G)$ такая,* *что*

$$p(t, x)=\frac{\partial S(t, x)}{\partial x}, \tag{22}$$

$$\frac{\partial S}{\partial t}+\mathscr{H}\left(t, x, \frac{\partial S}{\partial x}\right)\equiv 0; \tag{23}$$

б) *предполагая, что $p\in C^1(G)$, необходимо, а в односвязной области G и достаточно, чтобы в каждой точке $(t, x)\in G$ выполнялись соотношения*

$$\frac{\partial p_i}{\partial x_k}=\frac{\partial p_k}{\partial x_i}, \tag{24}$$

$$\frac{\partial p_i}{\partial t}=-\frac{\partial \mathscr{H}}{\partial x_i}-\sum_{k=1}^{n}\frac{\partial \mathscr{H}}{\partial p_k}\frac{\partial p_k}{\partial x_i}, \quad i, k=1, \ldots, n.$$

Доказательство. Подставляя $p=p(t, x)$ в (15), получаем форму

$$\tilde{\omega}=p(t, x)\,dx-\mathscr{H}(t, x, p(t, x))\,dt.$$

Лежандровость множества Σ означает, что $\oint\limits_{\Gamma}\tilde{\omega}=0$ для любого $\Gamma\subset G$. Согласно классической теореме из анализа это равносильно точности $\tilde{\omega}$, т. е. существованию такой S, что $\tilde{\omega}=dS$. Отсюда вытекает а).

Если $p\in C^1(G)$, то $\tilde{\omega}\in C^1(G)$ и $d\tilde{\omega}=d(dS)=0$, что эквивалентно (24). Достаточность этого условия при односвязности G следует из той же классической теоремы, что и а). ■

Уравнение (23) называется *уравнением Гамильтона — Якоби*. Если нам известно некоторое его решение в области G, то (22) определяет функцию с лежандровым графиком.

4.4.4. Достаточные условия абсолютного экстремума в простейшей задаче. Рассмотрим снова простейшую задачу классического вариационного исчисления

$$\mathscr{I}(x(\cdot))=\int_{t_0}^{t_1}L(t, x, \dot{x})\,dt\to\text{extr}, \tag{1}$$

$$x(t_0)=x_0, \quad x(t_1)=x_1. \tag{2}$$

Будет предполагаться, что лагранжиан L принадлежит классу C^2 в открытом множестве $V\subset\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^n$. Допустимыми для задачи (1), (2) будем считать функции

$x(\cdot)\in KC^1([t_0, t_1], \mathbf{R}^n)$ с кусочно-непрерывной производной, расширенный график которых $\{(t, x(t), \dot{x}(t) \mid t_0 \leqslant \leqslant t \leqslant t_1\} \subset V$. Для определенности все рассуждения будут проведены для случая минимума; необходимые изменения в формулировках в случае максимума читатель легко произведет, заменив L на $-L$.

В пп. 4.4.1 и 4.4.2 мы уже видели, что выполнение условия Лежандра $\hat{L}_{\dot{x}\dot{x}}(t)=L_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) \geqslant 0$ необходимо для того, чтобы допустимая функция $\hat{x}(\cdot)$ доставляла минимум в задаче (1), (2). С другой стороны, усиленное условие Лежандра $\hat{L}_{\dot{x}\dot{x}}(t) > 0$ является составной частью достаточных условий минимума. Из него следует, в свою очередь, выпуклость функции $\dot{x} \mapsto L(t, x, \dot{x})$ вблизи точек $(t, \hat{x}(t), \dot{\hat{x}}(t))$, и это наводит на мысль, что выпуклость лагранжиана L по $\dot{x}$ во всей области определения должна входить в достаточные условия абсолютного минимума. Теорема Боголюбова (см. п. 1.4.3) подтверждает это соображение. Поэтому мы предположим, что

$$L_{\dot{x}\dot{x}}(t, x, \xi) > 0, \quad \forall (t, x, \xi) \in V \tag{3}$$

и что для любых (t, x) сечение

$$V_{t,x}=\{\xi \mid (t, x, \xi) \in V\} \quad \text{выпукло.} \tag{4}$$

Из предложения 1 п. 4.4.3 вытекает тогда, что функция

$$\tilde{L}(t, x, \xi)=\begin{cases} L(t, x, \xi), & (t, x, \xi) \in V, \\ +\infty, & (t, x, \xi) \notin V \end{cases} \tag{5}$$

выпукла по ξ.

В соответствии с предложением 2 п. 4.4.3 условия (3) и (4) обеспечивают корректное определение гамильтониана $\mathscr{H} \in C^2(D)$ соотношениями

$$\begin{gathered} \mathscr{H}(t, x, p)=p\dot{x}-L(t, x, \dot{x}), \\ p=L_{\dot{x}}(t, x, \dot{x}), \end{gathered} \tag{6}$$

из которых надо исключить $\dot{x}$. По следствию 2 того же пункта

$$\begin{gathered} L(t, x, \xi)+\mathscr{H}(t, x, p) \geqslant p\xi, \\ \forall (t, x, \xi) \in V, \quad \forall (t, x, p) \in D. \end{gathered} \tag{7}$$

Наконец, мы снова будем иметь дело с дифференциальной формой Пуанкаре—Картана

$$\omega = p\,dx - \mathscr{H}(t, x, p)\,dt. \tag{8}$$

Теорема 1 (достаточные условия минимума). *Пусть L принадлежит классу C^2 в открытом множестве $V \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ и выполняются условия* (3) *и* (4). *Пусть, далее, D определено так же, как в предложении* 2 п. 4.4.3, *и G — открытое подмножество в $\mathbb{R} \times \mathbb{R}^n$.*

Если:

1) $\hat{x}(\cdot)\colon [t_0, t_1] \to \mathbb{R}^n$ — *допустимая экстремаль задачи* (1), (2) *и ее график*

$$\hat{\Gamma} = \{(t, \hat{x}(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset G;$$

2) *существует функция $p\colon G \to \mathbb{R}^{n*}$ класса $C^1(G)$, график которой*

$$\Sigma = \{(t, x, p(t, x)) \mid (t, x) \in G\} \subset D$$

и является лежандровым множеством;

3) $$\hat{p}(t) \overset{\text{def}}{=} p(t, \hat{x}(t)) \equiv L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t));$$

то

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)\,dt > \int_{t_0}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))\,dt \tag{9}$$

для любой допустимой $x(\cdot)\colon [t_0, t_1] \to \mathbb{R}^n$, удовлетворяющей граничным условиям (2) *и такой, что ее график*

$$\Gamma = \{(t, x(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset G.$$

Замечание. Из предложения 3 п. 4.4.3 вытекает, что производная $\dot{\hat{x}}(t) = \mathscr{H}_p(t, \hat{x}(t), \hat{p}(t))$ непрерывна вместе с $\hat{x}(t)$ и $\hat{p}(t) = p(t, \hat{x}(t))$. Поэтому $\hat{x}(\cdot)$ непрерывно дифференцируема, хотя допустимыми мы считаем все кусочно-дифференцируемые кривые. Это обстоятельство не ограничивает применимости теоремы. Действительно, в п. 4.4.1 было показано, что существование непрерывной (и даже кусочно-дифференцируемой) функции $\hat{p}(\cdot)$ такой, что $\hat{p}(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))$ является необходимым условием (сильного) минимума, вытекающим из принципа максимума. При условиях (3) и (4) это, как мы видим, означает, что $\dot{\hat{x}}(\cdot)$ должна быть непрерывной.

Доказательство. Обозначим $p(t)=p(t, x(t))$ и рассмотрим кривые

$$\gamma=\{(t, x(t), p(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset \Sigma,$$
$$\hat{\gamma}=\{(t, \hat{x}(t), \hat{p}(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset \Sigma.$$

Поскольку $x(t_i)=\hat{x}(t_i)$, $i=0, 1$, и, следовательно, $p(t_i)=$ $=p(t_i, x(t_i))=p(t_i, \hat{x}(t_i))=\hat{p}(t_i)$, концы этих кривых совпадают и можно рассмотреть замкнутый контур $\gamma-$ $-\hat{\gamma}\subset\Sigma$, который получается, если сначала пройти от $(t_0, x(t_0), p(t_0))$ до $(t_1, x(t_1), p(t_1))$ вдоль γ, а затем обратно вдоль $\hat{\gamma}$. Ввиду лежандровости Σ

$$0=\oint_{\gamma-\hat{\gamma}} \omega=\int_{\gamma} \omega-\int_{\hat{\gamma}} \omega. \tag{10}$$

Используя неравенство (7), получаем

$$\int_{\gamma} \omega=\int_{\gamma} p\,dx-\mathscr{H}\,dt=\int_{t_0}^{t_1}\{p(t)\dot{x}(t)-\mathscr{H}(t, x(t), p(t))\}\,dt \leqslant$$
$$\leqslant \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt=\mathscr{J}(x(\cdot)), \tag{11}$$

причем равенство здесь возможно, только если

$$p(t)\dot{x}(t)-\mathscr{H}(t, x(t), p(t)) \equiv L(t, x(t), \dot{x}(t)),$$

т. е. (снова используя (7)) если

$$p(t)\dot{x}(t)-\mathscr{H}(t, x(t), p(t))=$$
$$=\max_{\{p \mid (t, x(t), p) \in D\}}\{p\dot{x}(t)-\mathscr{H}(t, x(t), p)\}.$$

По теореме Ферма

$$\dot{x}(t)=\mathscr{H}_p(t, x(t), p(t)). \tag{12}$$

Поскольку график p лежандров, должны выполняться равенства (24) п. 4.4.3, откуда с учетом (12)

$$\frac{dp_i(t)}{dt}=\frac{d}{dt}p_i(t, x(t))=$$
$$=\frac{\partial p_i}{\partial t}+\sum_k \frac{\partial p_i}{\partial x_k}\frac{dx_k}{dt}=-\left(\frac{\partial\mathscr{H}}{\partial x_i}+\sum_k \frac{\partial\mathscr{H}}{\partial p_k}\frac{\partial p_k}{\partial x_i}\right)+$$
$$+\sum_k \frac{\partial p_k}{\partial x_i}\frac{\partial\mathscr{H}}{\partial p_k}=-\mathscr{H}_{x_i}(t, x(t), p(t)). \tag{13}$$

Согласно (12) и (13) $(x(\cdot), p(\cdot))$ — каноническая экстремаль. Но $(x(t_0), p(t_0)) = (\hat{x}(t_0), \hat{p}(t_0))$, и так как $(\hat{x}(\cdot), \hat{p}(\cdot))$ тоже является канонической экстремалью (предложение 3 п. 4.4.3), то по теореме единственности $x(t) \equiv \hat{x}(t)$, $p(t) \equiv \hat{p}(t)$.

Итак из (11)

$$\int_{\hat{\gamma}} \omega = \int_{t_0}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))\, dt = \mathcal{J}(\hat{x}(\cdot)),$$

а для остальных допустимых $x(\cdot)$, графики которых лежат в G и которые удовлетворяют условиям (2)

$$\int_{\gamma} \omega < \mathcal{J}(x(\cdot)).$$

Ввиду (10) получаем $\mathcal{J}(x(\cdot)) > \mathcal{J}(\hat{x}(\cdot))$. ■

Следствие. *Если в условиях теоремы G совпадает с проекцией D на* $\mathbb{R} \times \mathbb{R}^n = \{(t, x)\}$, *то* $\hat{x}(\cdot)$ *доставляет строгий абсолютный минимум в задаче* (1), (2).

Часть условий только что доказанной теоремы в курсах вариационного исчисления обычно представляют в несколько иной форме, связанной с понятием поля экстремалей.

Определение. Пусть Λ и G — открытые множества в $\mathbb{R}^n$ и $\mathbb{R} \times \mathbb{R}^n$ соответственно. Семейство функций $\{x(\cdot, \lambda)\colon [t_0(\lambda), t_1(\lambda)] \to \mathbb{R}^n \mid \lambda \in \Lambda\}$ образует *поле экстремалей, покрывающее* G, если:

1) $x(\cdot, \lambda)$ является экстремалью (решением уравнения Эйлера) задачи (1), (2) для любого $\lambda \in \Lambda$;

2) отображение $(t, \lambda) \mapsto (t, x(t, \lambda))$ взаимно однозначно и его образ

$$\{(t, x) \mid x = x(t, \lambda),\ t_0(\lambda) \leqslant t \leqslant t_1(\lambda),\ \lambda \in \Lambda\} \supset G.$$

3) функция $p\colon G \to \mathbb{R}^{n*}$, определяемая равенствами

$$\begin{gathered} p(t, x) = L_{\dot{x}}(t, x(t, \lambda), \dot{x}(t, \lambda)), \\ (t, x) = (t, x(t, \lambda)), \end{gathered} \tag{14}$$

принадлежит классу $C^1(G)$ и имеет лежандров график.

Функция $u\colon G \to \mathbb{R}^n$, определяемая равенствами

$$u(t, x) = \dot{x}(t, \lambda), \quad (t, x) = (t, x(t, \lambda)), \tag{15}$$

называется *функцией наклона поля*. Экстремаль $\hat{x}(\cdot)$: $[t_0, t_1] \to \mathbf{R}^n$ включается в поле $x(\cdot, \cdot)$, если ее график содержится в G и $\hat{x}(t) \equiv x(t, \hat{\lambda})$ для некоторого $\hat{\lambda} \in \Lambda$.

Подставляя (14) и (15) в (8) и, вспоминая определение лежандрова множества (п. 4.4.3), мы можем условие 3) переформулировать еще и так:

3′) В области G интеграл

$$\int_{(t_0, x_0)}^{(t_1, x_1)} \{L_{\dot{x}}(t, x, u(t, x))\,dx - \\ - [L_{\dot{x}}(t, x, u(t, x))\,u(t, x) - L(t, x, u(t, x))]\}\,dt \quad (16)$$

не зависит от пути интегрирования между точками (t_0, x_0) и (t_1, x_1) (согласно теореме классического анализа последнее равносильно тому, что интеграл по замкнутому контуру равен нулю). Интеграл (16) называется *инвариантным интегралом Гильберта*.

Теорема 1′. *Пусть функция L удовлетворяет тем же условиям, что и в теореме* 1, *и пусть $\hat{x}(\cdot)$: $[t_0, t_1] \to \mathbf{R}^n$ — допустимая экстремаль, график которой содержится в G.*

Если $\hat{x}(\cdot)$ может быть включена в поле экстремалей, покрывающее G, то $\hat{x}(\cdot)$ доставляет функционалу (1) *минимум в классе допустимых функций, удовлетворяющих условиям* (2) *и таких, что их графики содержатся в G.*

Читатели легко убедятся, что это всего лишь переформулировка теоремы 1. Особенно удобна она при $n = 1$, так как в этом случае условие 3) приведенного выше определения следует из 1) и 2) (см. второй из примеров лежандровых множеств в п. 4.4.3). Этим замечанием можно пользоваться при решении многих задач.

4.4.5. Сопряженные точки. Достаточные условия сильного и слабого экстремума. В этом пункте мы по-прежнему будем предполагать, что $L \in C^2(V)$; $\hat{x}(\cdot) \in C^1([t_0, t_1], \mathbf{R}^n)$ — допустимая экстремаль задачи (1), (2) п. 4.4.4, вдоль которой выполняется усиленное условие Лежандра

$$L_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) > 0, \quad \forall t \in [t_0, t_1]. \quad (1)$$

По непрерывности матрица $L_{\dot{x}\dot{x}}$ остается положительно определенной в некоторой окрестности $\hat{V}$ расширенного графика $\{(t, \hat{x}(t), \dot{\hat{x}}(t)) \mid t_0 \leqslant t \leqslant t_1\}$. Окрестность $\hat{V}$ выберем так, чтобы ее сечения $\hat{V}_{t,x}=\{\xi \mid (t, x, \xi)\in\hat{V}\}$ были выпуклы. Тогда выполняются условия предложения 2 п. 4.4.3 и соответственно мы можем определить область $\hat{D}\subset\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^{n*}$ и гамильтониан $\mathscr{H}\colon \hat{D}\to\mathbf{R}$ класса C^2.

Правые части канонической системы

$$\dot{x}=\mathscr{H}_p(t, x, p), \quad \dot{p}=-\mathscr{H}_x(t, x, p) \tag{2}$$

непрерывно дифференцируемы в $\hat{D}$. Каноническую экстремаль $(\hat{x}(\cdot), \hat{p}(\cdot))$, отвечающую экстремали $\hat{x}(\cdot)$, а следовательно, и самое $\hat{x}(\cdot)$ можно продолжить на некоторый интервал, содержащий отрезок $[t_0, t_1]$. После этого, сузив, если нужно, $\hat{V}$, мы можем считать, что $\hat{V}$ имеет вид

$$\hat{V}=\{(t, x, \xi) \mid |x-\hat{x}(t)|<\varepsilon, |\xi-\dot{\hat{x}}(t)|<\varepsilon, t_0-\varepsilon<t<t_1+\varepsilon\}.$$

Пусть $(t_0, x_0, p_0)=(t_0, \hat{x}(t_0), L_{\dot{x}}(t_0, \hat{x}(t_0), \dot{\hat{x}}(t_0)))$. Построим семейство канонических экстремалей $\{(x(\cdot, \lambda), p(\cdot, \lambda)) \mid |\lambda-p_0|<\delta\}$, определяемых начальными условиями

$$x(t_0, \lambda)=x_0, \quad p(t_0, \lambda)=\lambda \tag{3}$$

(ср. первый пример лежандрова множества в п. 4.4.3). При достаточно малом δ экстремали этого семейства вместе с $(\hat{x}(\cdot), \hat{p}(\cdot))=(x(\cdot, p_0), p(\cdot, p_0))$ содержатся в $\hat{D}$ при $t\in[t_0, t_1]$; функции $x(\cdot, \cdot)$, $p(\cdot, \cdot)$, очевидно, непрерывно дифференцируемы (п. 2.5.7).

Предложение 1. *Для того чтобы точка $\tau\in(t_0, t_1]$ была сопряженной к t_0, необходимо и достаточно, чтобы*

$$\det\left(\frac{\partial x(\tau, \lambda)}{\partial\lambda}\bigg|_{\lambda=p_0}\right)=0. \tag{4}$$

Доказательство. А) По теореме п. 2.5.7 совокупность производных решений системы (2) по начальным данным при $\lambda=p_0$ образует фундаментальную матрицу решений соответствующей системы уравнений в вариациях

$$\begin{aligned} \dot{\xi}&=\hat{\mathscr{H}}_{px}\xi+\hat{\mathscr{H}}_{pp}\eta, \\ \dot{\eta}&=-\hat{\mathscr{H}}_{xx}\xi-\hat{\mathscr{H}}_{xp}\eta. \end{aligned} \tag{5}$$

Интересующие нас производные $\partial x/\partial\lambda_i$, $\partial p/\partial\lambda_i$, расположенные в столбец, составляют половину (n из $2n$) столбцов этой матрицы и в соответствии с условиями (3)

$$\begin{pmatrix} \partial x/\partial\lambda_i \\ \partial p/\partial\lambda_i \end{pmatrix}\Bigg|_{t=t_0} = \begin{pmatrix} 0 \\ e_i \end{pmatrix}. \tag{6}$$

Равенство (4) эквивалентно существованию такого ненулевого вектора $c\in\mathbf{R}^n$, что

$$\frac{\partial x}{\partial\lambda}(\tau,\ p_0)\,c=\sum_i \frac{\partial x}{\partial\lambda_i}(\tau,\ p)\,c_i=0.$$

Обозначим

$$\begin{pmatrix} \bar\xi(t) \\ \bar\eta(t) \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{n} \dfrac{\partial x}{\partial\lambda_i}(t,\ p_0)\,c_i, \\ \sum\limits_{i=1}^{n} \dfrac{\partial p}{\partial\lambda_i}(t,\ p_0)\,c_i. \end{pmatrix}.$$

Будучи линейной комбинацией столбцов фундаментальной матрицы, это тоже решение системы (5), причем в силу (6) $\bar\xi(t_0)=0$, $\bar\eta(t_0)=c\neq 0$. Кроме того, $\bar\xi(\tau)=0$.

Итак, равенство (4) эквивалентно существованию нетривиального решения системы (5), у которого $\bar\xi(t_0)= =\bar\xi(\tau)=0$.

Б) Лемма. *Пара* $\{\xi(\cdot),\ \eta(\cdot)\}$ *является решением системы* (5), *тогда и только тогда, когда* $\xi(\cdot)$ *решение уравнения Якоби*

$$\frac{d}{dt}[\hat L_{\dot x\dot x}\dot\xi+\hat L_{\dot x x}\xi]=[\hat L_{x\dot x}\dot\xi+\hat L_{xx}\xi] \tag{7}$$

и

$$\eta(t)=\hat L_{\dot x\dot x}\dot\xi(t)+\hat L_{\dot x x}\xi(t). \tag{8}$$

Доказательство. По определению (см. п. 4.4.2) уравнение (7) есть уравнение Эйлера для вторичной экстремальной задачи с лагранжианом

$$\mathfrak{L}=\frac{1}{2}\{\xi^{\mathrm{T}}\hat L_{xx}\xi+2\xi^{\mathrm{T}}\hat L_{x\dot x}\dot\xi+\dot\xi^{\mathrm{T}}\hat L_{\dot x\dot x}\dot\xi\}. \tag{9}$$

Соответствующий гамильтониан $\mathfrak{H}$ получается преобразованием Лежандра по $\dot\xi$

$$\begin{aligned} \mathfrak{H}&=\eta^{\mathrm{T}}\dot\xi-\mathfrak{L}, \\ \eta&=\mathfrak{L}_{\dot\xi}=L_{\dot x\dot x}\dot\xi+L_{\dot x x}\xi. \end{aligned} \tag{10}$$

Чтобы получить второе из этих равенств, заметим, что

$$\xi^{\mathrm{T}}\hat{L}_{x\dot{x}}\dot{\xi}=\sum_{j,\,k}\frac{\partial^2\hat{L}}{\partial x_j\,\partial\dot{x}_k}\xi_j\dot{\xi}_k,$$

откуда

$$\frac{\partial}{\partial\dot{\xi}_i}(\xi^{\mathrm{T}}\hat{L}_{x\dot{x}}\dot{\xi})=\sum_{j}\frac{\partial^2\hat{L}}{\partial x_j\,\partial\dot{x}_i}\xi_j=\sum_{j}\frac{\partial^2\hat{L}}{\partial\dot{x}_i\,\partial x_j}\xi_j=(\hat{L}_{\dot{x}x}\xi)_i.$$

Аналогично

$$\frac{\partial}{\partial\dot{\xi}_i}(\dot{\xi}^{\mathrm{T}}\hat{L}_{\dot{x}\dot{x}}\dot{\xi})=\frac{\partial}{\partial\dot{\xi}_i}\left(\sum_{j,\,k}\frac{\partial^2\hat{L}}{\partial\dot{x}_j\,\partial\dot{x}_k}\dot{\xi}_j\dot{\xi}_k\right)=$$

$$=\sum_{k}\frac{\partial^2\hat{L}}{\partial\dot{x}_i\,\partial\dot{x}_k}\dot{\xi}_k+\sum_{j}\frac{\partial^2\hat{L}}{\partial\dot{x}_j\,\partial\dot{x}_i}\dot{\xi}_j=(2\hat{L}_{\dot{x}\dot{x}}\dot{\xi})_i.$$

Исключая из (10) $\dot{\xi}$, получаем

$$\mathfrak{H}=\frac{1}{2}\{\eta^{\mathrm{T}}\hat{L}_{\dot{x}\dot{x}}^{-1}\eta-2\xi^{\mathrm{T}}\hat{L}_{x\dot{x}}\hat{L}_{\dot{x}\dot{x}}^{-1}\eta+$$
$$+\xi^{\mathrm{T}}[\hat{L}_{x\dot{x}}\hat{L}_{\dot{x}\dot{x}}^{-1}\hat{L}_{\dot{x}x}-\hat{L}_{xx}\hat{L}_{\dot{x}\dot{x}}^{-1}]\xi\}. \quad (11)$$

Теперь воспользуемся формулами (11), (12) и (8) п. 4.4.3:

$$\begin{aligned}&\mathscr{H}_x(t,\,x,\,p)=-L_x(t,\,x,\,\Xi(t,\,x,\,p)),\\&\mathscr{H}_p(t,\,x,\,p)=\Xi(t,\,x,\,p),\\&\qquad p\equiv L_{\dot{x}}(t,\,x,\,\Xi(t,\,x,\,p)).\end{aligned}$$

Дифференцируя по x и p, получаем:

$$\begin{aligned}&\mathscr{H}_{xx}=-\hat{L}_{xx}-\hat{L}_{x\dot{x}}\hat{\Xi}_x=-\hat{L}_{xx}-\hat{L}_{x\dot{x}}\mathscr{H}_{px},\\&\mathscr{H}_{xp}=-\hat{L}_{x\dot{x}}\hat{\Xi}_p=-\hat{L}_{x\dot{x}}\mathscr{H}_{pp},\\&\quad E=\frac{\partial p}{\partial p}=\hat{L}_{\dot{x}\dot{x}}\hat{\Xi}_p=\hat{L}_{\dot{x}\dot{x}}\mathscr{H}_{pp},\\&\quad 0=\frac{\partial p}{\partial x}=\hat{L}_{\dot{x}x}+\hat{L}_{\dot{x}\dot{x}}\hat{\Xi}_x=\hat{L}_{\dot{x}x}+\hat{L}_{\dot{x}\dot{x}}\mathscr{H}_{px}.\end{aligned}$$

Отсюда

$$\begin{aligned}&\mathscr{H}_{pp}=\hat{L}_{\dot{x}\dot{x}}^{-1}, &&\mathscr{H}_{xp}=-\hat{L}_{x\dot{x}}\hat{L}_{\dot{x}\dot{x}}^{-1},\\&\mathscr{H}_{px}=-\hat{L}_{\dot{x}\dot{x}}^{-1}\hat{L}_{\dot{x}x}, &&\mathscr{H}_{xx}=-\hat{L}_{xx}+\hat{L}_{x\dot{x}}\hat{L}_{\dot{x}\dot{x}}^{-1}\hat{L}_{\dot{x}x}.\end{aligned} \quad (12)$$

Следовательно, (11) можно переписать так:

$$\mathfrak{H}=\frac{1}{2}\{\eta^{\tau}\mathscr{H}_{pp}\eta+2\xi^{\tau}\mathscr{H}_{xp}\eta+\xi^{\tau}\mathscr{H}_{xx}\xi\}.$$

Согласно предложению 3 п. 4.4.3 пара $\{\xi(\cdot),\ \eta(\cdot)\}$ тогда и только тогда является решением канонической системы

$$\dot{\xi}=\mathfrak{H}_{\eta},\quad \dot{\eta}=-\mathfrak{H}_{\xi}, \tag{13}$$

когда $\xi(\cdot)$ является решением (7) и выполняется второе из соотношений (10), совпадающее с (8). Остается заметить, что система (13) тождественна (5). ■

В) Завершение доказательства. В первой части доказательства было показано, что (4) эквивалентно существованию нетривиального решения $\{\xi(\cdot),\ \eta(\cdot)\}$ системы (5), у которого $\xi(t_0)=\xi(\tau)=0$. По лемме $\xi(\cdot)$ является решением уравнения Якоби и имеет место (8), откуда $\eta(\tau)=\hat{L}_{\dot{x}\dot{x}}(\tau)\dot{\xi}(\tau)$. Равенство $\eta(\tau)=0$ противоречит (по теореме единственности) нетривиальности решения, поэтому $\eta(\tau)\neq 0$. Вспоминая определение из п. 4.4.2, мы видим, что (4) эквивалентно сопряженности точки τ точке t_0. ■

Доказанное утверждение позволяет дать следующее геометрическое истолкование понятия сопряженной точки.

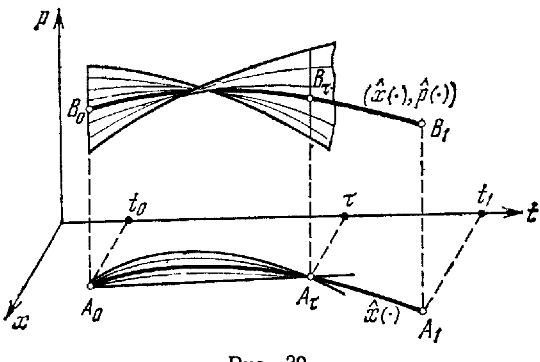


Рис. 39.

Пусть снова $(\hat{x}(\cdot),\ \hat{p}(\cdot))$—каноническая экстремаль, отвечающая рассматриваемой экстремали $\hat{x}(\cdot)$. При достаточно малом δ канонические экстремали, определяемые условиями (3), образуют полоску вдоль графика B_0B_1 (рис. 39). При $t=t_0$ край этой полоски «вертикален» и проектируется в точку $A_0=(t_0, x_0)$. Условие (4), эквивалентное, как мы видели, равенству $\frac{\partial x}{\partial \lambda}c=0$, означает, что при $t=\tau$ в точке $B_\tau=(\tau, \hat{x}(\tau), \hat{p}(\tau))$ полоска касается некоторого «вертикального» направления (т. е. направления, параллельного плоскостям $t=\text{const}$, $x=\text{const}$).

Если мы спроектируем полоску в пространство (t, x), то мы получим пучок экстремалей, выходящий из точки A_0 (рис. 39). Около точки $A_\tau=(\tau, \hat{x}(\tau))$ отображение $(t, \lambda)\to\to(t, x(t, \lambda))$ перестает быть взаимно однозначным («экстремали, бесконечно близкие к $\hat{x}(\cdot)$ и имеющие с ней общее начало́, пересекаются в A_τ»). Точка A_τ также называется сопряженной с A_0.

Упражнение. На сфере $S^2=\{(x, y, z)\mid x^2+y^2+z^2=R^2\}$ дуги большого круга суть экстремали функционала длины Проведем через точку $A_0\in S^2$ некоторый большой круг, Найдите на нем точку, сопряженную с A_0.

Предложение 2. *Если на экстремали $\hat{x}(\cdot)$: $[t_0, t_1]\to\mathbf{R}^n$ выполнено усиленное условие Лежандра* (1) *и усиленное условие Якоби (полуинтервал $(t_0, t_1]$ не содержит точек, сопряженных с t_0), то эту экстремаль можно включить в поле экстремалей, покрывающее некоторую окрестность G ее графика $\{(t, \hat{x}(t))\mid t_0\leqslant t\leqslant t_1\}$.*

Доказательство. А) Докажем сначала, что для достаточно малого $\delta>0$ ни одна из точек отрезка $[t_0, t_1]$ не будет сопряженной ни к одной точке $t_0'\in(t_0-\delta, t_0)$. В противном случае существуют такие $t_0^{(n)}\to t_0$, $t_1^{(n)}\in[t_0, t_1]$ и такие решения $\xi^{(n)}(\cdot)$ уравнения Якоби (7), что $\xi^{(n)}(t_0^{(n)})=\xi^{(n)}(t_1^{(n)})=0, \hat{L}_{\dot{x}\dot{x}}(t_1^{(n)})\dot{\xi}^{(n)}(t_1^{(n)})\neq 0$. Ввиду однородности уравнения можно положить $|\dot{\xi}^{(n)}(t_1^{(n)})|=1$. Переходя, если нужно, к подпоследовательности, мы можем считать, что $t_1^{(n)}\to t_1'\in[t_0, t_1]$,

$$\dot{\xi}^{(n)}(t_1^{(n)})\to\eta, \quad |\eta|=1.$$

При выполнении условия (1) уравнение Якоби сводится к линейной системе (5), коэффициенты которой непрерывны на отрезке $\Delta\supset[t_0, t_1]$, и задача Коши с начальными условиями

$$\xi(t_1')=0,$$

$$\eta(t_1')=\hat{L}_{\dot{x}\dot{x}}(t_1')\dot{\xi}(t_1')+\hat{L}_{\dot{x}x}(t_1')\xi(t_1')=\hat{L}_{\dot{x}\dot{x}}(t_1')\eta$$

имеет для нее решение $(\xi(t), \eta(t))$, определенное на Δ (п. 2.5.4). По теореме о непрерывной зависимости решения от начальных данных (п. 2.5.5) $\xi^{(n)}(t) \to \xi(t)$ и

$$\eta^{(n)}(t) = \hat{L}_{\dot{x}x}(t)\dot{\xi}^n(t) + \hat{L}_{\dot{x}x}(t)\xi^n(t) \to \eta(t)$$

равномерно на Δ. Но тогда и $\dot{\xi}^{(n)}(t) \to \dot{\xi}(t)$ равномерно на Δ.

Далее, $\xi(t_0) = \lim\limits_n \xi(t_0^{(n)}) = \lim\limits_n \xi^{(n)}(t_0^{(n)}) = 0$. При этом точки $t_0 = \lim\limits_n t_0^{(n)}$ и $t_1' = \lim\limits_n t_1^{(n)}$ не могут совпадать, поскольку в противном случае $\eta = \dot{\xi}(t_1') = 0$, ибо

$$|\dot{\xi}(t_1')| = \left| \frac{1}{t_1^{(n)} - t_0^{(n)}} \int\limits_{t_0^{(n)}}^{t_1^{(n)}} [\dot{\xi}^{(n)}(t) - \dot{\xi}(t_1')]\,dt \right| \leqslant$$

$$\leqslant \|\dot{\xi}^{(n)} - \dot{\xi}\|_C + \max_{[t_0', t_1']} |\dot{\xi}(t) - \dot{\xi}(t_1')| \to 0,$$

Так как $\xi(t_0) = \xi(t_1') = 0$ и $\hat{L}_{\dot{x}\dot{x}}(t_1')\dot{\xi}(t_1') \neq 0$, точка $t_1' \in (t_0, t_1]$ сопряжена t_0 вопреки условию.

Б) Выберем $t_0' < t_0$ так, чтобы на $[t_0, t_1]$ не было точек, сопряженных с t_0'. Построим семейство канонических экстремалей $(x(\cdot, \lambda), p(\cdot, \lambda))$, заменив в (3) t_0 на t_0'. В силу предложения 1 $\det \frac{\partial x}{\partial \lambda}(\tau, p_0) \neq 0$, $\forall \tau \in [t_0, t_1]$.

Определим отображение $\varphi\colon C_\delta \to \mathbf{R} \times \mathbf{R}^n$ цилиндра $C_\delta = [t_0 - \delta, t_1 + \delta] \times B(p_0, \delta)$ равенством $\varphi(t, \lambda) = (t, x(t, \lambda))$. В точках (τ, p_0), $\tau \in [t_0, t_1]$ его якобиан

$$\det \varphi' = \det \begin{pmatrix} 1 & 0 \\ x_t & x_\lambda \end{pmatrix} = \det x_\lambda \neq 0,$$

и мы можем выбрать δ так, чтобы это неравенство сохранилось во всех точках C_δ. Кроме того, по теореме об обратном отображении (п. 2.3.4) у каждой точки (τ, p_0) найдется окрестность U_τ, которую φ взаимно однозначно отображает на $W_\tau = \varphi(U_\tau)$ и $\varphi^{-1} \in C^1(W_\tau)$.

Покажем, что при достаточно малом δ отображение φ взаимно однозначно на всем C_δ. Если это не так, то существуют две последовательности (t_n', λ_n'), (t_n'', λ_n'') такие, что $\lambda_n' \to p_0$, $\lambda_n'' \to p_0$ и $(t_n', x(t_n', \lambda_n')) = \varphi(t_n', \lambda_n') = \varphi(t_n', \lambda_n'') = (t_n'', x(t_n'', \lambda_n''))$. Отсюда $t_n' = t_n''$.

Переходя к подпоследовательности, мы можем считать, что $t_n' = t_n'' \to \tau$. При достаточно больших n точки (t_n', λ_n')

и $(t_n'', \lambda_n'') \in U_\tau$ и равенство $\varphi(t_n', \lambda_n') = \varphi(t_n'', \lambda_n'')$ противоречит определению U_τ.

Теперь покажем, что $G = \varphi(\operatorname{int} C_\delta)$ является окрестностью графика $\hat{\Gamma} = \{(t, \hat{x}(t)) \mid t_0 \leqslant t \leqslant t_1\}$. Поскольку $\hat{\Gamma} = \varphi(\{(t, p_0) \mid t_0 \leqslant t \leqslant t_1\}) \subset G$, нужно доказать открытость G. Но если $(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{\lambda}) \in G$, $(\bar{t}, \bar{\lambda}) \in C_\delta$, то $\det \varphi'(\bar{t}, \bar{\lambda}) \neq 0$, и по той же теореме п. 2.3.4 некоторая окрестность точки $(\bar{t}, \bar{\lambda})$ отображается на окрестность точки $(\bar{t}, \bar{x})$. Поэтому $(\bar{t}, \bar{x})$ — внутренняя точка G и G открыто.

Таким образом, семейство экстремалей $\{x(\cdot, \lambda)\}$ взаимно однозначно покрывает G. То, что это семейство образует поле, было показано в первом примере лежандрова множества в п. 4.4.3. ■

Построенное поле состоит из экстремалей, проходящих через одну и ту же точку (t_0', x_0) и называется *центральным*.

Теперь мы возвратимся к достаточным условиям минимума. Прежде всего покажем, как из общей теоремы п. 4.4.4 получаются достаточные условия слабого минимума. Соответствующая теорема была уже сформулирована в п. 4.4.2.

Доказательство теоремы Якоби (теорема 2 п. 4.4.2). Если на допустимой экстремали $\hat{x}(\cdot)$ задачи (1), (2) п. 4.4.4 выполняются усиленные условия Лежандра и Якоби, то, согласно предложению 2, ее можно включить в поле экстремалей, покрывающее некоторую окрестность G графика $\hat{x}(\cdot)$. По теореме 1 (или 1′) п. 4.4.4 неравенство

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt > \mathcal{J}(\hat{x}(\cdot))$$

выполняется для всех $x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$, удовлетворяющих краевым условиям $x(t_0) = \hat{x}(t_0)$, $x(t_1) = \hat{x}(t_1)$ и таких, что график $x(\cdot)$ содержится в G, а расширенный график — в V. Если $\|x(\cdot) - \hat{x}(\cdot)\|_{C^1}$ достаточно мала, то последние два условия удовлетворяются. Следовательно, $\hat{x}(\cdot)$ доставляет в задаче (1), (2) п. 4.4.4 C^1-локальный, т. е. слабый, минимум. ■

Перейдем к достаточным условиям сильного минимума. В этом случае сами функции $x(\cdot)$ и $\hat{x}(\cdot)$ по-прежнему можно считать столь близкими, что график $x(\cdot)$ содержится в G, но $\dot{x}(\cdot)$ может быть любым и расширенный график $x(\cdot)$ не обязан помещаться в $\hat{V}$. Вне $\hat{V}$ условие (1) не гарантирует выпуклости функции $\xi \mapsto L(t, x, \xi)$, и неравенство Юнга, использованное в доказательстве теоремы 1 п. 4.4.4 (см. там формулу (7)), не обязано выполняться. Поэтому в формулировку теоремы придется ввести дополнительное предположение (условие Вейерштрасса) о неотрицательности функции

$$\mathscr{E}(t, x, u, \xi) = L(t, x, \xi) - L(t, x, u) - L_{\dot{x}}(t, x, u)(\xi - u).$$

Теорема Вейерштрасса о достаточных условиях сильного минимума. *Пусть функция L принадлежит классу C^2 в открытом множестве $V \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$, функция $\hat{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$* [1] *и ее расширенный график*

$$\overline{\Gamma} = \{(t, \hat{x}(t), \dot{\hat{x}}(t)) \mid t_0 \leqslant t \leqslant t_1\} \subset V.$$

Если:

1) $\hat{x}(\cdot)$ *удовлетворяет уравнению Эйлера*

$$\frac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x(t);$$

2) $\hat{x}(\cdot)$ *удовлетворяет краевым условиям*

$$\hat{x}(t_0) = x_0, \quad \hat{x}(t_1) = x_1;$$

3) *вдоль* $\hat{x}(\cdot)$ *выполнено усиленное условие Лежандра* (1);

4) *вдоль* $\hat{x}(\cdot)$ *выполнено усиленное условие Якоби: в полуинтервале* $(t_0, t_1]$ *нет точек, сопряженных с* t_0;

5) *существует окрестность* $\overline{V} \supset \overline{\Gamma}$, *для которой выполняется условие Вейерштрасса*

$$\mathscr{E}(t, x, u, \xi) \geqslant 0$$

для всех (t, x, u, ξ) *таких, что* $(t, x, u) \in \overline{V}$, $(t, x, \xi) \in V$, *то* $\hat{x}(\cdot)$ *доставляет сильный минимум в простейшей задаче* (1), (2) п. 4.4.4 (*строгий, если в условии Вейерштрасса* $\mathscr{E} > 0$ *при* $\xi \neq u$).

---

[1] Напомним, что $\hat{x}(\cdot)$, доставляющее сильный минимум, при выполнении условия (1) должно иметь непрерывную производную, хотя сейчас задача рассматривается в $KC^1([t_0, t_1], \mathbb{R}^n)$ (см. замечание после формулировки теоремы 1 п. 4.4.4).

Доказательство. Определим окрестность $\hat{V} \supset \overline{\Gamma}$ так же, как и в начале этого пункта, и построим окрестность $G \supset \hat{\Gamma} = \{(t, \hat{x}(t)) \mid t_0 \leqslant t \leqslant t_1\}$ и поле экстремалей, покрывающее G; без ограничения общности $\overline{V} = \hat{V}$. Обозначим через $u(t, x)$ наклон поля в точке $(t, x) = (t, x(t, \lambda)) \in G$:

$$u(t, x) = \dot{x}(t, \lambda) \quad \text{при} \quad x = x(t, \lambda),$$

и пусть

$$p(t, x) = L_{\dot{x}}(t, x, u(t, x)).$$

Согласно предложению 2 и примеру из п. 4.4.3 функция $p: G \to \mathbb{R}^{n*}$ имеет лежандров график

$$\Sigma = \{(t, x, p(t, x)) \mid (t, x) \in G\} \subset \hat{D}.$$

Теперь для допустимой $x(\cdot)$, график которой содержится в G,

$$\mathcal{J}(x(\cdot)) - \int_{\gamma} \omega = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt -$$
$$- \int_{t_0}^{t_1} \{p(t, x(t))\dot{x}(t) - \mathcal{H}(t, x(t), p(t, x(t)))\}\, dt =$$
$$= \int_{t_0}^{t_1} \{L(t, x(t), \dot{x}(t)) - L_{\dot{x}}(t, x(t), u(t, x(t)))\dot{x}(t) +$$
$$+ L_{\dot{x}}(t, x(t), u(t, x(t)))\, u(t, x(t)) - L(t, x(t), u(t, x(t)))\}\, dt =$$
$$= \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t))\, dt.$$

Поскольку экстремаль $\hat{x}(\cdot)$ содержится в поле, $\dot{\hat{x}}(t) = u(t, \hat{x}(t))$, а потому

$$\mathcal{J}(\hat{x}(\cdot)) - \int_{\hat{\gamma}} \omega = \int_{t_0}^{t_1} \mathcal{E}(t, \hat{x}(t), u(t, \hat{x}(t)), \dot{\hat{x}}(t))\, dt = 0. \tag{14}$$

Следовательно,

$$\mathcal{J}(x(\cdot)) - \mathcal{J}(\hat{x}(\cdot)) = \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t))\, dt \tag{15}$$

$\left(\text{как и в теореме 1 п. 4.4.4 } \int_{\gamma} \omega = \int_{\hat{\gamma}} \omega\right)$.

По построению (см. доказательство предложения 2 п. 4.4.5) точки $(t, x(t, \lambda), \dot{x}(t, \lambda)) = (t, x, u(t, x)) \in \hat{V} = \overline{V}$ при $(t, x) = (t, x(t, \lambda)) \in G$ и $(t, x(t), \dot{x}(t)) \in V$, поскольку $x(\cdot)$ — допустимая функция. Следовательно, применимо условие Вейерштрасса и из (15) $\mathcal{J}(x(\cdot)) - \mathcal{J}(\hat{x}(\cdot)) \geqslant 0$. Таким образом, если $x(\cdot) \in KC^1([t_0, t_1], \mathbb{R}^n)$ допустима, удовлетворяет краевым условиям и $\|x(\cdot) - \hat{x}(\cdot)\|_C$ столь мала, что график $x(\cdot)$ лежит в G, то $\mathcal{J}(x(\cdot)) \geqslant \mathcal{J}(\hat{x}(\cdot))$. Значит, $\hat{x}(\cdot)$ доставляет сильный минимум в задаче (1)—(4) п. 4.4.4.

Равенство $\mathcal{J}(x(\cdot)) = \mathcal{J}(\hat{x}(\cdot))$ возможно при этом только в случае, когда $\mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t)) = 0$ для всех t, кроме, быть может, точек разрыва $\dot{x}(\cdot)$. Если $u \neq \xi \Rightarrow \Rightarrow \mathcal{E} > 0$, то $\dot{x}(t) = u(t, x(t))$ и, в частности, $\dot{x}(\cdot)$ оказывается непрерывной. Далее, согласно соотношениям (8) и (12) п. 4.4.3,

$$p(t, x) = L_{\dot{x}}(t, x, u(t, x)) \Leftrightarrow u(t, x) = \\ = \Xi(t, x, p(t, x)) \Leftrightarrow u(t, x) = \mathcal{H}_p(t, x, p(t, x)) \Rightarrow \dot{x}(t) = \\ = \mathcal{H}_p(t, x, p(t, x)).$$

Полученная формула совпадает с (12) п. 4.4.4 и, рассуждая, как и там, мы устанавливаем, что $x(t) \equiv \hat{x}(t)$. ■

Замечание. Соотношение (15) позволяет показать, что квадратичный функционал

$$Q(\xi(\cdot)) = \int_{t_0}^{t_1} \mathcal{L}(t, \xi, \dot{\xi})\, dt$$

с лагранжианом (9) можно привести к «полному квадрату». Вычисляя функцию Вейерштрасса (и полагая $\hat{\xi}(\cdot) \equiv 0$) получаем

$$Q(\xi(\cdot)) = \int_{t_0}^{t_1} |P(t)\dot{\xi} + P(t)^{-1}[\hat{L}_{xx}(t) - \Pi(t)\Xi^{-1}(t)]\xi|^2 dt, \quad (16)$$

где $P(t)^2 = \hat{L}_{\dot{x}\dot{x}}(t)$, а $\Xi(\cdot)$ и $\Pi(\cdot)$ — решения матричной системы

$$\begin{aligned} \dot{\Xi}(t) &= \hat{\mathcal{H}}_{px}(t)\Xi + \hat{\mathcal{H}}_{pp}(t)\Pi, \\ \dot{\Pi}(t) &= -\hat{\mathcal{H}}_{xx}(t)\Xi - \hat{\mathcal{H}}_{xp}(t)\Pi \end{aligned} \quad (17)$$

с начальными условиями

$$\Xi(t_0')=0, \quad \Pi(t_0')=E. \tag{18}$$

4.4.6. Теорема Э. Нётер. В п. 1.4.1 мы уже познакомились с некоторыми ситуациями, в которых уравнение Эйлера имеет первый интеграл. Например, если лагранжиан не зависит явно от t, т. е. имеет вид $L(x, \dot{x})$, то $H=L_{\dot{x}}\dot{x}-L=\text{const}$ (интеграл энергии). В этом пункте мы познакомимся с общим приемом конструирования первых интегралов для систем дифференциальных уравнений, являющихся уравнениями Эйлера некоторых вариационных задач. Этот прием основан на простой и вместе с тем глубокой теореме, доказанной замечательной женщиной-математиком нашего века Эмми Нётер. Универсальный принцип, выражаемый теоремой Нётер, часто формулируют так: «Всякая симметрия в мире порождает закон сохранения», или еще: «Инвариантность системы относительно некоторой группы преобразований имеет следствием существование в системе первого интеграла». Так, например, инвариантность системы относительно сдвига по времени, проявляющаяся в том, что время t не входит в лагранжиан, дает нам интеграл энергии.

Перейдем к точным определениям. Пусть дано семейство отображений

$$S_\alpha\colon \mathbf{R}\times\mathbf{R}^n\to\mathbf{R}\times\mathbf{R}^n, \quad |\alpha|<\varepsilon_0,$$
$$S_\alpha(t, x)=(\mathfrak{T}(t, x, \alpha), \mathfrak{X}(t, x, \alpha)).$$

Относительно него мы будем предполагать, что:

1) функции $\mathfrak{T}$ и $\mathfrak{X}$ принадлежат классу C^2;

2) при $\alpha\to 0$

$$\begin{aligned}\mathfrak{T}(t, x, \alpha)&=t+\alpha T(t, x)+o(\alpha),\\ \mathfrak{X}(t, x, \alpha)&=x+\alpha X(t, x)+o(\alpha).\end{aligned} \tag{1}$$

Векторное поле $(T(t, x), X(t, x))$ будем называть *касательным векторным полем семейства* $\{S_\alpha\}$. Под действием преобразований S_α точка $P=(t, x)$ описывает некоторую кривую в $\mathbf{R}\times\mathbf{R}^n$ и $\{T(t, x), X(t, x)\}$—касательный вектор к этой кривой в точке P (т. е. при $\alpha=0$; рис. 40).

Лемма. *Пусть* $x(\cdot)\in C^2([t_0, t_1], \mathbf{R}^n)$. *Тогда существуют такие* $\delta>0$, $\varepsilon>0$, $x(\cdot, \cdot)\in C^2((t_0-\delta, t_1+\delta)\times\times(-\varepsilon, \varepsilon), \mathbf{R}^n)$, $t_k(\cdot)\in C^2((-\varepsilon, \varepsilon))$, $k=0, 1$, *что при*

$\alpha \in (-\varepsilon, \varepsilon)$ образ графика функции $x(\cdot)$ является графиком функции $x(\cdot, \alpha)$: $[t_0(\alpha,)\ t_1(\alpha)] \to \mathbf{R}^n$. При этом

$$x(t, 0) \equiv x(t), \quad x_\alpha(t, 0) \equiv X(t, x(t)) - \dot{x}(t)\, T(t, x(t)), \tag{2}$$

$$t_i'(0) = T(t_i, x(t_i)). \tag{3}$$

Доказательство. Образ графика $x(\cdot)$ при отображении S_α задается параметрическими формулами

$$\begin{aligned} t &= \tau(s, \alpha) = \mathfrak{T}(s, x(s), \alpha), \quad t_0 \leqslant s \leqslant t_1, \\ x &= \chi(s, \alpha) = \mathfrak{X}(s, x(s), \alpha), \quad |\alpha| \leqslant \varepsilon_0. \end{aligned} \tag{4}$$

Обе функции τ и χ как суперпозиции функций класса C^2 сами принадлежат этому классу.

Согласно (1)

$$\frac{\partial \mathfrak{T}}{\partial t}(t, x, 0) = 1, \quad \frac{\partial \mathfrak{T}}{\partial x}(t, x, 0) = 0,$$

и потому

$$\frac{\partial \tau}{\partial s}(s, 0) = \frac{\partial \mathfrak{T}}{\partial t}\frac{dt}{ds} + \frac{\partial \mathfrak{T}}{\partial x}\dot{x}(s) \equiv 1.$$

Используя компактность отрезка $t_0 \leqslant s \leqslant t_1$ и непрерывность $\partial\tau/\partial s$, находим такие $\delta_1 > 0$ и $\varepsilon_1 > 0$, что при

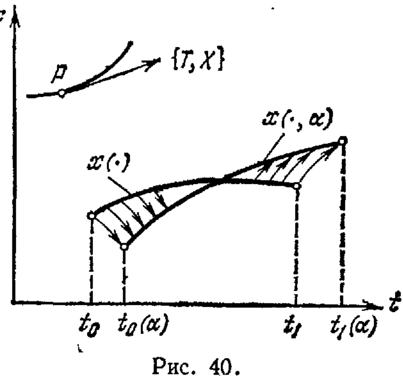


Рис. 40.

$|\alpha| < \varepsilon_1$, и $t_0 - \delta_1 < s < t_1 + \delta_1$ выполняется неравенство $\frac{\partial \tau}{\partial s}(s, \alpha) > 1/2$. Тогда при фиксированном $\alpha \in (-\varepsilon, \varepsilon)$ функция $s \mapsto \tau(s, \alpha)$ монотонно возрастает и отображает $[t_0, t_1]$ на

$$[t_0(\alpha), t_1(\alpha)] = [\tau(t_0, \alpha), \tau(t_1, \alpha)] = \\ = [\mathfrak{T}(t_0, x(t_0), \alpha), \mathfrak{T}(t_1, x(t_1), \alpha)]. \tag{5}$$

Из этих равенств видно, что $t_l(\alpha)$, $l=0, 1$,— функции класса C^2, причем из (1) следует (3).

В силу той же монотонности отображение A, определенное равенством $A(s, \alpha)=(\tau(s, \alpha), \alpha)$, взаимно однозначно отображает прямоугольник $\Pi=(t_0-\delta_1, t_1+\delta_1)\times\times(-\varepsilon_1, \varepsilon_1)$ на его образ $W=A(\Pi)$. Поскольку якобиан A

$$\det\begin{Bmatrix}\partial\tau/\partial s & \partial\tau/\partial\alpha\\ 0 & 1\end{Bmatrix}=\frac{\partial\tau}{\partial s}>\frac{1}{2},$$

то, рассуждая так же, как в доказательстве предложения 2 п. 4.4.5, устанавливаем, что W открыто и что обратное отображение A^{-1}: $W\to\Pi$, которое, очевидно, имеет вид $(t, \alpha)\mapsto(\sigma(t, \alpha), \alpha)$, непрерывно дифференцируемо и $\sigma\in C^1(W)$.

Из (4) и (1) имеем $\tau(s, 0)=\mathfrak{T}(s, x(s), 0)\equiv s$. Поэтому A оставляет на месте точки $(s, 0)$ и то же делает A^{-1}, откуда $\sigma(t, 0)\equiv t$.

Далее, как неявная функция, $\sigma(\cdot, \cdot)$ находится из уравнения $F(\sigma, t, \alpha)=\tau(\sigma, \alpha)-t=0$. Отсюда, во-первых, согласно (4) и (1),

$$\sigma_\alpha(t, 0)=-F_\sigma^{-1}F_\alpha=-(\tau_s(\sigma(t, 0), 0)^{-1}\tau_\alpha(\sigma(t, 0), 0)=$$
$$=-[\mathfrak{T}_t(t, x(t), 0)+\mathfrak{T}_x(t, x(t), 0)\dot{x}(t)]^{-1}\mathfrak{T}_\alpha(t, x(t), 0)=$$
$$=-T(t, x(t)), \qquad (6)$$

а, во-вторых, поскольку F класса C^2, то по сделанному в п. 2.3.4 замечанию $\sigma\in C^2(W)$.

Далее, A отображает отрезок $\{(s, 0)\,|\,t_0\leqslant s\leqslant t_1\}$ в себя и этот отрезок содержится в W. Следовательно, W содержит прямоугольник $(t_0-\delta, t_1+\delta)\times(-\varepsilon, \varepsilon)$, если $\delta>0$ и $\varepsilon>0$ достаточны малы, и на этом прямоугольнике

$$x(t, \alpha)=\mathfrak{X}(\sigma(t, \alpha), x(\sigma(t, \alpha)), \alpha) \qquad (7)$$

— функция класса C^2. При этом ε мы выберем так, чтобы при $|\alpha|<\varepsilon$ выполнялись неравенства $|t_i(\alpha)-t_i|<\delta$. Тогда, вспоминая определения функций $\sigma(\cdot, \cdot)$ и $t_i(\cdot)$, мы видим из (4) и (7), что образ графика $x(\cdot)$ при отображении S_α есть график функции $x(\cdot, \alpha)$: $[t_0(\alpha), t_1(\alpha)]\to\mathbf{R}^n$.

Наконец, дифференцируя (7) с учетом (1), (6) и равенства $\sigma(t, 0)=t$, находим

$$x_\alpha(t, 0)=[\mathfrak{X}_t(t, x(t), 0)+\mathfrak{X}_x(t, x(t), 0)\dot{x}(t)]\sigma_\alpha(t, 0)+ \\ +\mathfrak{X}_\alpha(t, x(t), 0)=\dot{x}(t)[-T(t, x(t))]+X(t, x(t));$$

т. е. верно (2). ■

Определение. Пусть функция $L\colon V\to\mathbf{R}$ по крайней мере непрерывна в открытом множестве $V\subset\mathbf{R}\times\times\mathbf{R}^n\times\mathbf{R}^n$. Интегральный функционал

$$\mathcal{J}(x(\cdot), t_0, t_1)=\int_{t_0}^{t_1}L(t, x(t), \dot{x}(t))\,dt \tag{8}$$

называется *инвариантным относительно семейства отображений* $\{S_\alpha\}$, если для любой функции $x(\cdot)\in C^2([t_0, t_1], \mathbf{R}^n)$ такой, что $\{(t, x(t), \dot{x}(t))\mid t_0\leqslant t\leqslant t_1\}\subset V$

$$\mathcal{J}(x(\cdot, \alpha), t_0(\alpha), t_1(\alpha))\equiv\mathcal{J}(x(\cdot), t_0, t_1) \tag{9}$$

для всех достаточно малых α.

(Здесь $x(\cdot, \alpha)$, $t_0(\alpha)$, $t_1(\alpha)$ построены по $x(\cdot)$, как в лемме, и интервал, в котором имеет место (9), может зависеть от $x(\cdot)$.)

Теорема Э. Нётер. *Пусть функции L, L_x, $L_{\dot{x}}$ непрерывны в открытом множестве $V\subset\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^n$ и интегральный функционал* (8) *инвариантен относительно семейства преобразований класса C^2, удовлетворяющих условию* (1). *Тогда функция*

$$\varphi(t, x, \dot{x})=L_{\dot{x}}(t, x, \dot{x})X(t, x)- \\ -[L_{\dot{x}}(t, x, \dot{x})\dot{x}-L(t, x, \dot{x})]T(t, x) \tag{10}$$

постоянна на каждом решении $x(\cdot)$ уравнения Эйлера

$$\frac{d}{dt}L_{\dot{x}}(t, x(t), \dot{x}(t))=L_x(t, x(t), \dot{x}(t)) \tag{11}$$

таком, что $x(\cdot)\in C^2([t_0, t_1], \mathbf{R}^n)$, $\{(t, x(t), \dot{x}(t))\mid t_0\leqslant\leqslant t\leqslant t_1\}\subset V$ и $L_{\dot{x}}(\cdot, x(\cdot), \dot{x}(\cdot))\in C^1([t_0, t_1], \mathbf{R}^{n*})$.

Доказательство. А) Пусть $x(\cdot)\colon[t_0, t_1]\to\mathbf{R}^n$ — то решение уравнения (11), о котором говорится в условии теоремы. Воспользовавшись леммой, построим семейство $(x(\cdot, \alpha), t_0(\alpha), t_1(\alpha))$, где $x(\cdot, \cdot)$ — функция класса C^2 на $(t_0-\delta, t_1+\delta)\times(-\varepsilon, \varepsilon)$. Фиксируем отрезок $\Delta=[\beta, \gamma]$ так, чтобы выполнялись неравенства $t_0-\delta<\beta<t_0<t_1<\gamma<t_1+\delta$, и уменьшим, если нужно,

ε так, чтобы при $|\alpha|<\varepsilon$ выполнялись неравенства $\beta<$ $<t_0(\alpha)<t_1(\alpha)<\gamma$.

Покажем, что отображение Φ: $(-\varepsilon, \varepsilon)\to C^1(\Delta, \mathbf{R}^n)\times$ $\times\mathbf{R}^2$, при котором $\alpha\mapsto\{x(\cdot, \alpha), t_0(\alpha), t_1(\alpha)\}$ дифференцируемо по Фреше в точке $\alpha=0$. Ясно, что достаточно проверить это для первой компоненты $\alpha\mapsto x(\cdot, \alpha)$. Фиксировав $\bar\varepsilon\in(0, \varepsilon)$ и применяя теорему о среднем (к дифференцируемым отображениям $\alpha\to x(t, \alpha)$, t — фиксировано), имеем при $|\alpha|\leqslant\bar\varepsilon$

$$\sup_{t\in\Delta}\left|\frac{x(t, \alpha)-x(t, 0)-\alpha x_\alpha(t, 0)}{\alpha}\right|\leqslant$$

$$\leqslant\sup_{t\in\Delta}\sup_{c\in[0, \alpha]}|x_\alpha(t, c)-x_\alpha(t, 0)|\to 0,$$

$$\sup_{t\in\Delta}\left|\frac{x_t(t, \alpha)-x_t(t, 0)-\alpha x_{t\alpha}(t, 0)}{\alpha}\right|\leqslant$$

$$\leqslant\sup_{t\in\Delta}\sup_{c\in[0, \alpha]}|x_{t\alpha}(t, c)-x_{t\alpha}(t, 0)|\to 0,$$

при $\alpha\to 0$, поскольку x_α и $x_{t\alpha}$ равномерно непрерывны на компакте $\Delta\times[-\bar\varepsilon, \bar\varepsilon]$. Следовательно, в пространстве $C^1(\Delta, \mathbf{R}^n)$

$$x(\cdot, \alpha)=x(\cdot, 0)+\alpha x_\alpha(\cdot, 0)+o(|\alpha|), \tag{12}$$

что и означает дифференцируемость по Фреше. Кроме того, из (12) вытекает, что

$$\Phi'(0)=\{x_\alpha(\cdot, 0), t_0'(0), t_1'(0)\}. \tag{13}$$

Б) Положим $\mathcal{I}(\alpha)=(\mathcal{I}\circ\Phi)(\alpha)=\mathcal{I}(x(\cdot, \alpha), t_0(\alpha), t_1(\alpha))$. Согласно (9) $\mathcal{I}(\alpha)\equiv\mathcal{I}(0)$ и, значит, $\mathcal{I}'(0)=$ $=\mathcal{I}'(x(\cdot), t_0, t_1)[\Phi'(0)]=0$. Воспользовавшись формулой (9) из п. 2.4.2 для производной интегрального отображения, получаем из (13)

$$0=\mathcal{I}'(x(\cdot), t_0, t_1)[\Phi'(0)]=$$

$$=\int_{t_0}^{t_1}\{L_x(t, x(t), \dot x(t))x_\alpha(t, 0)+$$

$$+L_{\dot x}(t, x(t), \dot x(t))x_{\alpha t}(t, 0)\}dt+$$

$$+L(t_i, x(t_i), \dot x(t_i))t_i'(0)\Big|_{i=0}^{i=1}. \tag{14}$$

По предположению, $L_{\dot x}(t, x(t), \dot x(t))$ непрерывно дифференцируема. Интегрируя в (14) по частям (как в п. 1.4.1) и учитывая (11), (2), (3) и (10), имеем

$$0=[L_{\dot x}(t_i,\ x(t_i),\ \dot x(t_i))\,x_\alpha(t_i,\ 0)+$$
$$+L(t_i,\ x(t_i),\ \dot x(t_i))\,t_i'(0)]_{i=0}^{i=1}=$$
$$=[L_{\dot x}(t_i,\ x(t_i),\ \dot x(t_i))\,X(t_i,\ x(t_i))+(L(t_i,\ x(t_i),\ \dot x(t_i))-$$
$$-L_{\dot x}(t_i,\ x(t_i),\ \dot x(t_i))\,\dot x(t_i))\,T(t_i,\ x(t_i))]_{i=0}^{i=1}=$$
$$=\varphi(t_1,\ x(t_1),\ \dot x(t_1))-\varphi(t_0,\ x(t_0),\ \dot x(t_0)).$$

Итак,

$$\varphi(t_1,\ x(t_1),\ \dot x_1(t_1))=\varphi(t_0,\ x(t_0),\ \dot x(t_0)).$$

Повторяя те же рассуждения для произвольного отрезка $[t_0,\ t]\subset[t_0,\ t_1]$, получаем равенство

$$\varphi(t,\ x(t),\ \dot x(t))\equiv\varphi(t_0,\ x(t_0),\ \dot x(t_0))\equiv\text{const}. \ \blacksquare$$

Следствие. *Если в условиях теоремы Нётер выполняются предположения* а) *и* б) *предложения* 2 п. 4.4.3, *то функция*

$$\psi(t,\ x,\ p)=pX(t,\ x)-\mathscr{H}(t,\ x,\ p)\,T(t,\ x) \tag{15}$$

постоянна на каждом решении $(x(t),\ p(t))$ *системы Гамильтона* (1) п. 4.4.3.

Доказательство предоставляется читателю. Из формулы (15) видно, что доставляемый теоремой Нётер первый интеграл есть значение дифференциальной формы Пуанкаре—Картана $\omega=p\,dx-\mathscr{H}\,dt$ на касательном векторном поле семейства $\{S_\alpha\}$.

В качестве примера покажем, как из теоремы Нётер получаются интегралы энергии и импульса (п. 1.4.1). Если L не зависит от t, то функционал (8) инвариантен относительно преобразований $(t,\ x)\mapsto(t+\alpha,\ x)$ (проверьте!). Здесь $T=1$, $X=0$, и потому $\varphi=L-L_{\dot x}\dot x=-\mathscr{H}$ постоянна. Если же L не зависит от одной из компонент вектора x, скажем от x_k, то функционал (8) инвариантен относительно преобразований $(t,\ x)\mapsto(t,\ x+\alpha e_k)$ (проверьте!). Здесь $T=0$, $X=e_k$, и потому $\varphi=L_{\dot x}e_k=L_{\dot x_k}$.

Другие примеры применения теоремы Нётер будут приведены в следующем пункте.

4.4.7. Вариационный принцип и законы сохранения в механике. Рассмотрим систему из n материальных точек («частиц») с массами $m_1,\ \ldots,\ m_n$, $r_i=(x_i,\ y_i,\ z_i)$ —

радиус-вектор i-й точки. Для краткости полагаем

$$r=(r_1, \ldots, r_n)\in\mathbb{R}^{3n}, \quad \frac{\partial}{\partial r}=\left(\frac{\partial}{\partial r_1}, \ldots, \frac{\partial}{\partial r_n}\right),$$
$$\frac{\partial}{\partial r_i}=\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i}\right).$$

Будем предполагать, что взаимодействие частиц между собой и с внешней средой описывается потенциальной энергией $U(t, r)$, так что уравнения движения имеют вид

$$m_i\ddot{r}_i=F_i=-\partial U/\partial r_i. \tag{1}$$

Кинетическая энергия системы определяется равенством

$$K=\frac{1}{2}\sum_{i=1}^{n} m_i(\dot{r}_i\,|\,\dot{r}_i). \tag{2}$$

Теорема (вариационный принцип Лагранжа). *Уравнения движения механической системы суть уравнения Эйлера вариационной задачи*

$$\int_{t_0}^{t_1} L(t, r, \dot{r})\,dt\to\mathrm{extr}, \qquad r(t_0)=r^0, \quad r(t_1)=r^1, \quad L=K-U. \tag{3}$$

Доказательство. Из (1) и (2) имеем

$$m_i\ddot{r}_i=\frac{d}{dt}(m_i\dot{r}_i)=\frac{d}{dt}\frac{\partial K}{\partial \dot{r}_i}=\frac{d}{dt}L_{\dot{r}_i}=L_{r_i}=\frac{\partial U}{\partial r_i}. \quad \blacksquare$$

Значение этой теоремы выходит далеко за рамки той простейшей ситуации, которой мы ограничимся. Физики и механики очень часто описывают свойства системы, задавая непосредственно ее лагранжиан L. При этом K, U и $L=K-U$ могут быть выражены не в декартовых, а в каких-то других координатах, в лагранжиан могут быть введены члены, описывающие магнитные или гироскопические (не потенциальные) силы и т. д. Как только лагранжиан задан, уравнения Эйлера вариационной задачи (3) определяют закон движения рассматриваемой системы.

Вариационный подход удобен, в частности, возможностью получения первых интегралов («законов сохранения») при помощи теоремы Нётер. Мы проиллюстрируем эту возможность на примере так называемых классических интегралов: энергии, импульса и момента.

а) Консервативная система: U не зависит от t. Группа преобразований: $(t, r) \mapsto (t+\alpha, r)$. Функция $r(\cdot)$: $[t_0, t_1] \to \mathbf{R}^{3n}$ переходит при этом в функцию $r_\alpha(\cdot)$: $[t_0+\alpha, t_1+\alpha] \to \mathbf{R}^{3n}$, $r_\alpha(t) = r(t-\alpha)$. Функционал (3) инвариантен:

$$\int\limits_{t_0+\alpha}^{t_1+\alpha} L(r_\alpha(t), \dot{r}_\alpha(t))\,dt = \int\limits_{t_0+\alpha}^{t_1+\alpha} L(r(t-\alpha), \dot{r}(t-\alpha))\,dt = \int\limits_{t_0}^{t_1} L(r(t), \dot{r}(t))\,dt.$$

Касательное векторное поле: $T=1$, $R=0$. Первый интеграл:

$$\varphi = \sum_{i=1}^{n} L_{\dot{r}_i} R_i - \left[\sum_{i=1}^{n} L_{\dot{r}_i}\dot{r}_i - L\right] T = -\sum_{i=1}^{n} m_i (\dot{r}_i \mid \dot{r}_i) + K - U = -K - U.$$

Следовательно, в консервативной системе полная энергия

$$\mathscr{H} = K + U$$

является первым интегралом («закон сохранения энергии»).

б) Свободная частица: U не зависит от радиуса-вектора r_k одной из частиц. Группа преобразований:

$$(t, r_1, \ldots, r_k, \ldots, r_n) \mapsto (t, r_1, \ldots, r_k + \alpha l, \ldots, r_n),$$

где $l \in \mathbf{R}^3$ произвольно. Инвариантность функционала (3) очевидна. Касательное векторное поле:

$$T=0, \quad R = (R_1, \ldots, R_k, \ldots, R_n) = (0, \ldots, l, \ldots, 0).$$

Первый интеграл:

$$\varphi = \sum_{i=1}^{n} L_{\dot{r}_i} R_i = \frac{\partial K}{\partial \dot{r}_k} l = m_k (\dot{r}_k \mid l).$$

Так как l произвольно, то $p_k = m_k \dot{r}_k$ должно быть постоянно. Таким образом, если k-я частица не взаимодействует с остальной частью системы (не вносит вклада в потенциальную энергию), то ее импульс $p_k = m_k \dot{r}_k$ сохраняется.

в) Внешние силы отсутствуют. Это означает, что U зависит только от разностей $r_i - r_j$ и L не меняется, если систему как единое твердое тело передвинуть в пространстве, т. е. сделать преобразование $(t, r_1, \ldots, r_n) \to (t, r_1 + \alpha l, \ldots, r_n + \alpha l)$, где l — произвольный вектор. Здесь $T = 0$, $R = (l, \ldots, l)$ и сохраняется $\sum_{i=1}^{n} \left(\frac{\partial L}{\partial \dot{r}_i} \middle| l \right) = \left(\sum_{i=1}^{n} m_i \dot{r}_i \middle| l \right)$. Так как l произвольно, то

$$P = \sum_{i=1}^{n} m_i \dot{r}_i = \text{const}.$$

Итак, если силы только внутренние, то имеет место *«закон сохранения полного импульса системы»*.

г) Вращательная симметрия: U зависит только от попарных расстояний $|r_i - r_j|$ между точками. Лагранжиан не меняется, если подвергнуть систему ортогональному преобразованию $\mathfrak{G}$, так как

$$L(\mathfrak{G} r_i, \mathfrak{G} \dot{r}_i) = K - U = \frac{1}{2} \sum_{i=1}^{n} m_i (\mathfrak{G} \dot{r}_i | \mathfrak{G} \dot{r}_i) - U(|\mathfrak{G} r_i - \mathfrak{G} r_j|) = \frac{1}{2} \sum_{i=1}^{n} m_i (\dot{r}_i | \dot{r}_i) - U(|r_i - r_j|)$$

(ортогональное преобразование не меняет скалярных произведений и длин). Фиксируем вектор ω и рассмотрим группу преобразований, соответствующую равномерному вращению системы около начала координат с угловой скоростью ω. Касательное векторное поле: $T = 0$, $R = (\omega \times r_1, \ldots, \omega \times r_n)$. Первый интеграл

$$\varphi = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{r}_i} R_i = \sum_{i=1}^{n} m_i (\dot{r}_i | \omega \times r_i) = \sum_{i=1}^{n} m_i (r_i \times \dot{r}_i | \omega).$$

Ввиду произвольности ω должен сохраняться *вектор момента количества движения*

$$M = \sum_{i=1}^{n} m_i r_i \times \dot{r}_i.$$

КОММЕНТАРИИ И ПУТЕВОДИТЕЛЬ ПО ЛИТЕРАТУРЕ

Литература по теории экстремальных задач огромна. Не претендуя на полноту, мы ограничились здесь упоминанием лишь некоторых изначальных работ, основных монографий, учебников и обзорных статей.

К главе I. § 1.1. История возникновения первых задач на максимум и минимум (помимо цитировавшейся книги Ван дер Вардена) изложена в [86]; о раннем этапе классического вариационного исчисления можно прочитать в [83] и [87]. Экстремальным свойствам круга и шара посвящена монография [21]. Транспортная задача исследовалась уже в работе Л. В. Канторовича [55] — первой работе по линейному программированию. Простейшие задачи автоматического регулирования были впервые исследованы Бушау [96]. Задача о быстродействии и ряд других подробно рассмотрены в основополагающей монографии Л. С. Понтрягина, В. Г. Болтянского, Р. В. Гамкрелидзе и Е. Ф. Мищенко [12] и во многих других книгах по оптимальному управлению ([23], [27] и др.).

§§ 1.2—1.5. История принципа Лагранжа рассказана в статье [45]. Этой же теме, но в рамках классического вариационного исчисления, посвящены работы [43], [44].

Укажем еще несколько книг и статей, посвященных затронутым в этих параграфах темам, и рассчитанных на широкого читателя: [71], [74], [76], [77], [115]. Основы классического вариационного исчисления изложены в учебниках [2], [3], [8].

К главе II. § 2.1. Помимо упоминавшихся уже учебников по функциональному анализу, укажем на книги [92], [107] специально ориентированные на «обслуживание» теории экстремальных задач.

§ 2.2. Дифференциальное исчисление в линейных нормированных пространствах (помимо [КФ]) изложено в учебниках [5], [56], [69] и монографиях [28], [54], [59].

§ 2.3. Бесконечномерные варианты теоремы о неявной функции изложены в учебниках [5], [56], [69].

Теорема п. 2.3.1 достаточно удобна для построения общей теории, ибо она содержит классическую теорему о неявной функции и теорему Люстерника о касательном пространстве и дает оценку отклонения от ядра отображения. Конструкции, использованные в доказательстве, фактически содержались в первоначальной работе [68] (см. также [69]). Другие доказательства и модификации см. в [54], [65].

§ 2.4. О дифференцируемости других важных конкретных функционалов см. в [28], [59].

§ 2.6. Основы конечномерного выпуклого анализа были заложены Минковским [113], [114] и Фенхелем [101], [102]. Выпуклый анализ в бесконечномерных пространствах был построен в шестидесятые годы в работах Бронстеда [95], А. Я. Дубовицкого и А. А. Милютина [46], Моро, Рокафеллара и др. Наиболее полный обзор конечномерной теории содержится в монографии Рокафеллара [81], а также (в части, касающейся лишь выпуклых множеств) в монографии Боннезена и Фенхеля [94]. Бесконечномерная теория изложена в [16], [52], [54], [62], [79], [84].

В самое последнее время делаются попытки создать синтетическое, «гладко-выпуклое» исчисление—см. [70], [99], [118].

К главе III. § 3.2. Трудно сказать, кем впервые было доказано правило множителей Лагранжа для гладких конечномерных задач с равенствами и неравенствами. В некоторых американских работах дается ссылка на работу Валентайна [120] и диссертацию Каруша [109]. Правило множителей для случая бесконечного числа неравенств было доказано Джоном [108]. Очень большую роль в этой тематике сыграла работа Куна и Таккера [111]. Бесконечномерные варианты правила множителей изложены в [37], [39], [46], [54], [79], [84].

§ 3.3. Линейное и выпуклое программирование очень широко представлено учебной и монографической литературой на русском языке: [38], [40], [41], [50], [57], [58], [67], [75], [76], [81], [89], [90].

§ 3.4. Одной из первых работ, посвященных необходимым условиям для задач с ограничениями типа неравенств, была [100]. Конечномерная теория прекрасно изложена в учебнике Хестенеса [106], см. также [51], [85]. В недавнее время Е. С. Левитиным, А. А. Милютиным и Н. П. Осмоловским была разработана полная теория условий второго порядка для задач с ограничениями [64], [65]. В последней из названных работ имеется подробная библиография. В нашем изложении использовались некоторые конструкции этой работы. См. также [105], [117].

К главе IV. §§ 4.1 и 4.4. Классическому вариационному исчислению посвящено много учебников и монографий. Кроме упомянутых выше см. [20], [63], [91], [93], [97], [103]. Наиболее полно теория необходимых и достаточных условий экстремума в задаче Лагранжа разработана в монографии Блисса [20]. Там же изложена история вопроса. О связях вариационного исчисления и классической механики см. [1].

§ 4.2. Первоначальный набросок теории оптимального управления был изложен в [26] и в обзорной статье Л. С. Понтрягина [78], затем теория оптимального управления составила содержание монографии Л. С. Понтрягина, В. Г. Болтянского, Р. В. Гамкрелидзе и Е. Ф. Мищенко [12], давшей толчок бурному развитию всего направления исследований, связанных с экстремальными задачами. Доказательство принципа максимума, изложенное в § 4.2, является обработкой первоначального доказательства В. Г. Болтянского, хотя в нем не используются никакие иные средства, кроме классического анализа. Метод центрированных систем использовался в работах А. Я. Дубовицкого и А. А. Милютина. Сейчас имеется много доказательств принципа максимума, см., например, [24], [31], [46]—[49], [54], [104], [112].

Теории оптимального управления посвящены учебники [23], [31], [35], а также монографии [25], [27], [32], [61], [66], [73] и ряд других. В назвайных монографиях много места уделено решению разнообразных конкретных прикладных задач.

§ 4.3. Принцип максимума для линейных систем был впервые доказан Р. В. Гамкрелидзе [36]. Завершенная теория линейиых систем была разработана Н. Н. Красовским [60]. Обзору проблематики, связанной с ляпуновскими задачами, посвящена статья [17], где содержатся обобщения результатов этого параграфа и имеется подробная библиография.

Развитие теории ляпуновских задач привело к изучению выпуклых интегральных функционалов в работах [54], [98], [119] и др.

В заключение отметим несколько монографий, в которых читатель сможет ознакомиться с некоторыми принципиальными вопросами теории экстремальных задач, не затронутыми в книге.

Многомерное вариационное исчисление: [110], [116].

Теоремы существования: [35], [54], [116].

Достаточные условия в задачах оптимального управления: [22], [23], [61].

Расширения экстремальных задач: [53], [88].

Двойственные методы в теории экстремальных задач: [52], [88].

Численные методы: [29], [42], [72], [80], [82].

Скользящие режимы: [34].

Динамическое программирование: [18], [19].

Фазовые и смешанные ограничения: [33], [47], [48].

См. также обзор [30].

ЛИТЕРАТУРА

Учебники и учебные пособия, цитируемые в основном тексте

КФ. Колмогоров А. Н., Фомин С. В. Элементы теории функций н функционального анализа.—М.: Наука, 1976.

1. Арнольд В. И. Математические методы классической механикн.—М.: Наука, 1974.
2. Ахиезер Н. И. Лекции по вариационному исчислению.—М.: Гостехиздат, 1955.
3. Гельфанд И. М., Фомин С. В. Вариационное исчисление.—М.: Физматгиз, 1961.
4. Данфорд Н., Шварц Дж. Линейные операторы. / Общая теория.—М.: ИЛ, 1962.
5. Картан А. Дифференциальное исчисление. Дифференциальные формы.—М.: ИЛ, 1971.
6. Коддингтон Э. А., Левинсон Н. Теория обыкновенных дифференциальных уравнений.—М.: ИЛ, 1958.
7. Курош А. Г. Курс высшей алгебры.—М.: Наука, 1971.
8. Лаврентьев М. А., Люстерник Л. А. Курс вариационного исчисления.—М.-Л.: Гостехиздат, 1950.
9. Никольский С. М. Курс математического анализа, тт. 1, 2.—М.: Наука, 1975.
10. Петровский И. Г. Лекции по теории обыкновенных дифференциальных уравнений.—М.-Л.: Гостехиздат, 1952.
11. Понтрягин Л. С. Обыкиовенные дифференциальные уравнения.—М.: Наука, 1965.
12. Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов.—М.: Физматгиз, 1961.
13. Привалов И. И. Введение в теорию функций комплексного перемениого.—М.: Наука, 1977.
14. Фихтенгольц Г. М. Курс дифференциального и интегрального исчисления, тт. 1, 2.—М.: Наука, 1969.
15. Хартман Ф. Обыкновенные дифференциальные уравнения.—М.: ИЛ, 1970.

Дополнительная литература

16. Акилов Г. П., Кутателадзе С. С. Упорядоченные векторные пространства.—Новосибирск: Наука, 1978.
17. Аркин В. И., Левин В. Л. Выпуклость значений векторных интегралов, теоремы измеримого выбора и вариационные задачи.—УМН **27**, вып. 3, 1972, с. 21—77.
18. Беллман Р. Динамическое программирование.—М.: ИЛ, 1960.
19. Беллман Р., Калаба Р. Динамическое программирование и современная теория управления.—М.: Наука, 1969.
20. Блисс Дж. Лекции по вариационному исчислению.—М.: ИЛ, 1950.
21. Бляшке В. Круг и шар.—М.: Наука, 1967.
22. Болтянский В. Г. Достаточные условия оптимальности и обоснование метода динамического программирования.— Изв. АН СССР, сер. матем. **28**, № 3, 1964, с. 481—514.
23. Болтянский В. Г. Математические методы оптимального управления.—М.: Наука, 1969.
24. Болтянский В. Г. Метод шатров в теории экстремальных задач.—Успехи матем. наук **30**, вып. 3, 1975, с. 3—55.
25. Болтянский В. Г. Оптимальное управление дискретными системами.—М.: Наука, 1973.
26. Болтянский В. Г., Гамкрелидзе Р. В., Понтрягин Л. С. К теории оптимальных процессов.—ДАН СССР **110**, № 1, 1956, с. 7—10.
27. Брайсон А., Хо Ю-ши. Прикладная теория оптимального управления.—М.: Мир, 1972.
28. Вайнберг М. М. Вариационный метод и метод монотонных операторов в теории нелинейных уравнений.—М.: Наука, 1972.
29. Васильев Ф. П. Лекции по методам решения экстремальных задач.—М.: Изд. МГУ, 1974.
30. Габасов Р., Кириллова Ф. М. Методы оптимального управления.—Итоги науки и техники. Современные проблемы математики. Т. 6.—М.: 1976, с. 133—206.
31. Габасов Р., Кириллова Ф. М. Методы оптимизации.—Минск: Изд-во БГУ, 1975.
32. Габасов Р., Кириллова Ф. М. Особые оптимальные управления.—М.: Наука, 1973.
33. Гамкрелидзе Р. В. Оптимальные процессы управления при ограниченных фазовых координатах.—Изв. АН СССР, сер. матем. **24**, № 3, 1960, с. 315—356.
34. Гамкрелидзе Р. В. О скользящих оптимальных режимах.—ДАН СССР **143**, № 6, 1962, с. 1243—1245.
35. Гамкрелидзе Р. В. Основы оптимального управления.—Тбилиси: Изд. ТГУ, 1977.
36. Гамкрелидзе Р. В. Теория оптимальных по быстродействию процессов в линейных системах.—Изв. АН СССР, сер. матем. **22**, № 4, 1958, с. 449—474
37. Гамкрелидзе Р. В., Харатишвили Г. Л. Необходимые условия первого порядка в экстремальных задачах. / В кн.: Международный конгресс математиков в Ницце.—М.: Наука, 1972.
38. Гасс С. Линейное программирование.—М : Физматгиз, 1961.
39. Гирсанов И. В. Лекции по математической теории экстремальных задач.—М.: Изд-во МГУ, 1970.

40. Гольштейн Е. Г. Теория двойственности в математическом программировании. — М.: Наука, 1971.
41. Данциг Д. Линейное программирование, его обобщения и применения. — М.: Прогресс, 1966.
42. Демьянов В. Ф., Рубинов А. М. Приближенные методы решения экстремальных задач. — Л.: Изд-во ЛГУ, 1968.
43. Дорофеева А. В. Вариационное исчисление во второй половине XIX в. — Историко-математические исследования XV, 1963, с. 99—128.
44. Дорофеева А. В. Развитие вариационного исчисления, как исчисления вариаций. — Историко-математические исследования XIV, 1961, с. 101—180.
45. Дорофеева А. В., Тихомиров В. М. От правила множителей Лагранжа до принципа максимума Понтрягина. — Историко-математические исследования XXV, 1979. — (в печати).
46. Дубовицкий А. Я., Милютин А.А. Задачи на экстремум при наличии ограничений. — ЖВМ и МФ 5, № 3, 1965, с. 395—453.
47. Дубовицкий А. Я., Милютин А. А. Необходимые условия слабого экстремума в общей задаче оптимального управления. — М.: Наука, 1971.
48. Дубовицкий А. Я., Милютин А. А. Необходимые условия экстремума в задачах оптимального управления со смешанными ограничениями типа неравенств. — ЖВМ и МФ 8, № 4, 1968, с. 725—770.
49. Дубовицкий А. Я., Милютин А. А. Трансляция уравнений Эйлера. — ЖВМ и МФ 9, № 6, 1969, с. 1263—1284.
50. Еремин И. И., Астафьев Н. Н. Введение в теорию линейного и выпуклого программирования. — М.: Наука, 1976.
51. Зангвилл У. И. Нелинейное программирование. — М.: Сов. радио, 1973.
52. Иоффе А. Д., Тихомиров В. М. Двойственность выпуклых функций и экстремальные задачи. — УМН 23, вып. 6, 1968, с. 51—116.
53. Иоффе А. Д., Тихомиров В. М. Расширение вариационных задач. — Труды ММО 18, 1968, с. 187—246.
54. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. — М.: Наука, 1974.
55. Канторович Л. В. Математические методы организации и планирования производства. — Л.: Изд-во ЛГУ, 1939.
56. Канторович Л. В., Акилов Г. П. Функциональный анализ. — М.: Наука, 1977.
57. Карлин С. Математические методы в теории игр, программировании и экономике. — М.: ИЛ, 1964.
58. Карманов В. Г. Математическое программирование. — М.: Наука, 1975.
59. Красносельский М. А., Забрейко П. П., Пустыльник Е. И., Соболевский П. Е. Интегральные операторы в пространствах суммируемых функций. — М.: Наука, 1966.
60. Красовский Н. Н. Теория управления движением. Линейные системы. — М.: Наука, 1968.
61. Кротов В. Ф., Гурман В. И. Методы и задачи оптимального управления. — М: Наука, 1973.
62. Кутателадзе С. С., Рубинов А. М. Двойственность Минковского и ее приложения. — Новосибирск: Наука, 1976.

63. Лаврентьев М. А., Люстерник Л. А. Основы вариационного исчисления. — М.-Л.: ОНТИ, 1935. — ТТ. I, II.
64. Левитин Е. С., Милютин А. А., Осмоловский Н. П. Об условиях локального минимума в задачах с ограничениями./ В кн. Математическая экономика и функциональный анализ. — М.: Наука, 1974, с. 139—202.
65. Левитин Е. С., Милютин А. А., Осмоловский Н. П. Условия высших порядков локального минимума в задачах с ограничениями. — УМН 33, вып. 6, 1978, с. 85—148.
66. Ли Э. Б., Маркус Л. Основы теории оптимального управления. — М.: Наука, 1972.
67. Линейные неравенства и смежные вопросы./Под ред. Г. Куна и А. Таккера. — М.: ИЛ, 1959.
68. Люстерник Л. А. Об условных экстремумах функционалов. — Матем. сб. 41, № 3, 1934, с. 390—401.
69. Люстерник Л. А., Соболев В. И. Элементы функционального анализа. — М.: Наука, 1965.
70. Магарил-Ильяев Г. Г. Теорема о неявной функции для липшицевых отображений. — УМН 33, вып. 1, 1978, с. 221—222.
71. Математика на службе инженера./Под ред. Н. Х. Розова. — М.: Знание, 1973.
72. Моисеев Н. Н. Численные методы в теории оптимальных систем. — М.: Наука, 1971.
73. Моисеев Н. Н. Элементы теории оптимальных систем. — М.: Наука, 1975.
74. Мышкис А. Д. Математика: Специальные курсы для втузов. — М.: Наука, 1971.
75. Никайдо Х. Выпуклые структуры и математическая экономика. — М.: Мир, 1972.
76. Нит И. В. Линейное программирование (с обсуждением некоторых нелинейных задач). — М.: Изд-во МГУ, 1978.
77. Оптимальное управление. — М.: Знание, 1978.
78. Понтрягин Л. С. Оптимальные процессы регулирования. — УМН 14, вып. 1, 1959, с. 3—20.
79. Пшеничный Б. Н. Необходимые условия экстремума. — М.: Наука, 1969.
80. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. — М.: Наука, 1975.
81. Рокафеллар Р. Выпуклый анализ. — М.: Мир, 1973.
82. Романовский И. В. Алгоритмы решения экстремальных задач. — М.: Наука, 1977.
83. Рыбников К. А. Первые этапы вариационного исчисления./ В кн.: Историко-математические исследования, вып. 2. — М.-Л.: ГИТТЛ, 1949.
84. Тихомиров В. М. Некоторые вопросы теории приближений. — М.: Изд-во МГУ, 1976.
85. Фиакко А., Мак-Кормик Г. Нелинейное программирование. Методы последовательной безусловной минимизации. — М.: Мир, 1972.
86. Цейтен Г. Г. История математики в древности и в средние века. — М.-Л.: ОНТИ, 1938.
87. Цейтен Г. Г. История математики в 16 и 17 столетиях. — М.-Л.: ОНТИ, 1938.
88 Экланд И., Темам Р. Выпуклый анализ и вариационные проблемы. — М.: Мир, 1979.

89. Эрроу К. Дж., Гурвиц Л., Удзава Х. Исследования по линейному и нелинейному программированию.—М.: ИЛ, 1962.
90. Юдин Д. Б., Гольштейн Е. Г. Линейное программирование (теория, методы, приложения).—М.: Наука, 1969.
91. Янг Л. Лекции по вариационному исчислению и теории оптимального управления.—М.: Мир, 1974.
92. Balakrishnan A. V. Applied Functional Analysis. N. Y.: Springer, 1976.
93. Bolza O. Vorlesungen über Variationsrechnung.—Leipzig: 1949.
94. Bonnesen T., Fenchel W. Theorie der konvexen Körper.—B.: Springer, 1934.
95. Brøndsted A. Conjugate convex functions in topological vector spaces.—Mat. Fys. Medd. Dansk. Vid. Selsk. **34**, 2, 1964, p. 1—26.
96. Bushaw D. W. Differential equations with a discontinuous forcing term.—Princeton: Dept. of Math. Princeton Univ., 1952.
97. Carathéodory C. Variationsrechnung und partielle Differentialgleichungen erster Ordnung.—Leipzig—Berlin: Teubner, 1935.
98. Castaing C., Valadier M. Convex analysis and measurable multifunctions. Lecture Notes in Mathematics, № 580.—B.: Springer, 1977.
99. Clarke F. H. Generalized gradients and applications.—Trans. Amer. Math. Soc. **205**, 1975, p. 247—262.
100. Cox M. J. On necessary conditions for relative minima.—Amer. J. Math. **66**, № 2, 1944, p. 170—198.
101. Fenchel W. Convex Cones, Sets and Functions.—Princeton: Princeton Univ., 1951.
102. Fenchel W. On conjugate convex functions.—Canadian J. Math. **1**, 1949, p. 73—77.
103. Hadamard J. Calcul des variations.—Paris: Hermann, 1910.
104. Halkin H. On necessary conditions for optimal control of nonlinear systems.—J. Analyse Math. **12**, 1964, p. 1—82.
105. Hestenes M. R. Calculus of Variations and Optimal Control Theory.—N. Y.: Wiley, 1966.
106. Hestenes M. R. Optimization theory. The finite dimensional case.—N. Y.: Wiley, 1975.
107. Holmes R. B. Geometric functional analysis and its applications.—N. Y.: Springer, 1975.
108. John F. Extremum problems with inequalities as subsidiary conditions./In: Studies and Essays. Courant Anniversary Volume.—N. Y.: Interscience, 1948, p. 187—204.
109. Karush W. E. Minima of functions of several variables with inequalities as side conditions.—Chicago: Univ. of Chicago, 1939.
110. Klötzler R. Mehrdimensionale Variationsrechnung.—B.: VEB Deutscher Verlag, 1971.
111. Kuhn H. W., Tucker A. W. Nonlinear programming./In: Proc. of Second Berkeley Symp.—Berkeley: Univ. of California Press, 1951, p. 481—492.
112. Michel P. Une démonstration élémentaire du principe du maximum de Pontriaguine.—Bull. Math. économiques. **14**, 1977.
113. Minkowski H. Geometrie der Zahlen.—Leipzig: Teubner, 1910.
114. Minkowski H. Theorie der konvexen Körper./In: Gesammelte Abhandlungen, Bd. II.—Leipzig-Berlin: Teubner, 1911.
115. Moisseyev N. N., Tihomirov V. M. Optimisation./In: Mathematics Applied to Physics.—N. Y.: Springer, 1970.

116. M o r r e y Ch. B. Multiple integrals in the calculus of variations. — N. Y.: Springer, •1966.
117. N e u s t a d t L. W. Optimization. A theory of necessary conditions. — Princeton: Princeton Univ. Press, 1976.
118. R o c k a f e l l a r R. T. The theory of subgradients and its applications to problems of optimization. — Lecture Notes Univer. of Montreal., 1978.
119. R o c k a f e l l a r R. T. Convex — Integral functuonals and duality./In: Contributions to Nonlinear Functional Analysis. — N. Y.: Acad. Press, 1971, p. 215—236.
120. V a l e n t i n e F. A. The problem of Lagrange with differential inequalities as added side conditions./In: Contributions to Calculus of Variations. — Chicago: Univ. of Chicago Press, 1933—1937.

СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

$\forall$ — квантор общности: «для всех»

$\exists$ — квантор существования: «существует»

$\Longrightarrow$ — знак импликации: «из ... следует ...»

$\Longleftrightarrow$ — знак эквивалентности

$\overset{\text{def}}{=}$ — равно по определению

$\overset{(4)}{=}$ — равно в силу (4)

$x \in A$ — элемент x принадлежит множеству A

$x \notin A$ — элемент x не принадлежит множеству A

$\varnothing$ — пустое множество

$A \cup B$ — объединение множеств A и B

$A \cap B$ — пересечение множеств A и B

$A \setminus B$ — разность множеств A и B

$A \subset B$ — множество A содержится в множестве B

$A \times B$ — декартово произведение множеств A и B

$A + B$ — арифметическая сумма множеств A и B

$\{x \mid P(x)\}$ — множество элементов x, обладающих свойством $P(\cdot)$

$\{x_1, \ldots, x_n, \ldots\}$ — множество, состоящее из элементов $x_1, \ldots, x_n, \ldots$

$F\colon X \longrightarrow Y$ — отображение F множества X в множество Y; функция F с областью определения X, принимающая значения во множестве Y

$x \longmapsto F(x)$ — отображение (функция) F сопоставляет элементу x элемент $F(x)$; обозначение для отображения (функции) F в случае, когда желательно указать обозначение его (ее) аргумента

$F(\cdot)$ — обозначение, которым подчеркивается, что F является отображением (функцией)

$F(A)$ — образ множества A при отображении F

$\operatorname{im} F = \{y \mid y = F(x),\quad x \in X\}$ — образ отображения $F\colon X \longrightarrow Y$

$F^{-1}(A)$ — прообраз множества A при отображении F

$F \mid A$ — ограничение отображения F на множество A

$F \circ G$ — суперпозиция отображений G и F: $(F \circ G)(x) = F(G(x))$

$\mathbf{R}$ — множество всех действительных чисел; числовая прямая

$\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$ — расширенная числовая прямая

$\inf A$ ($\sup A$) — нижняя (верхняя) грань чисел, входящих в множество $A \subset \mathbf{R}$

$\mathbf{R}^n$ — арифметическое n-мерное пространство, наделенное стандартной евклидовой структурой, элементы $\mathbf{R}^n$ следует представлять себе в виде вектор-столбцов, даже если они по типографским соображениям набраны в строчку

$\mathbf{R}^n_+ = \{x = (x_1, \ldots, x_n) \in \mathbf{R}^n \mid x_i \geqslant 0\}$ — неотрицательный ортант в $\mathbf{R}^n$

$e_1, \ldots, e_n$ — векторы стандартного ортонормированного базиса в $\mathbf{R}^n$: $e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 1)$

$\mathbf{R}^{n*}$ — арифметическое n-мерное пространство, сопряженное к $\mathbf{R}^n$; элементы $\mathbf{R}^{n*}$ следует представлять себе как вектор-строчки

$$px = \langle p, x \rangle = \sum_{i=1}^{n} p_i x_i \text{ для любых } x \in \mathbf{R}^n \text{ и } p \in \mathbf{R}^{n*}$$

x^{T} — вектор-строчка, транспонированная к вектор-столбцу x

$|x|$ — евклидова норма в $\mathbf{R}^n$

$|x|^2 = (x \mid x) = x^{\mathrm{T}} x$

$$(x \mid y) = \sum_{i=1}^{n} x_i y_i \text{ — скалярное произведение в } \mathbf{R}^n$$

$\|x\|$ — норма элемента x в нормированном пространстве

$\rho(x, y)$ — расстояние от элемента x до элемента y

$\rho(x, A) = \inf \{\rho(x, y) \mid y \in A\}$ — расстояние от элемента x до множества A

$B(x, r) = \{y \mid \rho(y, x) \leqslant r\}$ — замкнутый шар с центром x радиуса r

$\mathring{B}(x, r) = \{y \mid \rho(y, x) < r\}$ — открытый шар с центром x радиуса r

$\overline{A}$ — замыкание множества A

$\operatorname{int} A$ — внутренность множества A

$T_x M$ ($T_x^+ M$) — множество касательных векторов (односторонних касательных векторов) к множеству A в точке x

$\mathscr{L}_a f = \{x \mid f(x) \leqslant a\}$ — лебеговское множество функции f на уровне a

X^* — пространство, сопряженное с X

x^* — элемент сопряженного пространства X^*

$\langle x^*, x \rangle$ — значение линейного функционала $x^* \in X^*$ на элементе $x \in X$

$(x \mid y)$ — скалярное произведение элементов x и y гильбертова пространства

$A^{\perp} = \{x^* \mid \langle x^*, x \rangle = 0, \ x \in A\}$ — аннулятор множества A

$\dim L$ — размерность пространства L

X/L — фактор-пространство пространства X по подпространству L

$\mathscr{L}(X, Y)$ — пространство непрерывных линейных отображений пространства X в пространство Y; отображения из пространства $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ могут отождествляться с их матрицами относительно стандартных базисов в $\mathbf{R}^n$ и $\mathbf{R}^m$

$\mathscr{L}^n(X, Y)$ — пространство непрерывных полилинейных отображений пространства $X \times \ldots \times X$ в пространство Y

I — единичный оператор; E или E_n — матрица единичного оператора из $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$, т. е. единичная матрица n-го порядка

Λ^* — оператор, сопряженный с оператором Λ; $\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle$

$\operatorname{Ker}\Lambda = \{x \mid \Lambda x = 0\}$ — ядро линейного оператора

$\operatorname{Im}\Lambda = \{y \mid y = \Lambda x\}$ — образ линейного оператора

$x^{\mathrm{T}}Ax$ $(y^{\mathrm{T}}Ax)$ — обозначение для квадратичной (билинейной) формы с матрицей A: $y^{\mathrm{T}}Ax = \sum_{i,j=1}^{n} A_{ij}y_i x_j$.

$A > 0$ — матрица A положительно определена

$A \geqslant 0$ — матрица A неотрицательно определена

$C(K, Y)$ — пространство непрерывных отображений из K в Y; если Y нормировано, то для $x(\cdot) \in C(K, Y)$

$\|x(\cdot)\| = \|x(\cdot)\|_0 = \sup_{t \in K}\|x(t)\|$

$C(K) = C(K, Y)$, если ясно, о каком Y идет речь, или если $Y = \mathbf{R}$

$C([t_0, t_1])$ — пространство непрерывных функций на отрезке $[t_0, t_1]$

$C^r(U)$ — множество отображений, имеющих в U непрерывные производные до порядка r включительно; обычно употребляется в выражениях: «функция f принадлежит классу $C^r(U)$» или «отображение $F \in C^r(U)$»

$C^r(\Delta, Y)$ — пространство отображений отрезка $\Delta \subset \mathbf{R}$ в пространство Y (обычно $Y = \mathbf{R}^n$ или $\mathbf{R}^{n*}$), имеющих непрерывные производные до порядка r включительно; для $x(\cdot) \in C^r(\Delta, Y)$

$\|x(\cdot)\| = \|x(\cdot)\|_r = \max\{\|x(\cdot)\|_0, \|\dot{x}(\cdot)\|_0, \ldots, \|x^{(r)}(\cdot)\|_0\}$

$KC^1(\Delta, Y)$ — пространство отображений отрезка $\Delta \subset \mathbf{R}$ в пространство Y (обычно $Y = \mathbf{R}^n$ или $\mathbf{R}^{n*}$), имеющих кусочно-непрерывную производную. Снабжено нормой как подпространство пространства $C(\Delta, Y)$

$W^1_\infty(\Delta, Y)$ — пространство отображений отрезка $\Delta \subset \mathbf{R}$ в пространство Y (обычно $Y = \mathbf{R}^n$ или $\mathbf{R}^{n*}$), каждое из которых удовлетворяет условию Липшица с некоторой константой Снабжено нормой как подпространство пространства $C(\Delta, Y)$

$F'(x; h)$ — производная функции F в точке x по направлению вектора h

$\delta_+ F(x;\ \cdot)$ — первая вариация отображения F в точке x

$\delta F(x;\ \cdot)$ — первая вариация по Лагранжу отображения F в точке x

$\delta^n F(x;\ \cdot)$ — n-я вариация по Лагранжу отображения F в точке x

$F'_{\Gamma}(x)$ — производная отображения F в точке x по Гато

$F'(x)$ — производная отображения F в точке x по Фреше

$F'(x)[h]$ или $F'(x)h$ — значение производной отображения F в точке x на векторе h

$F \in D^1(x)$ — отображение F дифференцируемо по Фреше в точке x

$F \in SD^1(x)$ — отображение F строго дифференцируемо в точке x

$F''(x)$ — вторая производная отображения F в точке x по Фреше

$F''(x)[h_1,\ h_2]$ — значение второй производной (как билинейной функции) на паре векторов h_1 и h_2

$h \mapsto d^2F(x;\ h) = F''(x)[h,\ h]$ — второй дифференциал отображения F в точке x

$F_{x_1}(x_1,\ x_2)$ $(F_{x_2}(x_1,\ x_2))$ — частная производная по x_1 (по x_2) отображения F, т. е. производная отображения $x_1 \mapsto F(x_1,\ x_2)$ $(x_2 \mapsto F(x_1,\ x_2))$

$[x_1,\ x_2] = \{x \mid x = \alpha x_1 + (1-\alpha) x_2,\quad 0 \leqslant \alpha \leqslant 1\}$ — отрезок, соединяющий точки x_1, x_2

$\mathrm{dom}\, f = \{x \in X \mid f(x) < +\infty\}$ — эффективное множество функции $f\colon X \to \overline{\mathbf{R}}$

$\mathrm{epi}\, f = \{(x,\ \alpha) \in X \times \mathbf{R} \mid f(x) \leqslant \alpha,\ x \in \mathrm{dom}\, f\}$ — надграфик функции $f\colon X \to \overline{\mathbf{R}}$

δA — индикаторная функция множества A

sA — опорная функция множества A

μA — функция Минковского множества A

$\mathrm{lin}\, A$ — линейная оболочка множества A

$\mathrm{aff}\, A$ — аффинная оболочка множества A

$\mathrm{cone}\, A$ — коническая оболочка множества A

$\mathrm{conv}\, A$ — выпуклая оболочка множества A

$\overline{\mathrm{conv}}\, A$ — выпуклое замыкание, т. е замыкание выпуклой оболочки множества A

$\partial f(x)$ — субдифференциал функции f в точке x

$f^*(\ )$ — функция, сопряженная с функцией $f(\ )$, преобразование Юнга — Фенхеля функции f

$f(x) \to \inf\ (\sup,\ \mathrm{extr}),\ x \in C$ — обозначение экстремальной задачи

$\hat{x}$ — решение экстремальной задачи

absmin (absmax, absextr) — абсолютный минимум (максимум, экстремум)

$\hat{x} \in \mathrm{absmin}\,(\mathfrak{z})$ $(\mathrm{absmax}\,(\mathfrak{z}),\ \mathrm{absextr}\,(\mathfrak{z}))$ — $\hat{x}$ доставляет абсолютный минимум (максимум, экстремум) в задаче $\mathfrak{z}$

locmin (locmax, locextr) — локальный минимум (максимум, экстремум)

$\hat{x} \in \mathrm{locmin}\,(\mathfrak{z})$, $(\mathrm{locmax}\,(\mathfrak{z})$, $\mathrm{locextr}\,(\mathfrak{z}))$ — $\hat{x}$ доставляет минимум (максимум, экстремум) в задаче $\mathfrak{z}$

$\mathcal{J}$, $\mathcal{G}$, $\mathcal{B}$, $\mathcal{F}$, $\mathcal{P}$ — функционалы в задачах вариационного исчисления и оптимального управления

Q — квадратичный функционал вторичной задачи вариационного исчисления

$L = L(t, x, \dot{x})$ или $L = L(t, x, \dot{x}, u)$ — лагранжиан

$\mathcal{L}$ — функция Лагранжа

$\mathcal{H}$ — гамильтониан

H — функция Понтрягина

$\mathcal{E}$ — функция Вейерштрасса

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Алексеев Владимир Михайлович,
Тихомиров Владимир Михайлович,
Фомин Сергей Васильевич

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

М., 1979 г., 432 стр. с илл.

Редактор *Н. Н. Васина*

Техн. редактор *И. Ш. Аксельрод*

Корректор *О. М. Кривенко*

ИБ № 2344

Сдано в набор 20 06.79. Подписано к печати 25.10 79. Т-17355. Бумага $84\times108^{1}/_{32}$ Тип № 2. Высокая печать. Литературная гарнитура. Условн. печ л. 22,68. Уч.-изд л 24,64. Тираж 24 000 экз. Заказ № 314. Цена книги 1 р 10 к.

Издательство «Наука»
Главная редакция
физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Октябрьской Революции и ордена Трудового Красного Знамени Первая Образцовая типография имени А. А. Жданова Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли.
Москва, М-54, Валовая, 28