О необходимых условиях минимума

А. Д. ИОФФЕ
Технион — Израильский технологический институт
e-mail: alexander.ioffe38@gmail.com

УДК 517.97

Ключевые слова: оптимизационная альтернатива, релаксация, задача Больцы, оптимальное управление, дифференциальное включение.

Аннотация

В статье обсуждается общий подход к выводу необходимых условий в экстремальных задачах, основанный на так называемой «оптимизационной альтернативе», позволяющей свести задачу с ограничениями к одной задаче или последовательности задач на безусловный минимум. Используя предложенный формализм, мы доказываем необходимое условие экстремума в некоторой весьма общей задаче со смешанной (частично выпуклой) структурой, а затем предлагаем новое доказательство принадлежащего Ф. Кларку «стратифицированного» принципа максимума для задач оптимального управления с дифференциальными включениями.

Abstract

A. D. Ioffe, On necessary conditions for a minimum, Fundamentalnaya i prikladnaya matematika, vol. 19 (2014), no. 4, pp. 121—152.

We discuss a general approach to necessary optimality conditions based on so called “optimality alternative” that reduces a problem with constraints to one or a sequence unconstrained problems. The power of the approach is demonstrated by proofs of a necessary optimality condition in an abstract problem with mixed (convex vs. nonconvex) structure and a new proof of Clarke’s “stratified” maximum principle for optimal control of differential inclusions.

В. М. Тихомирову,
моему учителю, коллеге и другу,
с благодарностью и любовью

1. Введение

Цель настоящего краткого обзора — обсудить новый подход к выводу необходимых условий экстремума в (общем случае негладких) экстремальных задачах, использующий технические возможности, предоставляемые современным вариационным анализом. В классической теории необходимых условий экстремума центральное место бесспорно занимает восходящий к Эйлеру и Лагранжу метод вариаций (хотя, конечно, для частных задач с успехом использовались и более прямые методы, подобные прямым методам вариационного исчисления, или метод штрафных функций, впервые применённый, видимо, Р. Курантом в 1943 г.), который схематически можно представить с помощью следующей диаграммы:

$$\boxed{\text{исходная задача}} \quad\Longrightarrow\quad \boxed{\text{линеаризация}} \quad\Longrightarrow\quad \boxed{\text{необходимые условия}}\,.$$

Можно указать два основных способа аналитического воплощения метода вариаций. В рамках первого, который можно охарактеризовать как лагранжев метод или «принцип Лагранжа», целевая функция и все или некоторые функциональные ограничения объединяются в функции Лагранжа и необходимые условия в исходной задаче реализуются как необходимые условия минимума функции Лагранжа при не вошедших в неё ограничениях (см., например, $[1,7,33]$). При втором подходе, формализованном в знаменитой работе А. Я. Дубовицкого и А. А. Милютина [5], необходимые условия трактуются как условия непересечения некоторой системы множеств и независимо друг от друга анализируются целевая функция и все ограничения. Если говорить о необходимых условиях первого порядка — а здесь мы рассматриваем только такие условия, — оба подхода в классической теории в общем оказываются равносильными. Оба с успехом применялись к одним и тем же задачам и приводили к тождественным результатам.

Современный вариационный анализ предлагает два подхода к анализу необходимых условий оптимальности в негладких задачах, так или иначе связанные с идеей метрической регулярности. Как и в гладком случае, в одном из них целевая функция и все или некоторые выбранные ограничения рассматриваются как единый объект, в то время как второй предоставляет свободу независимого рассмотрения целевой функции и всех ограничений. При этом, как кажется, второй подход оказывается несколько более универсальным. Замечательно однако, что каждый из них, применённый в классической ситуации, существенно упрощает доказательство, полностью исключая необходимость варьирования и перехода к линеаризованной задаче. Более того, в то время как в рамках первого подхода явно используются те или иные инфинитезимальные критерии регулярности, второй подход не нуждается даже в этом.

Первый подход опирается на критерии накрывания или метрической регулярности, обобщающие теорему Люстерника—Грейвса. (Приведём стандартную формулировку этой теоремы: если $F\colon X \to Y$ строго дифференцируемо в $\bar{x}$, то F накрывает в окрестности $\bar{x}$, если (и только если, на самом деле) выполнено условие Люстерника $\operatorname{Im} F'(\bar{x}) = Y$. Теорема Грейвса, на самом деле, несколько сильнее. Она не предполагает дифференцируемости F, а требует вместо этого, чтобы разность F с некоторым накрывающим линейным оператором удовлетворяла условию Липшица с константой, меньшей модуля накрывания оператора.) Например, если $\bar{x}$ есть решение задачи о минимизации $f(x)$ при условии $F(x) = 0$, то отображение $x \mapsto \big(f(x), F(x)\big)$ не может быть накрывающим в окрестности $\bar{x}$. В гладком случае это наблюдение вместе с теоремой

Люстерника—Грейвса немедленно приводит к доказательству принципа Лагранжа. Впервые такой подход был применён к доказательству принципа максимума, видимо, в работе Дж. Варги [36]. Дж. Варга, впрочем, ссылается не на теорему Люстерника—Грейвса, а на результат Дж. А. Йорка [37], представляющий собой её ослабленный вариант применительно к интегральным операторам, ассоциированным с обыкновенными дифференциальными уравнениями. Аналогичная идея лежит в основе доказательства принципа Лагранжа для гладких аппроксимативно-выпуклых задач в недавней работе Е. Р. Аваков, Г. Г. Магарил-Ильяев и В. М. Тихомиров [1].

Поразительно, но в то время как теорема Люстерника уже в середине 1960-х, начиная с работ А. Я. Дубовицкого и А. А. Милютина, стала играть центральную роль в доказательствах необходимых условий, значение теоремы Грейвса о накрывании для теории необходимых условий не было осознано, видимо, до конца 1970-х годов. (Первое известное мне явное использование теоремы Грейвса имеет место в обзоре А. В. Дмитрука, А. А. Милютина и Н. П. Осмоловского [4], где рассматриваются уже «полугладкие» отображения с конечномерными негладкостями.) Возможно, это объясняется тем, что связь оценок метрической регулярности (подобных, например, полученным в [7] или двумя годами позже в более общей ситуации С. М. Робинсоном в [31]) с накрыванием была осознана лишь в 1980-е годы.

Разумеется, принцип «оптимальность убивает накрывание» никоим образом с гладкостью не связан и подходящий критерий ненакрывания может быть базой для получения принципа Лагранжа (см., например, [21]). Следует, однако, заметить, что необходимые условия в форме принципа Лагранжа могут быть получены лишь при достаточно ограничительных предположениях, по-видимому исключающих наличие функциональных включений типа $0 \in F(x)$.

Второй подход идейно связан с более слабым свойством метрической субрегулярности (введённым под другим названием в [6] и обстоятельно изученным в [15]), хотя и не нуждается ни в самом этом свойстве и ни в каких-либо гарантирующих его критериях. Непосредственным предшественником этого подхода была простая и естественная идея о возможности использования «точных штрафов», высказанная впервые, видимо, Ф. Кларком в [11]: если f имеет локальный минимум на замкнутом множестве Q в точке $\bar{x} \in Q$ и при этом удовлетворяет условию Липшица в этой точке, то $\bar{x}$ — точка безусловного локального минимума функции $g(x) = f(x) + K\,d(x, Q)$, при всяком K превышающем липшицеву константу f в $\bar{x}$. Несомненно, переход к безусловной минимизации может предоставить много удобств: см., например, доказательство принципа максимума для задач со свободным правым концом в [28]. Однако реализация этой идеи в чистом виде требует искусства и изобретательности, в частности для подсчёта расстояния, когда множество Q задаётся с помощью тех или иных функциональных соотношений (ср. [12]). Но и здесь связанные с регулярностью соображения позволяют существенно упростить ситуацию и ввести точные штрафы, явно использующие структуру функциональных ограничений. Например, в простейшей задаче минимизации f при условии $F(x) = 0$ мы можем перейти к безусловной минимизации функции $\varphi(x) = f(x) + K\|F(x)\|$, если f липшицева, F субрегулярно в $\bar{x}$, а критерий ненакрывания, применённый только к F, может рассматриваться как необходимое условие минимума в «вырожденном» случае. Дополнительным удобством в этом случае является возможность единообразного подхода к получению необходимых условий не только первого, но и более высоких порядков [19]. Более того, в рамках подобного подхода оказывается возможным отказаться даже от необходимости априорной проверки субрегулярности ограничений и использования критериев ненакрывания. Именно справедлива следующая теорема.

Теорема 1.1 (оптимизационная альтернатива). *Рассмотрим абстрактную задачу минимизации*

$$\text{минимизировать } \ f(x) \ \text{ при условии } \ x \in M \subset X.$$

Мы предполагаем, что (X, d) *— полное метрическое пространство,* $\bar{x} \in M$ *и* f *удовлетворяет условию Липшица в окрестности* $\bar{x}$*. Пусть* $\varphi(x)$ *— неотрицательная функция, равная нулю в* $\bar{x}$*. Предположим, наконец, что* $\bar{x}$ *— локальное решение задачи. Тогда имеет место следующая альтернатива:*

— *или функция* $\lambda f + \varphi$ *имеет безусловный минимум в* $\bar{x}$ *при некотором* $\lambda > 0$;
— *или найдётся последовательность* $(x_n) \to \bar{x}$*, такая что*

$$\varphi(x_n) < n^{-1} d(x_n, M)$$

и для всякого n *функция* $x \mapsto \varphi(x) + n^{-1} d(x, x_n)$ *имеет локальный минимум в* x_n.

Мы будем говорить о *регулярной ситуации* в первом случае и о *нерегулярной* во втором.

Доказательство. Действительно, либо существует такое $R > 0$, что $R\varphi(x) \geqslant d(x, M)$ для всех x из некоторой окрестности $\bar{x}$, либо имеется последовательность (z_n), сходящаяся к $\bar{x}$ и такая, что $n^2\varphi(z_n) < d(x_n, M)$. В первом случае, поскольку функция f липшицева, для всякого x, близкого к $\bar{x}$, и всякого $u \in M$, такого что, скажем, $d(x, u) \leqslant 2d(x, M)$, имеем

$$f(x) \geqslant f(u) - Ld(x, u) \geqslant f(\bar{x}) - 2LR\varphi(x),$$

где L — липшицева константа f.

Так как X полно и φ полунепрерывна снизу, мы можем применить вариационный принцип Экланда к $\varphi(x)$, принимая во внимание, что $\varphi(z_n) \leqslant \inf \varphi + n^{-2} d(z_n, M)$, и найти (x_n), такие что $d(x_n, z_n) \leqslant n^{-1} d(z_n, Q)$, $\varphi(x_n) \leqslant \varphi(z_n)$ и $\varphi(x) + n^{-1} d(x, x_n) > \varphi(x_n)$ для $x \neq x_n$. Наконец,

$$d(x_n, Q) \geqslant d(z_n, Q) - d(x_n, z_n) \geqslant (1 - n^{-1}) d(z_n, Q) \geqslant \\ \geqslant (1 - n^{-1}) n^2 \varphi(z_n) \geqslant n\varphi(x_n). \quad \square$$

Таким образом, исходная задача с ограничениями заменяется одной задачей или последовательностью задач безусловной минимизации. Свобода выбора функции φ (в дальнейшем мы будем называть её *пробной функцией*) даёт основание надеяться, что последние могут оказаться проще для анализа. Альтернатива, ещё до того как она была явным образом сформулирована в [24], лежала в основе доказательств принципа максимума для дифференциальных включений в [22] (см. также [34]). Возникавшие при этом задачи безусловной минимизации имели форму классической вариационной задачи Больцы, но с негладким интегрантом.

В разделе 5 мы приводим в качестве достаточно простого примера основанное на оптимизационной альтернативе полное доказательство принципа максимума Понтрягина для классической задачи оптимального быстродействия (с закреплёнными концами). Это доказательство, хотя и использует технику субдифференциального исчисления, кажется короче и технически проще большинства (если не всех!) доступных доказательств этого результата.

В заключительной части статьи мы рассматриваем задачу оптимального управления с дифференциальными включениями. Наиболее общий принцип максимума для таких задач был доказан Ф. Кларком в [13]. В разделах 6 и 7 мы предлагаем другое доказательство этого результата, использующее оптимизационную альтернативу. Наше доказательство следует общей схеме доказательства в [22], но использует другую пробную функцию, построенную с использованием остроумного технического приёма, заимствованного из [13]. Оно короче и, как нам кажется, проще в чисто идейном плане. Читатель без труда убедится, что это, по существу, то же доказательство, что и в разделе 5.

В обоих случаях, как и в [22, 34], задача минимизации пробной функции является вариационной задачей Больцы с негладким и, вообще говоря, невыпуклым (по производной) интегрантом. Следует отметить, что задача Больцы как результат применения точного штрафа в задаче оптимального управления со свободным правым концом возникала уже в [28]. Однако техника, используемая в [28], позволяет получить лишь необходимые условия, заметно более слабые, чем полученные позднее в [26, 34] (в дополнение к тому очевидному факту, что точного штрафа может не быть в задаче с закреплёнными концами).

Существенная доля технических проблем в общей структуре доказательства принципа максимума в цитированных [22] и [34] была связана именно с доказательством необходимых условий в негладкой задаче Больцы. В [26] они впервые были получены с помощью весьма сложных построений, использующих большой арсенал средств, включая среди прочего двойственность, сглаживание и релаксацию. В дальнейшем доказательство было упрощено в [34], но всё ещё оставалось технически весьма сложным. В разделе 4 мы даём заметно более простое и короткое доказательство необходимых условий для варианта задачи Больцы при предположениях, более простых по сравнению с [22] и [34], но всё же достаточных для доказательства принципа максимума на основе оптимизационной альтернативы. Несколько заключительных замечаний и комментариев собраны в коротком последнем разделе.

Начинаем мы с напоминания в следующем разделе некоторых фактов из вариационного анализа, необходимых для последующих рассуждений.

2. Необходимые сведения из локального вариационного анализа

Следующие обозначения являются более или менее стандартными (здесь f — функция на банаховом пространстве X и $Q \subset X$):

$$\operatorname{dom} f = \{x \colon |f(x)| < \infty\} \text{ —}$$

область определения f;

$$\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbb{R} \colon \alpha \geqslant f(x)\} \text{ —}$$

эпиграф f;

$$d(x, Q) = \inf\{\|x - u\| \colon u \in Q\} \text{ —}$$

расстояние от x до Q;

$$\delta_Q(x) = \begin{cases} 0, & x \in Q, \\ \infty, & x \notin Q, \end{cases} \text{ —}$$

индикаторная функция S;

$$s_Q(x^*) = \sup\{\langle x^*, x\rangle \colon x \in Q\} \text{ —}$$

опорная функция Q; $F \colon X \rightrightarrows Y$ — многозначное отображение из X в Y; $B(x, r)$ — замкнутый шар радиуса r с центром в x.

2.1. Субдифференциалы

Следующие определения формально имеют смысл в любом банаховом пространстве:

$$\partial_{\mathrm{F}} f(x) = \left\{x^* \colon \langle x^* h\rangle \leqslant \liminf_{h \to 0} \frac{f(x+h) - f(x)}{\|h\|}\right\} \text{ —}$$

субдифференциал Фреше функции f в точке x;

$$f^-(x; h) = \liminf_{t \to 0} t^{-1}\bigl(f(x + th) - f(x)\bigr) \text{ —}$$

нижняя производная Адамара в x;

$$\partial_{\mathrm{H}} f(x) = \{x^* \colon \langle x^*, h\rangle \leqslant f^-(x; h) \text{ для всех } h\} \text{ —}$$

субдифференциал Дини—Адамара функции f в x. Равенство $\partial_F f(x) = \partial_H f(x)$ справедливо для полунепрерывных функций на конечномерном пространстве. В общем случае $\partial_F f(x) \subset \partial_H f(x)$.

Специфическим для конечномерных пространств, а также для гильбертова пространства является *проксимальный субдифференциал*: $x^* \in \partial_{\mathrm{P}} f(x)$, если найдётся такое $\rho > 0$, что $f(x+h) - f(x) \geqslant \langle x^*, h\rangle - (1/2)\rho\|h\|^2$ для всех h из окрестности нуля. Разумеется, $\partial_{\mathrm{P}} f(x) \subset \partial_{\mathrm{F}} f(x)$.

Все функции в дальнейшем предполагаются по крайней мере полунепрерывными снизу, но могут принимать бесконечные значения обоих знаков.

Следующий результат играет фундаментальную роль в исчислении конечномерных субдифференциалов (см. [26] для проксимального субдифференциала и [20] для субдифференциала Фреше).

Теорема 2.1. *Пусть $f_1, \ldots, f_k$ — полунепрерывные снизу функции на $\mathbb{R}^n$, конечные в точке x. Пусть ∂ обозначает либо проксимальный субдифференциал, либо субдифференциал Фреше. Если $x^* \in \partial(f_1 + \ldots + f_k)(x)$, то для всякого $\varepsilon > 0$ найдутся $x_1, \ldots, x_k$, ε-близкие к x, и $x_i^* \in \partial f_i(x_i)$, такие что $\|x_1^* + \ldots + x_k^* - x^*\| < \varepsilon$.*

Предельный субдифференциал функции f в точке x — это

$$\partial f(x) = \limsup_{u \to x} \partial f(u) = \{x^* \colon \text{ найдутся } x_k \to_f x,\ x_k^* \in \partial_{\mathrm{P}} f(x_k),\ x_k^* \to x^*\}.$$

Здесь посредством $x_k \to_f x$ обозначена f-сходимость: $x_k \to x$ и $f(x_k) \to f(x)$. Если функция полунепрерывна снизу, то в определении можно использовать субдифференциал Фреше — результат не изменится.

Отметим следующие элементарные свойства предельного субдифференциала:

— если функция f K-липшицева в окрестности x, то $\partial f(x) \neq \varnothing$ и $\|x^*\| \leqslant K$ для всякого $x^* \in \partial f(x)$;
— если функция f строго дифференцируема в x, то $\partial f(x) = \{f'(x)\}$;
— если функция f выпукла, то
$$\partial f(x) = \{x^* \colon \langle x^*, h\rangle \leqslant f(x+h) - f(x) \text{ для всех } h\};$$
— если $f(x) = d(x, Q)$ и $x^* \in \partial f(x)$ и при этом $f(x) > 0$, то $\|x^*\| = 1$.

Из более глубоких результатов, относящихся к исчислению предельных субдифференциалов, мы упомянем здесь следующую теорему о субдифференциале композиции (см., например, [30, предложение 6.17 (d)]).

Теорема 2.2. *Пусть $A\colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор ранга m и φ — функция на $\mathbb{R}^m$. Тогда*
$$\partial(\varphi \circ A)(x) = A^* \partial f(Ax).$$

Бесконечномерную ситуацию мы рассмотрим *только для случая сепарабельных банаховых пространств*. (Впрочем, все аргументы с небольшими изменениями переносятся на случай слабо компактно порождённых пространств.)

Главное и существенное отличие от конечномерного случая состоит в том, что субдифференциалы Дини—Адамара и Фреше могут не совпадать даже у достаточно просто устроенных функций. Более того, если в пространстве не существует эквивалентной нормы, дифференцируемой по Фреше в ненулевых точках, то даже у вогнутой непрерывной функции субдифференциал Фреше может быть тождественно пуст. Наоборот, субдифференциал Дини—Адамара полунепрерывной снизу функции на сепарабельном пространстве непуст на плотном подмножестве её области определения. Иначе говоря, мы можем доверять субдифференциалу Дини—Адамара на любом сепарабельном пространстве, а субдифференциалу Фреше — только на пространстве с дифференцируемой по Фреше нормой. (Заметим, что понятие «субдифференциал, которому можно доверять на некотором классе банаховых пространств» допускает точную формализацию; см. [25, 30].)

Подобно теореме 2.1, следующий результат играет центральную роль в исчислении субдифференциалов в бесконечномерном пространстве. (Обстоятельные обсуждения этого результата вместе с доказательствами его эквивалентности ряду других утверждений можно найти в [10, 23, 27, 38].)

Теорема 2.3 (неопределённый минимизационный принцип). *Предположим, что*

— *либо X — гильбертово пространство и ∂ — проксимальный субдифференциал;*
— *либо в X существует эквивалентная норма, дифференцируемая по Фреше, и ∂ — субдифференциал Фреше;*
— *либо в X существует эквивалентная норма, дифференцируемая по Гато, и ∂ — субдифференциал Дини—Адамара.*

Предположим, что $f_1, \ldots, f_k$ полунепрерывны снизу на X, конечны в x и все, за исключением, возможно, одной из них, удовлетворяют условию Липшица в окрестности этой точки. Если $f = f_1 + \ldots + f_k$ имеет локальный минимум в x, то для всякого $\varepsilon > 0$ найдутся $x_1, \ldots, x_k$ и $x_1^*, \ldots, x_k^*$, такие что

$$\|x_i - x\| < \varepsilon, \quad |f_i(x_i) - f_i(x)| < \varepsilon, \quad x_i^* \in \partial f_i(x_i), \quad \|x_1^* + \ldots + x_k^*\| < \varepsilon.$$

Следующий результат является специфическим для проксимационного субдифференциала и в значительной мере определяет предпочтение, отдаваемое ему при работе с интегральными функционалами.

Предложение 2.4 [14, теорема 5.10]. *Рассмотрим на пространстве $L^2(0,1)$ интегральный функционал*

$$\Phi(x(\cdot)) = \int_0^1 \varphi\bigl(t, x(t)\bigr)\, dt,$$

где φ задана на $[0,T] \times \mathbb{R}^n$, измерима по t при всяком x и непрерывна по x почти при всяком t. Тогда

$$\partial_P \Phi\bigl(x(\cdot)\bigr) \subset \bigl\{p(\cdot) \in L^2 \colon p(t) \in \partial_P \varphi\bigl(t, x(t)\bigr)\ dt \text{ п. в.}\bigr\}.$$

Мы не будем далее углубляться в обсуждение бесконечномерных аспектов субдифференциального исчисления и отсылаем читателя к [25]. Теорема 2.3 и предложение 2.4 — это всё, что нам нужно в данной работе.

2.2. Регулярность

Пусть $F\colon X \rightrightarrows Y$ и $\bar y \subset F(\bar x)$. В теории регулярности рассматриваются следующие три основные свойства многозначных отображений.

— Отображение F *накрывает в окрестности* $(\bar x, \bar y)$ с константой r, если для всякого положительного $r' < r$ найдётся такое $\varepsilon > 0$, что для всяких $(x, y) \in \operatorname{Graph} F$, ε-близких к $(\bar x, \bar y)$, включение $B(y, tr') \cap B(\bar y, \varepsilon) \subset \subset F\big(B(x, t)\big)$ выполняется при $0 \leqslant t < \varepsilon$.

— Отображение F *метрически регулярно в окрестности* $(\bar x, \bar y)$ с константой K, если для всякого $K' > K$ существует такое $\varepsilon > 0$, что неравенство $d\big(x, F^{-1}(y)\big) \leqslant K' d\big(y, F(x)\big)$ выполняется при всех (x, y), ε-близких к $(\bar x, \bar y)$.

— Отображение F обладает *псевдолипшицевым свойством* (или *свойством Обэна*) в окрестности $(\bar x, \bar y)$ с константой K, если для всякого $K' > K$ существует такое $\varepsilon > 0$, что включение $F(x) \cap B(\bar y, \varepsilon) \subset F(x') + K'\|x - x'\| B$ выполняется, если x и x' ε-близки к $\bar x$.

Три свойства тесно связаны друг с другом. Именно, F накрывает в окрестности $(\bar x, \bar y)$ с константой r тогда и только тогда, когда F метрически регулярно в окрестности $(\bar x, \bar y)$ с константой r^{-1}, и метрическая регулярность в окрестности $(\bar x, \bar y)$ эквивалентна свойству Обэна обратного отображения F^{-1} в окрестности $(\bar y, \bar x)$ с той же константой.

В разделе 6 нам понадобится следующий простой факт.

Предложение 2.5. *Предположим, что многозначное отображение $F\colon X \rightrightarrows Y$ псевдолипшицево с константой K в окрестности точки $(\bar x, \bar y) \in \operatorname{Graph} F$. Тогда для всякой точки $(x, y) \in \operatorname{Graph} F$, близкой к $(\bar x, \bar y)$, и всякой пары $(x^*, y^*) \in \in N_G\big(\operatorname{Graph} F, (x, y)\big)$ выполняется соотношение $\|x^*\| \leqslant K\|y^*\|$.*

3. Релаксация

Релаксация — это один из важнейших инструментов анализа задач вариационного исчисления и оптимального управления. Мы обсудим в заключительном разделе возникновение и развитие этого понятия и, в частности, отличия нашего подхода в настоящей статье от использовавшихся ранее.

Определение 3.1. Пусть X и Y — банаховы пространства, $A\colon X \to Y$ — замкнутый линейный оператор и f — функция на $X \times Y$. Скажем, что f *релаксируема относительно* A в окрестности точки $(\bar x, \bar y)$ $(A\bar x = \bar y)$ на подпространстве $Y_0 \subset Y$, если для всякого конечного набора $\{e_1, \ldots, e_k\} \subset Y_0$ существует такое $\varepsilon > 0$, что для всякой пары (x, y) с $Ax = y$ и $\|x - \bar x\| < \varepsilon$ и произвольных $\alpha_1 \geqslant 0, \ldots, \alpha_k \geqslant 0$ с $\sum \alpha_i \leqslant \varepsilon$ существует последовательность (x_n, y_n), такая что $Ax_n = y_n$, x_n сходятся к некоторому u, y_n слабо сходятся к $y + \sum \alpha_i e_i$, а $f(x_n, y_n)$ сходятся к

$$g(u, \alpha_1, \ldots, \alpha_k, y) = f(u, y) + \sum \alpha_i \big(f(u, y + e_i) - f(u, y)\big).$$

Следующее простое предложение, которое можно рассматривать как некоторый специальный случай оптимизационной альтернативы, оказывается весьма полезным во многих ситуациях, в частности для доказательства приводимой ниже теоремы о релаксации.

Предложение 3.2 [19]. *Пусть X, Y — банаховы пространства, $Q \subset X$ замкнуто, $F\colon X \to Y$ — отображение и f — функция на X, конечная в точке $\bar{x}$. Рассмотрим задачу о минимуме f при условии $F(x) = 0$, $x \in Q$. Пусть $\bar{x}$ — локальный минимум в этой задаче. Предположим, что f удовлетворяет условию Липшица в окрестности $\bar{x}$, F определено и липшицево в окрестности $\bar{x}$ и ограничение F на Q субрегулярно на Q в точке $\bar{x}$, т. е. $d\big(x, F^{-1}(0)\big) \geqslant K\|F(x)\|$ для всех $x \in Q$ из окрестности $\bar{x}$. Тогда существует такое $N > 0$, что $\bar{x}$ доставляет безусловный локальный минимум функции $f + N\big(\|F(\cdot)\| + d(x, Q)\big)$.*

Доказательство. По условию найдётся такое K, что для всякого $u \in Q$, достаточно близкого к $\bar{x}$, справедливо неравенство $d\big(u, F|_Q^{-1}(0)\big) \leqslant K\|F(u)\|$. Иначе говоря, для всякого такого u можно подобрать $x \in Q$ так, чтобы $F(x) = 0$ и $\|u - x\| \leqslant K\|F(u)\|$. Если λ — липшицева константа f, то

$$f(u) \geqslant f(x) - \lambda\|u - x\| \geqslant f(\bar{x}) - \lambda K\|F(u)\|,$$

т. е. $g(u) = f(u) + \lambda K\|F(u)\| \geqslant f(\bar{x}) = g(\bar{x})$. Если теперь c — липшицева константа g, то $\inf\limits_x \big(g(x) + cd(x, Q)\big) \geqslant \inf\limits_{x \in Q} g(x)$. $\square$

Предложение 3.3 (теорема о релаксации). *Пусть X, Y, Z — банаховы пространства и $Y_0 \subset Y$ — замкнутое подпространство. Предположим, что*

- *$A\colon X \to Y$ — замкнутый линейный оператор с плотной областью определения, отображающий X на всё Y и имеющий компактный обратный $\Lambda\colon Y \to X$;*
- *$\Lambda_0\colon Y \to Z$ — компактный линейный оператор;*
- *f — функция на $X \times Y$, удовлетворяющая условию Липшица;*
- *f_0 — функция на Z, удовлетворяющая условию Липшица.*

Если функция f релаксируема относительно A в окрестности точки $(\bar{x}, \bar{y})$ на Y_0 и функция $\varphi(x, y) = f_0(\Lambda_0 y) + f(x, y)$ имеет в этой точке локальный минимум при условии $Ax = y$, то для всякого конечного набора $e_1, \ldots, e_k$ элементов Y_0 функция

$$\psi(x, \alpha_1, \ldots, \alpha_k, y) = f_0\bigg(\Lambda_0\Big(y + \sum \alpha_i e_i\Big)\bigg) + g(x, \alpha_1, \ldots, \alpha_k, y)$$

имеет в точке $(\bar{x}, 0, 0, \ldots, 0, \bar{y})$ локальный минимум на множестве

$$Q = \left\{(x, \alpha_1, \ldots, \alpha_k, y)\colon \alpha_i \geqslant 0,\ i = 1, \ldots, k,\ Ax = y + \sum_{i=1}^{k} \alpha_i e_i\right\}.$$

Более того, существует $N > 0$, такое что $(\bar{x}, 0, 0, \ldots, 0, \bar{y})$ доставляет безусловный локальный минимум функции

$$\begin{aligned}
&h(x,\alpha_1,\dots,\alpha_k,y) = \\
&\quad = \psi(x,\alpha_1,\dots,\alpha_k,y) + N\left(\left\|x - \Lambda\left(y + \sum_{i=1}^k \alpha_i e_i\right)\right\| + \sum_{i=1}^k \alpha_i^-\right) = \\
&\quad = f_0\left(\Lambda_0\left(y + \sum \alpha_i e_i\right)\right) + f(x,y) + \sum_{i=1}^k \alpha_i\big(f(x,y+e_i) - f(x,y)\big) + \\
&\quad + N\left(\left\|x - \Lambda\left(y + \sum_{i=1}^k \alpha_i e_i\right)\right\| + \sum_{i=1}^k \alpha_i^-\right).
\end{aligned}$$

(*Здесь, как обычно,* $\alpha^- = \max\{-\alpha, 0\}$.)

Доказательство. Фиксируем $e_1,\dots,e_k \in Y_0$. Поскольку по условию $\Lambda = A^{-1}$ — компактный оператор, равенство $Au = y + \sum_i \alpha_i e_i$ выполняется, если $x_n \to u$, y_n слабо сходятся к $y + \sum_i \alpha_i e_i$ и $Ax_n = y_n$.

Определим оператор T из $X \times \mathbb{R}^k_+ \times Y$ (где $\mathbb{R}^k_+$ — неотрицательный ортант в $\mathbb{R}^k$) в X соотношением

$$T(x,y,\alpha_1,\dots,\alpha_k) = x - \Lambda\left(y + \sum_{i=1}^k \alpha_i e_i\right).$$

Если мы наделим $X \times \mathbb{R}^k \times Y$ нормой

$$\|(x,\alpha_1,\dots,\alpha_k,y)\| = \|x\| + |\alpha_1| + \ldots + |\alpha_k| + \|y\|\},$$

то, очевидно, T накрывает с константой, не меньшей чем 1, в частности,

$$d\big((x,\alpha_1,\dots,\alpha_k,y), T^{-1}(0)\big) \leqslant \left\|x - \Lambda\left(y + \sum_{i=1}^k \alpha_i e_i\right)\right\|. \tag{3.1}$$

Пусть теперь $Ax = y$ и $\|x - \bar{x}\| < \varepsilon/2$, $\|y - \bar{y}\| < \varepsilon/2$. Пусть $\delta \in (0,\varepsilon)$ таково, что $\|x-u\| < \varepsilon/2$ и $\alpha_i\|e_i\| < \varepsilon/2$, если $u = \Lambda\left(y+\sum_{i=1}^k \alpha_i e_i\right)$ и $\sum \alpha_i < \delta$. По определению (и в силу замечания в начале доказательства) для всяких таких u и α_i существует последовательность (x_n, y_n), такая что y_n слабо сходятся к $y+\sum \alpha_i e_i$, $x_n \to u$ и $f(x_n,y_n) \to g(u,\alpha_1,\dots,\alpha_k,y)$. Тогда $f_0(\Lambda_0 y_n) \to f_0\Big(\Lambda_0\big(y + \sum \alpha_i e_i\big)\Big)$ (в силу компактности Λ_0). Но $\varphi(x_n,y_n) \geqslant \varphi(\bar{x},\bar{y})$, и следовательно,

$$\psi(x,\alpha_1,\dots,\alpha_k,y) \geqslant \varphi(\bar{x},\bar{y}) = \psi(\bar{x},0,\dots,0,\bar{y}).$$

Этим завершается доказательство первого утверждения. Второе следует из первого ввиду (3.1) и предложения 3.2. □

4. Задача Больцы

В этом разделе мы рассмотрим классическую задачу Больцы о минимуме функционала

$$J\big(x(\cdot)\big) = \ell\big(x(0), x(T)\big) + \int\limits_0^T L\big(t, x(t), \dot{x}(t)\big)\, dt \quad x(\cdot)\colon [0,T] \to \mathbb{R}^n, \tag{4.1}$$

(B_1) ℓ удовлетворяет условию Липшица в окрестности $\big(\bar{x}(0), \bar{x}(1)\big)$ в $\mathbb{R}^{2n}$;

(B_2) интегрант L измерим по t и почти при каждом t удовлетворяет условию Липшица с константой $k(t)$ по x и абсолютной (не зависящей от значений переменных) константой K по y, при этом $k(\cdot) \in L^1$:

$$|L(t,x,y) - L(t,x',y')| \leqslant k(t)\|x - x'\| + K\|y - y'|.$$

Наша цель — доказать следующую теорему.

Теорема 4.1. *Если $\bar{x}(t) \equiv 0$ доставляет локальный минимум функционалу J в $W^{1,1}$, то существует функция $p(\cdot) \in W^{1,1}$ со значениями в $\mathbb{R}^n$, такая что выполнены следующие три соотношения:*

— $\big(p(0), -p(1)\big) \in \partial\ell\big(\bar{x}(0), \bar{x}(1)\big)$ *(условие трансверсальности);*

— $\dot{p}(t) \in \operatorname{conv}\big\{q\colon \big(q, p(t)\big) \in \partial L_t\big(\bar{x}(t), \dot{\bar{x}}(t)\big)\big\}$ *почти всюду (эйлерово включение);*

— $L(x, \bar{x}(t), y) - L\big(t, \bar{x}(t), \dot{\bar{x}}(t)\big) - \langle p(t), y - \dot{\bar{x}}(t)\rangle \geqslant 0$ *для всех y почти всюду (условие Вейерштрасса).*

Мы начнём с проверки релаксируемости функционала J при условиях (B_1) и (B_2).

Лемма 4.2. *Пусть заданы $\alpha_i > 0$, $i = 1, \dots m$, в сумме равные единице. Тогда на отрезке $[0,1]$ существуют измеримые функции $\alpha_{1k}(t), \dots, \alpha_{mk}(t)$, $k = 1, 2, \dots$, такие что*

— $\sum\limits_i \alpha_{ik}(t) = 1$ *почти всюду для всех k и каждая функция $\alpha_{ik}(t)$ принимает лишь значения 0 и 1;*

— *для всякого $x(\cdot) \in L^1$ и всякого i $\lim\limits_{k\to\infty} \int\limits_0^1 \alpha_{ik}(t)x(t)\,dt = \alpha_i \int\limits_0^1 x(t)\,dt$.*

В частности, каждая последовательность $\big(\alpha_{ik}(\cdot)\big)$, $i = 1, \dots, m$, слабо сходится в каждом L^p, $p \in [1, \inf y)$, к постоянной функции, тождественно равной α_i.

Доказательство. Положим $\beta_0 = 0$, $\beta_i = \alpha_1 + \ldots + \alpha_i$ для $i = 1, \dots, m$ и определим α_{ik} следующим образом:

$$\alpha_{ik} = \begin{cases} 1, & t \in k^{-1}(l + \beta_{i-1}, l + \beta_i),\ l = 0, 1 \dots, k-1, \\ 0 & \text{иначе.} \end{cases}$$

Проверка декларированных соотношений при этом не составляет труда. $\square$

Определим оператор $A\colon L^2 \to \mathbb{R}^n \times L^2$ и функционал f на $L^2 \times (\mathbb{R}^n \times L^2)$:

$$A\bigl(x(\cdot)\bigr) = \bigl(x(0), \dot{x}(\cdot)\bigr), \quad f\Bigl(x(\cdot), \bigl(a, y(\cdot)\bigr)\Bigr) = \int\limits_0^1 L\bigl(t, x(t), y(t)\bigr)\, dt.$$

Очевидно, A — замкнутый оператор (его область определения совпадает с $W^{1,2}$) на всё пространство $\mathbb{R}^n \times L^2$, его обратный

$$\Lambda\bigl(a, y(\cdot)\bigr) = a + \int\limits_0^t y(\tau)\, d\tau \;—$$

компактный оператор. Пусть, наконец, L_t обозначает функцию $(x, y) \mapsto L(t, x, y)$.

Предложение 4.3. *При выполнении условий* (B_1) *и* (B_2) *функция* f *релаксируема относительно* A *на* L^2.

Доказательство. Пусть заданы некоторые $\alpha_i > 0$, $i = 1, \dots, m$, $\sum \alpha_i < 1$. Положим $\alpha_0 = 1 - \sum\limits_1^m \alpha_i$. Пусть функции $\alpha_{ik}(t)$, $i = 0, \dots, m$, $k = 1, 2, \dots$, построены для $\alpha_0, \dots, \alpha_m$ согласно лемме 4.2. Для заданных $y(\cdot)$ и $e_i(\cdot) \in L^2$, $i = 1, \dots, m$, положим

$$y_k(t) = \sum_{i=1}^m \alpha_{ik}(t)\bigl(y(t) + e_i(t)\bigr).$$

По лемме 4.2 $y_k(\cdot)$ слабо сходятся в L^2 к $y(\cdot) + \sum \alpha_i e_i(\cdot)$. С другой стороны (так как $\alpha_{ik}(t)$ принимают лишь значения 0 и 1),

$$L\bigl(t, x, y_k(t)\bigr) = \sum_{i=1}^m \alpha_{ik}(t) L\bigl(t, x, y(t) + e_i(t)\bigr) =$$

$$= L\bigl(t, x, y(t)\bigr) + \sum_{i=1}^m \alpha_{ik}(t)\Bigl(L\bigl(t, x, y(t) + e_i(t)\bigr) - L\bigl(t, x, y(t)\bigr)\Bigr).$$

Таким образом, если

$$x_k(t) = a + \int\limits_0^t y_k(s)\, ds,$$

то в силу слабой сходимости $y_k(\cdot)$ $x_k(\cdot)$ равномерно сходится к

$$x(t) = a + \int\limits_0^t \Bigl(y(t) + \sum_{i=1}^m \alpha_i e_i(s)\Bigr) ds.$$

Следовательно, по лемме 4.2 $f\Bigl(x_k(\cdot), \bigl(a, y_k(\cdot)\bigr)\Bigr)$ сходится к

$$f\bigl(x(\cdot), y(\cdot)\bigr) + \sum \alpha_i \biggl(f\Bigl(x(\cdot), \Bigl(a, y(\cdot) + \sum \alpha_i e_i(\cdot)\Bigr)\Bigr) - f\bigl(x(\cdot), y(\cdot)\bigr)\biggr). \qquad \square$$

Доказательство теоремы 4.1. Без ограничения общности можно считать, что $\bar{x}(t) \equiv 0$. Поскольку вложение $W^{1,2} \to W^{1,1}$ непрерывно, J определён на $W^{1,2}$ и $\bar{x}$ доставляет локальный минимум J на $W^{1,2}$. Мы докажем теорему 4.1 с помощью теоремы о релаксации, полагая $X = L^2$, $Y = \mathbb{R}^n \times L^2$, $Y_0 = L^2$, $Z = L^2 \times \mathbb{R}^k$ $f_0 = \ell$ и

$$\Lambda_0\big(a, y(\cdot)\big) = \left(a, a + \int\limits_0^1 y(t)\, dt\right),$$

так что в обозначениях теоремы о релаксации

$$\varphi\Big(x(\cdot), \big(a, y(\cdot)\big)\Big) = \ell\left(a, a + \int\limits_0^1 y(t)\, dt\right) + \int\limits_0^1 L\big(t, x(t), y(t)\big)\, dt.$$

По теореме о релаксации $\big(\bar{x}(\cdot), 0, \ldots, 0, \dot{\bar{x}}(\cdot)\big)$ — локальный минимум, соответствующей функции h. Мы можем представить h в виде суммы пяти функций $J_1, \ldots, J_5$ переменных $x(\cdot) \in L^2$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, $a \in \mathbb{R}^n$, $y(\cdot) \in L^2$ (мы в дальнейшем для краткости обозначаем кортежи $\big(x(\cdot), \alpha_1, \ldots, \alpha_k, a, y(\cdot)\big)$ через z):

$$\begin{aligned}
J_1(z) &= \ell\left(a, a + \int\limits_0^T \left(y(t) + \sum_{i=1}^k \alpha_i e_i(t)\right) dt\right),\\
J_2(z) &= \int\limits_0^T L\big(t, x(t), y(t)\big)\, dt,\\
J_3(z) &= \sum_{i=1}^k \alpha_i \Big(L\big(t, x(t), y(t) + e_i(t)\big) - L\big(t, x(t), y(t)\big)\Big),\\
J_4(z) &= N \int\limits_0^T \left\| x(t) - a - \int\limits_0^t \left(y(\tau) + \sum_{i=1}^k \alpha_i e_i(\tau)\right) d\tau \right\| dt,\\
J_5(z) &= N \sum_{i=1}^k \alpha_i^-.
\end{aligned}$$

Каждый из пяти функционалов удовлетворяет условию Липшица на $Z = L^1 \times \mathbb{R}^k \times \mathbb{R}^n \times L^2$. Пространство Z гильбертово, следовательно,

$$0 \in \partial_{\mathrm{P}}\left(\sum_{j=1}^k J_j\right)(\bar{z}), \tag{4.2}$$

и для анализа этого соотношения мы можем использовать неопределённый минимизационный принцип. Дальнейшие рассуждения весьма схожи с аргументацией третьего этапа доказательства теоремы 5 в [26].

Мы будем обозначать через z^* кортежи $\big(w(\cdot), \xi_1, \dots, \xi_k, b, v(\cdot)\big) \in Z$ с индексами того функционала J_j, к чьему проксимационному субдифференциалу относится данный z^*. Имеем

$$\begin{aligned}
&z_1^* \in \partial_{\mathrm{P}} J_1(z) \iff \\
&\qquad \iff w_1(\cdot) = 0, \quad \xi_{1j} = \Bigg(\int\limits_0^T e_j(t)\,dt\Bigg) v_1, \quad j = 1, \dots, k, \quad v_1(\cdot) \equiv v_1 = \mathrm{const}, \\
&\qquad\qquad (b_1 - v_1, v_1) \in \partial_{\mathrm{P}} \ell\Bigg(a, a + \int\limits_0^T \Bigg(y(t) + \sum_{i=1}^k \alpha_i e_i(t)\Bigg) dt\Bigg);
\end{aligned}$$

$$\begin{aligned}
&z_2^* \in \partial_{\mathrm{P}} J_2(z) \iff \\
&\qquad \iff \xi_{2j} = 0, \quad b_2 = 0, \quad \big(w_2(t), v_2(t)\big) \in \partial_{\mathrm{P}} L_t\big(x(t), y(t)\big) \text{ почти всюду}
\end{aligned}$$

(в последнем соотношении, которое следует из предложения 2.4, $L_t(\cdot) = L(t, \cdot)$);

$$\begin{aligned}
&z_3^* \in \partial_{\mathrm{P}} J_3(z) \implies \\
&\quad \implies \|w_3(\cdot)\|, \|v_3(\cdot)\| \leqslant c\Bigg(\sum_{i=0}^k \alpha_i\Bigg), \quad b_3 = 0, \\
&\qquad \xi_{3i} = \int\limits_0^T \Big(L\big(t, x(t), y(t) + e_i(t)\big) - L\big(t, x(t), y(t)\big)\Big)\, dt, \\
&z_4^* \in \partial_{\mathrm{P}} J_4(z) \implies \\
&\quad \implies w_4(t) = q(t), \quad \xi_{4i} = -\int\limits_0^T q(t) \int\limits_0^t e_i(\tau)\,d\tau\,dt, \quad b_4 = \Bigg(\sum_{i=1}^k \alpha_i - 1\Bigg) \int\limits_0^T q(t)\,dt, \\
&\qquad v_4(t) = -\int\limits_t^T q(\tau) d\tau
\end{aligned}$$

(здесь c — абсолютная константа, зависящая только от липшицевой константы интегранта L_t, $\|q(t)\| \leqslant N$ почти при всяком t и $\|q(t)\| = N$, если выражение под знаком интеграла в J_4 отлично от нуля при данном t); наконец,

$$z_5^* \in \partial_{\mathrm{P}} J_5(z) \implies \xi_{5i} \in [-N, 0], \quad i = 1, \dots, k, \quad w_5(\cdot) = v_5(\cdot) = 0, \quad b_5 = 0.$$

Согласно неопределённому минимизационному принципу из (4.2) следует, что для всякого $\varepsilon > 0$ можно найти z_j, произвольно близкие к $\bar z$, и $z_j^* \in \partial_{\mathrm{P}} J_j(z_j)$, такие что

$$\Bigg\|\sum_{j=1}^5 w_j(\cdot)\Bigg\| < \varepsilon, \quad \Bigg|\sum_{j=1}^5 \xi_{ji}\Bigg| < \varepsilon, \quad i = 1, \dots, k, \quad \Bigg\|\sum_{j=1}^5 b_j\Bigg\| < \varepsilon, \quad \Bigg\|\sum_{j=1}^5 v_j(\cdot)\Bigg\| < \varepsilon. \tag{4.3}$$

(Разумеется, имеются в виду нормы $w(\cdot)$ и $v(\cdot)$ в L^2.)

Таким образом, существуют последовательности (z_{js}), $j = 1, \dots, 5$, $s = 1, 2, \dots$, сходящиеся к $\bar z$, и $z_{js}^* \in \partial_{\mathrm{P}} J_j(z_{js})$, такие что нормы сумм соответствующих компонент z_j^* стремятся к нулю. Принимая во внимание, что

$(\alpha_{ij})_s \to 0$, и полагая

$$p_s(t) = -v_{1s} + \int\limits_t^1 q_s(\tau)\,d\tau,$$

получаем, в частности, для $x(\cdot)$, $y(\cdot)$ и a-компонент, что

$$w_{2s}(\cdot) - \dot p_s(\cdot) \to 0, \quad v_{2s}(\cdot) - p_s(\cdot) \to 0, \quad b_{1s} + \int\limits_0^1 q_s(t)\,dt \to 0, \tag{4.4}$$

и для α_i-компонент, что

$$\Bigl(\int\limits_0^1 e_i(t)\,dt\Bigr)v_{1s} + \int\limits_0^1 \Bigl(L\bigl(t, x_{3s}(t), y_{3s}(t) + e_i(t)\bigr) - L\bigl(t, x_{3s}(t), y_{3s}(t)\bigr)\Bigr)\,dt - \\ - \int\limits_0^1 \Bigl(\int\limits_t^1 q_s(\tau)\,d\tau\Bigr)e_i(t)\,dt + \xi_{5s} \to 0, \quad i = 1, \dots, k. \tag{4.5}$$

Поскольку v_{1s} не превосходят липшицевой константы ℓ и $q_s(\cdot)$ являются равномерно ограниченными, последовательность $\bigl(p_s(\cdot)\bigr)$ относительно компактна в пространстве непрерывных функций, так что, выбирая подходящую подпоследовательность, мы можем считать, что $p_s(\cdot)$ равномерно сходятся к некоторой $p(\cdot)$, а $\dot p_s(\cdot)$ слабо сходятся к $\dot p(\cdot)$. По теореме Мазура некоторая последовательность выпуклых комбинаций $\dot p_s(\cdot)$ (с наименьшим индексом, тоже стремящимся к бесконечности) сильно сходится к $\dot p(\cdot)$. Согласно (4.4) отсюда следует, что $v_{2s}(t) \to p(t)$ почти везде, а некоторая последовательность выпуклых комбинаций $w_{2s}(\cdot)$ сходится к $\dot p(\cdot)$ почти при каждом t.

Для всякого такого t и всякого $\varepsilon > 0$ рассмотрим множество $\hat Q_\varepsilon$ состоящее из всевозможных векторов q, таких что $(q, p) \in \partial_t(x, y)$ для некоторых (x, y), удалённых не более чем на ε от $\bigl(\bar x(t), \dot{\bar x}(t)\bigr)$, и p, такого что $\|p - p(t)\| < \varepsilon$. Пусть Q_ε — замыкание $\hat Q_\varepsilon$. Тогда, как следует из объяснений в предыдущем абзаце, $\dot p(\cdot) \subset \operatorname{conv} Q_\varepsilon$.

Согласно (B_2) Q_ε лежит в шаре радиуса $k(t)$ и, значит, компактно. Очевидно, что эти множества убывают с уменьшением ε и

$$\bigcap_{\varepsilon>0} (\operatorname{conv} Q_\varepsilon) = \operatorname{conv}(\cap Q_\varepsilon) = \operatorname{conv}\bigl\{q \colon \bigl(q, p(t)\bigr) \in \partial L_t,\ \bigl(\bar x(t), \dot{\bar x}(t)\bigr)\bigr\}.$$

Последнее справедливо, поскольку отображение $(x, y) \mapsto \partial L_t(x, y)$ компактнозначно и полунепрерывно сверху. Этим завершается доказательство эйлерова включения.

Далее, $v_{1s} \to -p(1)$, $b_{4s} \to p(1) - p(0)$ и $b_{1s} + b_{4s} \to 0$ согласно (4.3). Таким образом, $b_{1s} - v_{1s} \to p(0)$ и, переходя к пределу при $s \to \infty$, мы получаем условие трансверсальности $\bigl(p(0), -p(1)\bigr) \in \partial\ell\bigl(\bar x(0), \bar x(T)\bigr)$.

Наконец, поскольку $-N \leqslant \xi_{5i} \leqslant 0$ и $y_s(\cdot) \to \dot{\bar x}(\cdot)$, получаем из второго соотношения в (4.3) с учётом полученных выше представлений для ξ_{ji}, что

$$\int_0^1 \Big(L\big(t,\bar{x}(t),\dot{\bar{x}}(t)+e_i(t)\big) - L\big(t,\bar{x}(t),\dot{\bar{x}}(t)\big) - \langle p(t),e_i(t)\rangle\Big)\,dt \geqslant 0, \qquad i=1,\dots,k. \tag{4.6}$$

Таким образом, для каждого конечного набора $\{e_1(\cdot),\dots,e_k(\cdot)\}\subset L^2$ множество функций $p(\cdot)$, удовлетворяющих условию трансверсальности, эйлерову включению и неравенствам (4.6) для всех $i=1,\dots,k$, не пусто. Очевидно, всякое такое множество слабо компактно в $W^{1,2}$, и расширение набора может привести лишь к сужению этого множества. Поэтому пересечение всех таких множеств не пусто, и значит, существует $p(\cdot)$, такая что

$$\int_0^T \Big(L\big(t,\bar{x}(t),\dot{\bar{x}}(t)+e(t)\big) - L\big(t,\bar{x}(t),\dot{\bar{x}}(t)\big) - \langle p(t),e(t)\rangle\Big)\,dt \geqslant 0$$

для всякого $e(\cdot)\in L^2$. Но это значит, что

$$L(t,\bar{x}(t),y) - L\big(t,\bar{x}(t),\dot{\bar{x}}(t)\big) - \langle p(t),y-\dot{\bar{x}}(t)\rangle)\,dt \geqslant 0 \quad \text{для всех} \quad y\in\mathbb{R}^n$$

почти при всех t. Теорема полностью доказана. □

5. Пример: классическая задача о быстродействии

Имеется в виду задача

$$T\to\min\colon \dot{x}=f(x,u),\ u\in U,\ x(0)=x_0,\ x(T)=x_1.$$

Здесь U — замкнутое ограниченное подмножество $\mathbb{R}^k$ и $f\colon\mathbb{R}^n\times U\to\mathbb{R}^n$ вместе со своей производной по x непрерывна по совокупности переменных. Предположение об ограниченности множества U допустимых управлений, равно как и то, что мы рассматриваем автономную систему с f, не зависящим от t, не являются слишком ограничительными условиями в контексте классической задачи, и мы принимаем их, чтобы не вдаваться в излишние технические детали.

Для дальнейшего нам удобнее переформулировать задачу как задачу с фиксированным временем. Стандартный способ состоит во введении дополнительной фазовой фазовой координаты:

$$\text{(P)}\quad \tau(1)\to\min\colon \dot{x}=vf(x,u),\ \dot{\tau}=v,\ u\in U,\ v\in[0,N], \qquad x(0)=x_0,\ x(1)=x_1,\ \tau(0)=0,$$

где N достаточно велико (скажем, $N>2\bar{T}$). Если $\big(\bar{x}(t),\bar{u}(t)\big)$, $t\in[0,\bar{T}]$, — оптимальный процесс в исходной задаче, то $\hat{x}(t)=\bar{x}(\bar{T}t)$, $\hat{\tau}(t)=\bar{T}t$, $\hat{u}(t)=\bar{u}(\bar{T}t)$, $\hat{v}(t)\equiv\bar{T}$ дают решение последней задачи.

Положим $u_0(\cdot)=\hat{u}(\cdot)$, и пусть измеримые управления $u_1(\cdot),u_2(\cdot),\dots$ (со значениями в U) таковы, что почти при всяком $t\in[0,1]$

— $f\big(\hat{x}(t),u_i(t)\big)\neq\dot{\hat{x}}(t)$;

— $\{u_i(t)\colon i=1,2,\dots\}$ плотно в $U\setminus\{u\colon \bar{T}f(\hat{x}(t),u)=\dot{\hat{x}}(t)\}$.

Построение такой системы функций $u_i(t)$ не представляет труда. Далее, для всякого $t \in [0,1]$ определим множества

$$U_m(t) = \{u_i(t),\ i = 0, 1, \dots, m\},$$
$$V_m = \left\{ v_j = \left(\frac{j}{m}\right) \bar{T},\ j = 1, \dots, Nm \right\},$$
$$F_m(t, x) = \{v f(x, u) \colon v \in V_m,\ u \in U_m(t)\}$$

и для каждого m рассмотрим вспомогательную задачу (P_m), получающуюся из задачи (P) заменой условия $u \in U$ условием $u \in U_m(t)$ почти при всяком $t \in [0,1]$, а условия $v \geqslant 0$ условием $v \in V_m$. Очевидно, $\big(\hat{x}(\cdot),\ \hat{\tau}(\cdot),\ \hat{u}(\cdot),\ \hat{v}(\cdot)\big)$ — решение и в этой вспомогательной задаче. Мы применим оптимизационную альтернативу к вспомогательным задачам (P_m), а затем перейдём к пределу по m.

Итак, рассмотрим задачу (P_m). В качестве пробной функции берём

$$\varphi\big(x(\cdot), \tau(\cdot)\big) = \int\limits_0^1 L\big(t, x(t), \tau(t), \dot{x}(t), \dot{\tau}(t)\big)\, dt + \|x(0) - x_0\| + |\tau(0)| + \|x(1) - x_1\|,$$

где

$$L(t, x, \tau, y, \xi) = \min_{\substack{0 \leqslant i \leqslant m,\\ 0 \leqslant j \leqslant Nm}} \big(\big\|y - v_j f\big(x, u_i(t)\big)\big\| + \|\xi - v_j\| \big).$$

Регулярный случай. Предположим, что при некотором положительном λ пара $\big(\hat{x}(\cdot), \hat{\tau}(\cdot)\big)$ доставляет локальный минимум функционалу

$$\Psi\big(x(\cdot), \tau(\cdot)\big) = \lambda \tau(1) + \|x(0) - x_0\| + |\tau(0)| + \|x(1) - x_1\| + \\ + \int\limits_0^1 L\big(t, x(t), \hat{\tau}(t), \dot{x}(t), \dot{\tau}(t)\big)\, dt.$$

Это функционал типа Больцы, для которого условия (B_1) и (B_2) очевидно выполняются, и значит, применима теорема 4.1 и существуют суммируемые $p(\cdot)$ со значениями в $\mathbb{R}^n$ и $\eta(\cdot)$, гарантирующие выполнение трёх декларируемых теоремой условий.

Условие трансверсальности сводится к соотношениям

$$\eta(1) = -\lambda, \quad \|p(0)\| \leqslant 1, \quad \|p(1)\| \leqslant 1, \quad |\eta(0)| \leqslant 1. \tag{5.1}$$

Условие Вейерштрасса, применённое к $y = \bar{T} f\big(\hat{x}(t), u_i(t)\big)$, $i = 1, \dots, m$, и $\xi = \bar{T}$, почти при всяком t даёт

$$0 \geqslant \langle p(t), f\big(\hat{x}(t), u_i(t)\big) - f\big(\hat{x}(t), \hat{u}(t)\big) \rangle, \ \forall\, i = 1, \dots, m. \tag{5.2}$$

Переходя к анализу эйлерова включения, отмечаем, что (по выбору $u_i(t)$) $L_t\big(\hat{x}(t), \hat{\tau}(t), \dot{\hat{x}}(t), \dot{\hat{\tau}}(t)\big) = 0$ почти при всех $t \in [0,1]$ и, более того, при (x, y, τ, ξ), близких к $\big(\hat{x}(t), \hat{\tau}(t), \dot{\hat{x}}(t), \dot{\hat{\tau}}(t)\big)$, выполняется равенство

$$L_t(x, \tau, y, \xi) = \|y - \bar{T} f(x, \hat{u}(t)\| + |\xi - \bar{T}|. \tag{5.3}$$

Другими словами, каждое из двух слагаемых в L_t (как функция (x,y,ξ)) есть композиция выпуклой (норма и модуль) и гладкой (во втором слагаемом даже линейной) функций. Поэтому

$$(q,p,\mu,\eta)\in\partial L_t\big(\hat x(t),\hat\tau(t),\dot{\hat x}(t),\dot{\hat\tau}(t)\big)$$

тогда и только тогда, когда для некоторых $a\in\mathbb{R}^n$, $\|a\|\leqslant 1$ и $\alpha\in[-1,1]$

$$q=-\bar T f_x\big(\hat x(t),\hat u(t)\big)a,\quad p=a,\quad \mu=0,\quad \eta=\alpha.$$

Таким образом, эйлерово включение в данном случае

$$\dot p=-\bar T f'_x\big(\hat x(t),\hat u(t)\big)p,\quad \eta(t)\equiv\mathrm{const}=-\lambda. \tag{5.4}$$

Наконец, применяя условие Вейерштрасса к $y=v_j f\big(\hat x(t),\hat u(t)\big)$, $\xi=v_j$, $j=1,\dots,Nm$, и принимая во внимание, что $\dot{\hat x}(t)=\bar T f\big(\hat x(t),\hat u(t)\big)$, получаем

$$\bar T\big(\big\langle p(t),f\big(\hat x(t),\hat u(t)\big)\big\rangle-\alpha\big)\geqslant v_j\big(\big\langle p(t),f\big(\hat x(t),\hat u(t)\big)\big\rangle-\alpha\big)\quad \text{для всех}\ \ j=1,\dots,Nm, \tag{5.5}$$

что может случиться, лишь если

$$\big\langle p(t),f\big(\hat x(t),\hat u(t)\big)\big\rangle=\alpha. \tag{5.6}$$

Нерегулярный случай. В этом случае имеется последовательность пар $\big(x_s(\cdot),\tau_s(\cdot)\big)$, таких что $x_s(t)\to\hat x(t)$, $\tau_s(t)\to\hat\tau(t)$ равномерно, $\dot x_s(t)\to\dot{\hat x}(t)$, $\dot\tau_s(t)\to\dot{\hat\tau}(t)$ почти везде и при всяком s $\big(x_s(\cdot),\tau_s(\cdot)\big)$ доставляет минимум функционалу

$$\Psi_s\big(x(\cdot),\tau(\cdot)\big)=\|x(0)-x_0\|+|\tau(0)|+\|x(1)-x_1\|+\\+\delta_s(\|x(0)-x_s(0)\|+|\tau(0)-\tau_s(0)|)+\int_0^1 L_s\big(t,x(t),\tau(t),\dot x(t),\dot\tau(t)\big)\,dt$$

где

$$L_s(t,x,\tau,y,\xi)=L(t,x,\tau,y,\xi)+\delta_s(\|y-\dot x_s(t)\|+|\xi-\dot\tau_s(t)|)$$

и $\delta_s\to 0$. При этом либо на множестве положительной меры $\dot x_s(t)\notin F_m\big(t,x_s(t)\big)$ и/или $\dot\tau_s(t)\notin V_m$, либо $\|x_s(0)-x_0\|+|\tau_s(0)|+\|x_s(1)-x_1\|\neq 0$, либо то и другое вместе.

В этом случае функционал удовлетворяет всем условиям теоремы 4.1, и следовательно, существуют $p_s(\cdot)$ и $\xi(\cdot)$, удовлетворяющие трём условиям декларируемым этой теоремой для функционала Ψ_s. Иначе говоря, существуют $p_s(\cdot)$ и $\eta_s(\cdot)$, такие что

$$\|p_s(0)\| \leqslant 1+\delta_s, \quad \|p_s(1)\| \leqslant 1, \quad |\eta_s(0)| \leqslant 1+\delta_s, \quad \eta_s(1)=0,$$
$$\big(\dot p_s, \dot\eta_s(t)\big) \in \operatorname{conv}\big\{(q,\omega)\colon \big(q,\omega,p_s(t),\eta_s(t)\big) \in \partial L_t\big(x_s(t),\tau_s(t),\dot x_s(t),\eta_s(t)\big) + {} + \{0\}\times\{0\}\times(\delta_s B)\times[-\delta_s,\delta_s]\big\},$$
$$\delta_s\big(\big\|\dot x_s(t) - f\big(x_s(t),u_i(t)\big)\big\| + |\dot\tau_s(t) - v_j(t)|\big) + \langle p_s(t),\dot x_s(t)\rangle + \eta_s(t)\dot\tau_s(t) \geqslant {} \geqslant v_j(t)\big(\big\langle p_s(t), f\big(x_s(t),u_i(t)\big)\big\rangle + \eta_s(t)\big), \quad i=0,1,\dots,m,\ j=1,\dots,Nm. \tag{5.7}$$

Последние два соотношения, конечно, выполнены почти при всех t. Поскольку L не зависит от τ, $\dot\eta_s(t)\equiv 0$, и значит, $\eta_s(t)\equiv 0$.

Посмотрим, что происходит с этими соотношениями при $s\to\infty$. Если (при некотором t) $x_s(t)$, $\tau_s(t)$ и их производные настолько близки к $\hat x(t)$, $\hat\tau(t)$, $\dot{\hat x}(t)$, $\dot{\hat\tau}(t)$, что минимум в определении L при близких значениях x, τ, y, ξ реализуется при $u=\hat u(t)$, $v=\bar T$, то те же аргументы, что и в регулярном случае, показывают, что второе соотношение в (5.7) редуцируется к

$$\dot p_s(t) = -\bar T f'_x\big(x_s(t),\hat u(t)\big)a, \quad p(t)=a+b,$$

с некоторыми a и b, такими что $\|a\|\leqslant 1$, $\|b\|\leqslant\delta_s$, или иначе

$$\big\|\dot p_s(t) - \bar T f'_x\big(x_s(t),\hat u(t)\big)p_s(t)\big\| \leqslant K\delta_s. \tag{5.8}$$

Заметим, что в силу условий, наложенных на f, интегрант L как функция от x удовлетворяет условию Липшица с константой $k(t)$. Поэтому $\|\dot p_s(t)\|\leqslant k(t)$ почти везде и, следовательно, можно считать (ввиду ограниченности $p_s(0)$), что $p_s(\cdot)$ сходятся равномерно к некоторой функции $p(\cdot)$, а производные $\dot p_s(\cdot)$ сходятся слабо к $\dot p(\cdot)$. Используя опять теорему Мазура о выпуклых комбинациях слабо сходящихся последовательностей, получаем из (5.8), что $p(\cdot)$ удовлетворяет дифференциальному уравнению (5.4).

Наконец, переходя к пределу при $s\to\infty$ в последнем соотношении в (5.7), получаем (снова принимая во внимание, что $\dot{\hat x}(t) = \bar T f\big(\hat x(t),\hat u(t)\big)$), что

$$\bar T\big\langle p(t), f\big(\hat x(t),\hat u(t)\big)\big\rangle \geqslant v_j\big\langle p(t), f\big(\hat x(t),u_j(t)\big)\big\rangle \quad \text{для всех}\ \ i=1,\dots,m,\ j=1,\dots,Nm.$$

Полагая в последнем неравенстве $v_j=\bar T$, приходим к (5.2), а при $u_j(t)=\hat u(t)$ видим, что неравенство может выполняться при всех v_j, только если $\big\langle p(t), f\big(\hat x(t),\hat u(t)\big)\big\rangle = 0$ почти при всех t.

Нам осталось проверить, что $p(t)\neq 0$. Мы уже отмечали, что либо на множестве положительной меры $\dot x_s(t)\notin F_m(t,x)$ и/или $\dot\tau_s(t)\notin V_m$, либо $\|x_s(0)-x_0\| + |\tau_s(0)| + \|x_s(1)-x_1\| \neq 0$. Заметим, что $\tau_s(0)=0$ и $\dot\tau_s(t)\in V_m$ почти всюду, поскольку $\eta(t)$ тождественно равна нулю. Если $x_s(0)\neq x_0$, то $\|p_s(0)-1\|\leqslant\delta_s$. Аналогичная ситуация возникает, если $x(1)\neq x_1$. Наконец, если $\dot x_s(t)\notin F_m(t,x)$ на множестве положительной меры, то $\|p_s(t)\|=1$ почти при всех t на этом множестве. Таким образом, при каждом s равенство $\|p_s(t)\|=1$ выполняется хотя бы при одном t, и поскольку $p_s(\cdot)$ сходятся равномерно, предельная функция $p(\cdot)$ отлична от нуля хотя бы в одной точке и, значит, на всём интервале, так как она удовлетворяет линейному однородному дифференциальному уравнению (5.4).

Завершение доказательства. Положим $H(x,u,p) = \langle p, f(x,u)\rangle$. Возвращаясь к исходной задаче и суммируя полученные результаты, мы можем утверждать, что существуют $\lambda \geqslant 0$ и абсолютно непрерывная функция $p(\cdot)$, обладающие следующими свойствами:

a) $\lambda + \|p(t)\| > 0$ при всех $t \in [0,\bar{T}]$ (*нетривиальность*);

b) $\dot{p}(t) = -H_x\big(\bar{x}(t),\bar{u}(t),p(t)\big)$ при всех t (*сопряжённое уравнение*);

c) $H_x\big(\bar{x}(t),\bar{u}(t),p(t)\big) = \max\limits_{u\in U_m(t)} H\big(\bar{x}(t),u,p(t)\big)$ почти всюду (*принцип максимума*);

d) $H\big(\bar{x}(t),\bar{u}(t),p(t)\big) = \lambda$ почти всюду на $[0,\bar{T}]$ (*условие трансверсальности*).

Для доказательства принципа максимума Понтрягина в полном объёме нам осталось проверить, что максимум по $U_m(t)$ в c) можно заменить максимумом по всему множеству U и что в условии трансверсальности равенство выполнено во всех точках отрезка. Второе есть немедленное следствие первого ввиду непрерывности f и компактности U. Первое же доказывается с помощью стандартного рассуждения, подобного использованному в доказательстве теоремы 4.1. Множество $\mathcal{P}_m$ пар $\big(\lambda, p(\cdot)\big)$, удовлетворяющих (a)—(d) и нормированных условием
$$\lambda + \max_{t\in[0,\bar{T}]} \|p(t)\| = 1,$$
компактно в слабой топологии пространства $\mathbb{R}\times W^{1,1}$. Следовательно, пересечение этих множеств не пусто. Это значит, что мы можем распространить максимум в (c) на объединение множеств $U_m(t)$, обозначим его $U(t)$. Но $f\big(\hat{x}(t),U(t)\big)$ плотно в $f(\hat{x},U)$ почти при всех t по выбору $u_i(\cdot)$, и непрерывность f позволяет перейти к максимуму по всему U. Этим завершается доказательств принципа максимума Понтрягина для классической задачи о быстродействии автономной системы управления.

6. Оптимальное управление дифференциальным включением

Мы рассматриваем следующую задачу: минимизировать
$$\ell\big(x(0),x(T)\big) \tag{6.1}$$
на траекториях дифференциального включения
$$\dot{x} \in F(t,x), \tag{6.2}$$
удовлетворяющих концевым условиям
$$\big(x(0),x(T)\big) \in S. \tag{6.3}$$
Под траекторией (6.2) мы, как обычно, понимаем абсолютно непрерывную функцию $x(t)$, такую что $\dot{x}(t) \in F\big(t,x(t)\big)$ почти при всех t. Нас будут интересовать

необходимые условия локального минимума в пространстве $W^{1,1}(0,T)$. Напомним, что норма в этом пространстве может быть, например, задана следующим образом:

$$\|x(\cdot)\|_{1,1} = \|x(0)\| + \int\limits_0^T |\dot{x}(t)|\,dt.$$

В дальнейшем мы фиксируем некоторую траекторию $\bar{x}(t)$, удовлетворяющую (6.2), (6.3), и предполагаем, что выполнены следующие условия:

(A_1) ℓ локально липшицева; F измеримо по t и график $F(t,\cdot)$ замкнут почти при всех t; S замкнутое множество в $\mathbb{R}^n \times \mathbb{R}^n$;

(A_2) существуют $\rho > 0$, $\varepsilon > 0$, такие что $F(t,x) \cap B(\dot{\bar{x}}(t),\rho) \neq \varnothing$, если $\|x - \bar{x}(t)\| < \varepsilon$ при всех t из некоторого множества полной меры;

(A_3) для всякого достаточно большого R существуют $\varepsilon > 0$ и суммируемая функция $k_R(t)$, такие что почти при всех t

$$F(t,x) \cap B(\dot{\bar{x}}(t),R) \subset F(t,x') + k_R(t)\|x - x'\|B$$

если $x, x' \in B(\bar{x}(t),\varepsilon)$.

Последние два условия заслуживают некоторого комментария. Оба в какой-то степени ограничивают способность F меняться слишком резко с изменением x. В соответствии с более или менее устоявшейся терминологией (A_3) означает, что $F(t,\cdot)$ *псевдолипшицево* радиуса R и с константой $k(t)$ на $B(\bar{x}(t),\varepsilon)$. Легко заметить, что ($\mathrm{A}_2$) следует из ($\mathrm{A}_3$), если $k_R(t)$ ограниченная.

Наша цель — доказать следующую теорему.

Теорема 6.1. *Предположим что выполнены условия* (A_1)—(A_3). *Если* $\bar{x}(\cdot)$ *доставляет локальный минимум в задаче* (6.1)—(6.3) *(в сильной топологии пространства* $W^{1,1}$*), то найдутся* $\lambda \geqslant 0$ *и абсолютно непрерывная функция* $p(\cdot)$ *со значениями в* $\mathbb{R}^n$*, такие что*

$$\bigl(\lambda_0, p(\cdot)\bigr) \neq 0$$

(*нетривиальность*);

$$\bigl(p(0), -p(T)\bigr) \in \lambda_0 \partial \ell\bigl(\bar{x}(0), \bar{x}(T)\bigr) + N\Bigl(S, \bigl(\bar{x}(0), \bar{x}(T)\bigr)\Bigr)$$

(*условие трансверсальности*);

$$\dot{p}(t) \in \operatorname{conv}\Bigl\{q\colon \bigl(q, p(t)\bigr) \in N\Bigl(\operatorname{Graph} F_t, \bigl(\bar{x}(t), \dot{\bar{x}}(t)\bigr)\Bigr)\Bigr\}$$

(*включение Эйлера—Лагранжа*);

$$\langle p(t), y\rangle \leqslant \langle p(t), \dot{\bar{x}}(t)\rangle \text{ для всех } y \in F\bigl(t, \bar{x}(t)\bigr) \text{ и для почти всех } t$$

(*принцип максимума*).

Мы выведем теорему из следующего более специализированного результата.

Теорема 6.2. *Предположим, что выполнены условия* (A_1) *и* (A_2) *наряду с условием*

(A_3') *найдутся* $\eta > 0$, $\varepsilon > 0$, *измеримая функция* $R(t)$ *и суммируемая функция* $k(t)$, *такие что почти при всех* t

$$(1-2\eta)R(t) > \rho, \quad F(t,x) \cap B\bigl(\dot{\bar{x}}(t), R(t)\bigr) \subset F(t,x') + k(t)\|x-x'\|B,$$

если $x, x' \in B(\bar{x}(t), \varepsilon)$.

Положим $f_t(x,y) = d\bigl(y, F(t,x)\bigr)$. *Если* $\bar{x}(\cdot)$ *доставляет локальный минимум в задаче* (6.1)—(6.3), *то найдутся* $\lambda_0 \geqslant 0$ *и абсолютно непрерывная функция* $p(\cdot)$ *со значениями в* $\mathbb{R}^n$, *удовлетворяющие условию нетривиальности теоремы* 6.1 *и такие, что*

$$\bigl(p(0), -p(T)\bigr) \in \partial\bigl(\lambda_0 \ell + d(\cdot, S)\bigr)\bigl(\bar{x}(0), \bar{x}(T)\bigr), \tag{6.4}$$

$$\dot{p}(t) \in \operatorname{conv}\bigl\{q\colon \bigl(q, p(t)\bigr) \in \partial f_t(\cdot)\bigl(\bar{x}(t), \dot{\bar{x}}(t)\bigr)\bigr\}, \tag{6.5}$$

$$\langle p(t), y\rangle \leqslant \langle p(t), \dot{\bar{x}}(t)\rangle \quad \textit{для всех} \quad y \in F\bigl(t, \bar{x}(t)\bigr) \cap (1-\eta)R(t). \tag{6.6}$$

Сформулированный результат близок к «стратифицированному принципу максимума», полученному Ф. Кларком в [13]. Приведённая формулировка несколько отличается от содержащейся в [13], где η отсутствует. (Впрочем, модификация приводимого ниже доказательства, позволяющая получить результат Ф. Кларка в полном объёме не представляет труда.) С другой стороны, условия (6.4), (6.5) у нас несколько точнее: в [13] они имеют ту же форму, что и условие трансверсальности и включение Эйлера—Лагранжа из предыдущей теоремы.

Итак, предположим, что теорема 6.2 уже доказана. Это значит, что для всякого достаточно большого R и некоторого фиксированного малого η мы располагаем нетривиальной парой $\bigl(\lambda_0, p(\cdot)\bigr)$, удовлетворяющей условиям (6.4)—(6.6). Заметим, что все четыре соотношения в утверждении теоремы 6.1 однородны относительно $\bigl(\lambda_0, p(\cdot)\bigr)$. Поэтому если мы нормируем такую пару, разделив, например, на $\lambda_0 + \max\limits_{t\in[0,T]} \|p(t)\|$, то полученная новая пара будет по-прежнему удовлетворять всем соотношениям. Обозначим через $\mathcal{L}(R)$ замыкание в $\mathbb{R} \times W^{1,1}$ совокупности таким образом полученных нормированных пар. Ясно, что и условие трансверсальности, и включение Эйлера—Лагранжа при замыкании сохраняются, т. е. элементы $\mathcal{L}(R)$ также обладают всеми четырьмя свойствами из теоремы 6.1.

В силу (A_3) $F(t,\cdot)$ почти при каждом t обладает свойством Обэна в точке $\bigl(\bar{x}(t), \dot{\bar{x}}(t)\bigr)$ с константой $k(t)$, где в качестве $k(t)$ можно взять любую из функций $k_R(t)$. Поэтому по предложению 2.5 $\|q\| \leqslant k(t)\|p\|$, если $(q,p) \in N\Bigl(\operatorname{Graph} F(t,\cdot), \bigl(\bar{x}(t), \dot{\bar{x}}(t)\bigr)\Bigr)$, т. е., в частности, $\|\dot{p}(t)\| \leqslant k(t)\|p(t)\|$, если $\bigl(\lambda, p(\cdot)\bigr) \in \mathcal{L}(R)$. Отсюда немедленно следует, что множества $\mathcal{L}(R)$ компактны в слабой топологии пространства $\mathbb{R} \times W^{1,1}$. С другой стороны, ясно, что множества $\mathcal{L}(R)$ с ростом R могут только убывать. Следовательно, пересечение этих множеств на пусто. Каждый элемент в пересечении удовлетворяет свойствам нетривиальности и трансверсальности, равно как и включению Эйлера—Лагранжа и принципу максимума для всякого R. Этим завершается доказательство теоремы 6.1.

7. Доказательство теоремы 6.2

Доказательство основано на сведении задачи с помощью оптимизационной альтернативы к серии вариационных задач Больцы, подобных рассмотренной в разделе 4.

7.1. Построение пробной функции

Мы начинаем доказательство с построения пробной функции типа функционала Больцы с интегрантом, удовлетворяющими условию Липшица. Для этой цели мы используем упрощённый вариант конструкции, предложенной Ф. Кларком в [13].

Лемма 7.1. *Рассмотрим многозначное отображение* $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$. *Пусть* $\bar{y} \in G(\bar{x})$ *и существуют* $\varepsilon > 0$, $k > 0$, $R > r > 0$, *такие что при* $\|x, x' - \bar{x}\| < \varepsilon$ *выполнены соотношения*

$$G(x) \cap B(\bar{y}, r) \neq \varnothing, \quad G(x) \cap B(\bar{y}, R) \subset G(x') + k\|x' - x\|B. \tag{7.1}$$

Пусть $\eta > 0$ *таково, что* $(1 - 2\eta)R > r$. *Определим многозначное отображение* $\Gamma\colon \mathbb{R}^n \rightrightarrows \mathbb{R} \times \mathbb{R}^m$:

$$\Gamma(x) = \{(\lambda, y)\colon \lambda \in [0, 1 + \eta],\ y = \lambda(z - \bar{y}),\ z \in G(x) \cap B(\bar{y}, (1 - \lambda\eta)R)\}. \tag{7.2}$$

Тогда Γ *липшицево (относительно метрики Хаусдорфа) на* $B(\bar{x}, \varepsilon)$ *с константой, не превышающей*

$$c = k\eta^{-1} \max\{R^{-1}, 1 + \eta + \eta^2\}.$$

Именно, если $\|x, x' - \bar{x}\| < \varepsilon$ *и* $\mathbb{R} \times \mathbb{R}^m$ *рассматривается с нормой* $\|(\alpha, z)\| = |\alpha| + \|z\|$, *то*

$$\Gamma(x) \subset \Gamma(x') + c\|x - x'\|B.$$

Доказательство. Как следует из (7.1), $G(x) \neq \varnothing$, и значит, $\Gamma(x) \neq \varnothing$ для всякого $x \in B(\bar{x}, \varepsilon)$ и $(0, 0) \in \Gamma(x)$ по определению. Для доказательства леммы нам нужно для всяких x, x', ε-близких к $\bar{x}$, и $(\lambda, y) \in \Gamma(x)$ подобрать $(\lambda', y') \in \Gamma(x')$ так, чтобы

$$\|(\lambda, y) - (\lambda', y')\| \leqslant c\|x - x'\|.$$

Положим $\delta = \|x - x'\|$, и пусть $y = \lambda(z - \bar{y})$, где $z \in G(x)$, $\|z - \bar{y}\| \leqslant (1 - \lambda\eta)R \leqslant R$. Если $\lambda \leqslant \delta k(\eta R)^{-1} < c\delta$, то $\|y\| = \lambda\|z - \bar{y}\| \leqslant \lambda R \leqslant \delta k/\eta$, и мы можем взять $\lambda' = 0$, $y' = 0$.

Если же $\lambda > \delta k(\eta R)^{-1}$, то, поскольку по условиям леммы

$$z \in G(x) \cap B(\bar{y}, R) \subset G(x') + k\delta B,$$

мы можем найти такое $z' \in G(x')$, что $\|z - z'\| \leqslant k\delta$.

Положим $\lambda' = \lambda - (\eta R)^{-1}\|z - z'\|$, $y' = \lambda'(z' - \bar{y})$. Тогда $0 < \lambda' < \lambda \leqslant 1 + \eta$ и

$$\|z' - \bar{y}\| \leqslant \|z - \bar{y}\| + \|z - z'\| \leqslant (1 - \lambda\eta)R + (\lambda - \lambda')\eta R = (1 - \lambda'\eta)R,$$

т. е. $(\lambda', y') \in \Gamma(x')$. С другой стороны, так как $|\lambda - \lambda'| \leqslant k\delta(\eta R)^{-1}$ и $\|z' - \bar{y}\| < R$,

$$\|y - y'\| = \|\lambda(z - \bar{y}) - \lambda'(z' - \bar{y})\| \leqslant \lambda\|z - z'\| + \|z' - \bar{y}\|\,|\lambda - \lambda'| \leqslant$$
$$\leqslant (1 + \eta)k\delta + \left(\frac{k}{\eta}\right)\delta = \left(\frac{k}{\eta}\right)(1 + \eta + \eta^2)\delta,$$

т. е. $\|(\lambda, y) - (\lambda', y')\| \leqslant c\delta$, что и требовалось. □

В дальнейшем мы рассматриваем многозначное отображение

$$\Gamma_t \colon [0, T] \times \mathbb{R}^n \rightrightarrows \mathbb{R} \times \mathbb{R}^n,$$

определяемое при каждом t соотношением (7.2), в котором роль G играет $F(t, \cdot)$ и $R = R(t)$.

Положим

$$\varphi\big(\mu(\cdot), x(\cdot)\big) = d\Big(\big(x(0), x(T)\big), S\Big) + \int\limits_0^T d\Big(\big(\dot{\mu}(t), \dot{x}(t)\big), \Gamma_t\big(x(t)\big)\Big)\, dt. \tag{7.3}$$

Это и есть искомая пробная функция. В дальнейшем, когда это удобно, мы считаем (очевидным образом, не нарушая общности), что $\bar{x}(t) \equiv 0$.

По лемме интегрант функционала φ почти при каждом t удовлетворяет условию Липшица по (μ, x) с константой

$$c(t) = k(t)\eta^{-1} \max\{R(t)^{-1}, 1 + \eta + \eta^2\}.$$

(Ясно, что $c(t)$ — суммируемая функция.) Цена, которую за это пришлось заплатить — введение дополнительного переменного μ. Очевидно, функция φ неотрицательна и $\varphi\big(\mu(\cdot), \bar{x}(\cdot)\big) = 0$ при всяком μ с производной $|\dot{\mu}(t)| \leqslant 1 + \eta$ почти везде.

В дальнейшем мы фиксируем $\bar{\mu}(t) \equiv t$. По лемме φ удовлетворяет условию Липшица на $W^{1,1}$, и мы можем применить альтернативу в полном объёме. Это значит, что либо реализуется регулярная ситуация и существует $\lambda_0 > 0$, такое что $\big(\bar{\mu}(\cdot), \bar{x}(\cdot)\big)$ есть точка локального минимума функционала

$$\Phi\big(\mu(\cdot), x(\cdot)\big) =$$
$$= \lambda_0 \ell\big(x(0), x(T)\big) + d\Big(\big(x(0), x(T)\big), S\Big) + \int\limits_0^T d\Big(\big(\dot{\mu}(t), \dot{x}(t)\big), \Gamma_t\big(x(t)\big)\Big)\, dt$$

(в сильной топологии пространства $W^{1,1}$), либо существует последовательность $\big(\mu_m(\cdot), x_m(\cdot)\big) \subset W^{1,1}$, сходящаяся по норме к $\big(\bar{\mu}(\cdot), \bar{x}(\cdot)\big)$, такая что при всяком m

а) или $\big(x_m(0), x_m(T)\big) \notin S$, или $d\Big(\big(\dot{\mu}_m(t), \dot{x}_m(t)\big), \Gamma_t\big(x(t)\big)\Big) > 0$ на множестве положительной меры, или и то и другое и при этом б) функционал

$$\Phi_m\big(\mu(\cdot),x(\cdot)\big)=$$
$$= d\Big(\big(x(0),x(T)\big),S\Big)+m^{-1}(\|x(0)-x_m(0)\|+|\mu(0)-\mu_m(0)|)+$$
$$+\int\limits_0^T\Big[d\Big(\big(\dot\mu(t),\dot x(t)\big),\Gamma_t\big(x(t)\big)\Big)+m^{-1}(|\dot\mu(t)-\dot\mu_m(t)|)+\|\dot x(t)-\dot x_m(t)\|\Big]\,dt$$

имеет в $\big(\mu_m(\cdot),x_m(\cdot)\big)$ локальный минимум в сильной топологии $W^{1,1}$.

7.2. Анализ

Легко убедиться, что и Φ, и все Φ_m удовлетворяют условиям теоремы 4.1. Мы начнём расшифровку соотношений, предлагаемых теоремой 4.1, с анализа субдифференциала интегранта $d\big((\nu,y),\Gamma_t(x)\big)$, когда x близко к $\bar x(t)$, y близко к $\dot{\bar x}(t)$ и ν близко к единице.

Имеем

$$\begin{aligned}&d\big((\nu,y),\Gamma_t(x)\big)=\\&\quad=\inf\{|\nu-\lambda|+\|y-\lambda z\|\colon \lambda\in[0,1+\eta],\ z\in F(t,x)\cap(1-\lambda\eta)R(t)B\}\leqslant\\&\quad\leqslant\inf\{\nu\|\nu^{-1}y-z\|\colon z\in F(t,x)\cap(1-\nu\eta)R(t)B\}=\\&\quad=\nu d(\nu^{-1}y,F(t,x)\cap(1-\nu\eta)R(t)B).\end{aligned}\tag{7.4}$$

Фиксируем некоторое малое положительное $\delta<\min\{1,\rho\}$ (тогда, очевидно, $\delta<(1-(1+\delta)\eta)R(t))$, и пусть

$$\|x\|<\min\left\{\varepsilon,\frac{\delta}{3}k(t)\right\},\quad \|y\|<\frac{\delta}{3},\quad |\nu^{-1}-1|<\frac{\delta}{3}.$$

Тогда согласно (A_3')

$$d\big(\nu^{-1}y,F(t,x)\big)\leqslant|1-\nu^{-1}|\,\|y\|+d\big(y,F(t,x)\big)\leqslant\frac{\delta^2}{9}+k(t)\|x\|+\|y\|<\delta,$$

так что ближайшая к $\nu^{-1}y$ точка $F(t,x)$ лежит строго внутри шара радиуса $(1-\nu\eta)R(t)$, и следовательно,

$$d\big((\nu,y),\Gamma_t(x)\big)\leqslant\nu d\big(\nu^{-1}y,F(t,x)\big)<\delta\nu.$$

Мы можем считать δ достаточно малым, чтобы гарантировать, что и при λ, отличающихся от ν менее чем на δ, ближайшая к $\lambda^{-1}y$ точка $F(t,x)$ лежит строго внутри шара радиуса $(1-\lambda\eta)R(t)$, т. е.

$$\inf\{\|y-\lambda z\|\colon z\in F(t,x)\cap(1-\nu\eta)R(t)B\}=\lambda d\big(\lambda^{-1}y,F(t,x)\big)<\delta\lambda.$$

(Заметим, что левая часть неравенства в любом случае не может быть меньше правой.) Но в этом случае инфимум в правой части первого равенства в (7.4) достигается при $\lambda=\nu$, и мы имеем на самом деле равенство:

$$d\big((\nu,y),\Gamma_t(x)\big)=\nu d\big(\nu^{-1}y,F(t,x)\big)<\delta\nu.\tag{7.5}$$

Определим функции g и ψ на $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ и оператор

$$A\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$$

соотношениями

$$g_t(x, y, \nu) = \nu d\big(\nu^{-1} y, F(t, x)\big), \quad \psi_t(x, w, \nu) = \nu d\big(w, F(t, x)\big),$$
$$A(x, y, \nu) = (x, \nu^{-1} y, \nu).$$

Тогда $g_t = \psi_t \circ A$, оператор A непрерывно дифференцируем в любой точке с $\nu > 0$ и его производная $A'(x, y, z)$ — сюръективный линейный оператор в $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$. Тогда по теореме 2.2 для $\mathbf{z} = (x, y, \nu)$ с $\nu > 0$ справедливо равенство $\partial g_t(\mathbf{z}) = A'^* \partial \psi_t\big(A(\mathbf{z})\big)$, т. е.

$$\partial g_t(x, y, \nu) = \big\{\big(q, \nu^{-1} p, \nu^{-2}(\langle p, y\rangle + \xi)\big) \colon (q, p, \xi) \in \partial \psi_t(x, \nu^{-1} y, \nu)\big\}. \tag{7.6}$$

Наконец,

$$\partial \psi_t(x, w, \nu) = \big\{(q, p, \xi) \colon (q, p) \in \partial d\big(\cdot, F(t, \cdot)\big)(x, w),\ \xi = d\big(w, F(t, x)\big)\big\}. \tag{7.7}$$

Подчеркнём ещё раз, что (7.6) и (7.7) справедливы, лишь если x, y, ν достаточно близки к $\bar{x}(t) = 0$, $\dot{\bar{x}}(t) = 0$, $\dot{\bar{\mu}}(t) = 1$ соответственно.

Теперь мы можем перейти непосредственно к доказательству теоремы.

7.3. Регулярный случай

Так как Φ имеет локальный минимум в $\big(\bar{\mu}(\cdot), \bar{x}(\cdot)\big)$, существует пара абсолютно непрерывных функций $\big(\xi(\cdot), p(\cdot)\big)$, обладающая тремя свойствами, декларируемыми теоремой 4.1. Условие трансверсальности в нашем случае имеет вид

$$\big(p(0), p(T)\big) \in \partial\big(\lambda_0 \ell + d(\cdot, S)\big)\big(x(0), x(T)\big), \quad \xi(0) = \xi(T) = 0,$$

последнее соотношение справедливо, так как терминальный член в Φ не зависит от μ, а первое соотношение — это в точности (6.4).

Включение Эйлера—Лагранжа (6.5) следует из второго соотношения в теореме 4.1. Действительно, заметим прежде всего, что и интегрант в Φ не зависит явно от μ, а зависит только от производной $\dot{\mu}$. Поэтому $\dot{\xi}(t) \equiv 0$, т. е. $\xi(t) = \mathrm{const} = 0$, и следовательно, с учётом (7.5) условие Эйлера—Лагранжа принимает вид

$$(\dot{p}(t), 0) \in \operatorname{conv}\big\{(q, \xi) \colon (q, p, \xi) \in \partial g_t\big(\bar{x}(t), \dot{\bar{x}}(t), \dot{\bar{\mu}}(t)\big)\big\}.$$

Требуемое включение (6.5) теперь следует из (7.6), (7.7) с учётом того, что $\bar{\mu}(t) = t$ и $g_t(x, y, 1) = f_t(x, y)$.

Наконец, из последнего соотношения в теореме 4.1 (очевидным образом классического условия Вейерштрасса), применённого к Φ (опять же с учётом того, что $\xi(t) \equiv 0$), получаем, что $\langle p(t), \dot{\bar{x}}(t)\rangle \geqslant \langle p(t), y\rangle$ для всякого $y \in F\big(t, \bar{x}(t)\big)$ с нормой, не превосходящей $(1 - \eta)R$.

Итак, если реализуется регулярная ситуация, то найдутся $\lambda_0 > 0$ и абсолютно непрерывная $p(\cdot)$, такие что утверждение теоремы 6.2 верно.

7.4. Нерегулярный случай

Как и в регулярном случае, из теоремы 4.1 следует существование пары $\big(\xi_m(\cdot), p_m(\cdot)\big)$, обладающей определёнными свойствами. В частности $\dot{\xi_m}(t) = 0$, поскольку интегрант в Φ_m зависит лишь от $\dot{\mu}$. Таким образом, $\xi_m(t) = \mathrm{const}$ и в данном случае. Условие трансверсальности для Φ_m имеет вид: существует u с нормой, не превосходящей m^{-1}, такое что

$$\big(p_m(0) + u, -p_m(T)\big) \in \partial d(\cdot, S)\big(x_m(0), x_m(T)\big), \quad |\xi_m(0)| \leqslant m^{-1}.$$

Следовательно, $|\xi_m(t)| \leqslant m^{-1}$ при всех t.

Полагая $h(x, y, \nu) = d\big((\nu, y), \Gamma_t(x)\big)$, мы можем записать включение Эйлера—Лагранжа в виде

$$\dot{p}_m(t) \in \mathrm{conv}\big\{q\colon \big(q + v(t), p_m(t), \xi_m(t) + \zeta(t)\big) \in \partial h\big(x_m(t), \dot{x}_m(t), \dot{\mu}_m(t)\big)\big\},$$

где $\|v(t)\| \leqslant m^{-1}$, $|\zeta|(t) \leqslant m^{-1}$. Если при этом $\big(x_m(0), x_m(T)\big) \notin S$, то $\max\{\|p_m(0) + u\|, \|p_m(T)\|\} = 1$, т. е. $\max\{\|p_m(0)\|, \|p_m(T)\| \geqslant 1 - m^{-1}$.

Если же $\big(\dot{\mu}_m(t), \dot{x}_m(t)\big) \notin \Gamma_t\big(x_m(t)\big)$ на множестве положительной меры, то почти всюду на этом множестве $\max\{|\xi_m(t)|, \|p_m(t)\|\} \geqslant 1 - m^{-1}$. Таким образом, при $m > 2$ необходимо $p_m(t) \geqslant 1 - 2m^{-1}$ при некотором $t \in [0, 1]$. С другой стороны, поскольку $h_t(\cdot, y, \nu)$ удовлетворяет условию Липшица с константой $c(t)$ (не зависящей от m), $\|\dot{p}_m(t)\| \leqslant c(t)$ почти всюду для каждого m.

Отсюда следует, что последовательность $\big(p_m(\cdot)\big)$ слабо компактна в $W^{1,1}$, и можно считать, что она сходится в $W^{1,1}$ (и значит, равномерно) к некоторой $p(\cdot)$, такой что $\|\dot{p}(t)\| \leqslant c(t)$ почти для всех t и $\|p(t)\| = 1$ хотя бы в одной точке. Очевидно, $\big(p(0), -p(T)\big) \in \partial d(\cdot, S)\big(\bar{x}(0), \bar{x}(T)\big)$. С другой стороны, поскольку $\big(\dot{\mu}_m(t), \dot{x}_m(t)\big) \to \big(1, \dot{\bar{x}}(t)\big)$ почти везде, стандартные рассуждения, основанные на теореме Мазура (связанные с рассмотрением последовательности выпуклых комбинаций $p_m(\cdot)$, сходящихся к $p(\cdot)$ сильно — см. доказательство теоремы 4.1) и полунепрерывности сверху предельного субдифференциала, позволяют заключить, как и в регулярном случае, с учётом (7.6) и (7.7), что $p(\cdot)$ удовлетворяет включению Эйлера—Лагранжа (6.5).

Наконец, из условия Вейерштрасса теоремы 4.1 получаем, что

$$\langle p_m(t), \dot{x}_m(t)\rangle + \xi_m(t)\dot{\mu}_m(t) \geqslant \langle p_m(t), y\rangle + \xi_m(t)\lambda \quad \text{для всех} \quad (\lambda, y) \in \Gamma_t(x_m(t))$$

почти при всех t. Переходя к пределу при $m \to \infty$, получаем тот же принцип максимума, что и в регулярном случае. Теорема полностью доказана.

8. Несколько заключительных замечаний

1. Идея релаксации (скользящих режимов) в контексте необходимых условий в оптимальном управлении восходит к работам Р. В. Гамкрелидзе и Дж. Варги 1962 года [3, 35]. (Похожие идеи в контексте достаточных условий примерно в то же время развивались в работах В. Ф. Кротова [8].) Но

для задач классического вариационного исчисления возможность «овыпукления» была отмечена Н. Н. Боголюбовым уже в 1930 году [9]. Для проверки релаксируемости задачи Больцы из раздела 4 нужна лишь незначительная модификация доказательства теоремы Боголюбова (например, приведённого в [7]).

В абстрактной ситуации релаксируемость впервые обсуждалась, видимо, в работах Х. Халкина и Л. Нойштадта [18, 29]. Дальнейшее развитие этих идей можно проследить в [1, 7, 16]. В этих работах рассматриваются отображения $F(x,u)$ со смешанной структурой: гладкие по x и допускающие «смешивание» разных значений u в следующем смысле. Для всякого $\delta > 0$ и всякого конечного набора $(u_1,\ldots,u_m)$ элементов U существует отображение $M_\delta(x,\alpha_1,\ldots,\alpha_m)$ (в U), такое что $F\big(x, M_\delta(x,\alpha_1,\ldots,\alpha_k)\big)$ аппроксимирует $F(x,\bar u) + \sum \alpha_i\big(F(x,u_i) - F(x,\bar u)\big)$ в том или ином смысле «с точностью до δ».

Предлагаемый здесь подход (являющийся некоторым видоизменением предложенного ранее в [24]) подчёркивает решающую роль линейных операторов типа дифференциальных (или им обратных) для практического осуществления подобного смешивания. При этом отпадает нужда в конструировании какого-либо специального аппроксимирующего отображения.

2. Подчеркнём ещё раз, что доказательство теоремы 6.1 и доказательство классического принципа максимума в разделе 5 — это варианты одного и того же доказательства, отличающиеся лишь техническими деталями, связанными с различиями рассматриваемых задач.

Разумеется, принцип максимума для стандартной задачи оптимального управления является простым следствием теоремы 6.1(как, впрочем, и более ранних результатов для управляемых систем с дифференциальными включениями [22, 34]). Действительно, пусть уравнение управляемого объекта имеет вид

$$\dot x = f(t,x,u), \quad u \in U.$$

Положим $F(t,x) = f(t,x,U)$. Если U ограниченное, то при естественных предположениях отображение $F(t,x)$ удовлетворяет всем условиям теоремы 6.1. Если же U неограниченное, то следует рассмотреть последовательность задач, в которых U заменено на $U_k = U \cap B(0,k)$, а затем стандартным образом использовать тот факт, что совокупность множителей $\big(\lambda_0, p(\cdot)\big)$, обеспечивающих выполнение принципа максимума, не возрастает с ростом k. Другой способ редукции (технически, возможно, наиболее простой) предполагает рассмотрение отображений $F(t,x,\Omega)$, соответствующих объединениям $\Omega \subset U$ всевозможных конечных подмножеств U с оптимальным управлением, как разделе 5.

3. На самом деле теорема 6.1 — существенно более сильный результат, чем стандартный принцип максимума, ибо она позволяет оптимизировать управление системами с обратной связью (т. е. такими, в которых возможные управления зависят от состояния системы), по крайней мере если эта связь в той или иной мере является липшицевой. Для иллюстрации рассмотрим автономную систему $\dot x = f(x,u)$, $u \in U(x)$, предполагая, что f удовлетворяет условию

Липшица в окрестности каждой точки, а многозначное отображение $U(x)$ является липшицевым относительно метрики Хаусдорфа. И в этом случае $F(x) = f\big(x, U(x)\big)$ удовлетворяет условиям теоремы 6.1.

4. Неясно, можно ли применить первый описанный во введении подход (связанный с использованием критерия ненакрывания) к доказательству теоремы 6.1. В любом случае, как кажется, это доказательств технически будет более сложным, поскольку оно неизбежно будет связано с оценкой копроизводных некоторого многозначного отображения пространства $W^{1,1}$, скажем соотносящего каждому $x(\cdot) \in W^{1,1}$ множество

$$x(t) - x(0) - \int_0^t F\big(\tau, x(\tau)\big)\, d\tau,$$

(под первообразной многозначного отображения, как обычно, понимается совокупность первообразных его измеримых селекторов), или, например, подобного отображения, построенного с использованием леммы 7.1.

Литература

[1] Аваков Е. Р., Магарил-Ильяев Г. Г., Тихомиров В. М. О принципе Лагранжа для задач на экстремум при наличии ограничений // УМН. — 2013. — Т. 68, № 3. — С. 5—38.

[2] Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. — М.: Наука, 1979.

[3] Гамкрелидзе Р. В. Оптимальные скользящие режимы // ДАН СССР. — 1962. — Т. 143. — С. 1243—1245.

[4] Дмитрук А. В., Милютин А. А., Осмоловский Н. П. Теорема Люстерника и теория экстремума // УМН. — 1980. — Т. 35, № 6. — С. 3—46.

[5] Дубовицкий А. Я., Милютин А. А. Задачи на экстремум при наличии ограничений // Ж. вычисл. матем. и матем. физ. — 1965. — Т. 5, № 3. — С. 395—453.

[6] Иоффе А. Д. Метрическая регулярность и субдифференциальное исчисление // УМН. — 2000. — Т. 55, № 3. — С. 103—162.

[7] Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. — М.: Наука, 1974.

[8] Кротов В. Ф. Методы решения вариационных задач на основе достаточных условий абсолютного минимума // Автоматика и телемеханика. — 1962. — Т. 23. — С. 1571—1583; 1963. — Т. 24. — С. 581—598.

[9] Bogolyubov N. N. Sur quelques méthodes nouvelles dans le calcul des variations // Ann. Math. Pure Appl. Ser. 4. — 1930. — Vol. 7. — P. 249—271.

[10] Borwein J. M., Zhu J. Techniques of Variational Analysis. — Berlin: Springer, 2006. — (CMS Books Math.; Vol. 20).

[11] Clarke F. H. A new approach to Lagrange multipliers // Math. Oper. Res. — 1976. — Vol. 1. — P. 165—174.

[12] Clarke F. H. Optimization and Nonsmooth Analysis. — Wiley; Interscience, 1983.

[13] Clarke F. H. Necessary Conditions in Dynamic Optimization. — Amer. Math. Soc., 2005. — (Mem. Amer. Math. Soc.; Vol. 816).

[14] Clarke F. H., Ledayev Yu. S., Stern R. J., Wolenski P. R. Nonsmooth Analysis and Control Theory. — Springer, 1998.

[15] Dontchev A. L., Rockafellar R. T. Implicit Function and Solution Mapping. — Springer, 2009.

[16] Ginsburg B., Ioffe A. D. Maximum principle for general semilinear optimal control problems // Nonsmooth Analysis and Geometric Methods in Optimal Control / B. Mordukhovich, H. Sussmann, eds. — Springer, 1995. — P. 81—110.

[17] Graves L. M. Some mapping theorems // Duke Math. J. — 1950. — Vol. 17. — P. 111—114.

[18] Halkin H., Neustadt L. General necessary conditions for optimization problems // Proc. Nat. Acad. Sci. — 1966. — Vol. 56. — P. 1066—1071.

[19] Ioffe A. D. Necessary and sufficient conditions for a local minimum. 1—3 // SIAM J. Control Optim. — 1979. — Vol. 17. — P. 245—288.

[20] Ioffe A. D. On subdifferentiability spaces // Ann. N.Y. Acad. Sci. — 1983. — Vol. 410. — P. 107—121.

[21] Ioffe A. D. On the local surjection property // Nonlinear Anal. — 1987. — Vol. 11. — P. 565—592.

[22] Ioffe A. D. Euler—Lagrange and Hamiltonian formalisms in dynamic optimization // Trans. Amer. Math. Soc. — 1997. — Vol. 349. — P. 2871—2900.

[23] Ioffe A. D. Fuzzy principles and characterization of trustworthiness // Set-Valued Analysis. — 1998. — Vol. 6. — P. 265—276.

[24] Ioffe A. D. Optimality alternative: a non-variational approach to necessary conditions // Variational Analysis and Applications (a volume dedicated to the memory of G. Stampacchia) / F. Giannessi, A. Maugeri, eds. — Kluwer, 2005. — P. 531—552.

[25] Ioffe A. D. On the general theory of subdifferentials // Adv. Nonlinear Anal. — 2012. — Vol. 1. — P. 47—120.

[26] Ioffe A. D., Rockafellar R. T. The Euler and Weierstrass conditions for nonsmooth variational problems // Calculus of Variations and PDEs. — 1996. — Vol. 4. — P. 59—87.

[27] Lassonde M. First order rules for nonsmooth constrained optimization // Nonlinear Analysis. TMA. — 2001. — Vol. 44. — P. 1031—1056.

[28] Loewen P. D. Optimal Control via Nonsmooth Analysis. — Amer. Math. Soc., 1993. — (CRM Proc. & Lect. Notes; Vol. 2).

[29] Neustadt L. An abstract variational theory with applications // SIAM J. Control. — 1966. — Vol. 4. — P. 505—527; 1967. — Vol. 5. — P. 90—137.

[30] Penot J. P. Calculus Without Derivatives. — Springer, 2012. — (Grad. Texts Math.; Vol. 266).

[31] Robinson S. M. Stability theory for system of inequalities. II: differentiable nonlinear systems // SIAM J. Num. Anal. — 1976. — Vol. 13. — P. 497—513.

[32] Rockafellar R. T., Wets R. J. B. Variational Analysis. — Springer, 1998.

[33] Tikhomirov V. M. Fundamental Principles of the Theory of Extremal Problems. — Wiley, 1982.

[34] Vinter R. B. Optimal Control. — Birkhäuser, 2000.

[35] Warga J. Relaxed variational problems // J. Math. Anal. Appl. — 1962. — Vol. 4. — P. 111—128.

[36] Warga J. Controllability and necessary conditions in unilateral problems without differentiability assumptions // SIAM J. Control Optim. — 1976. — Vol. 14. — P. 546—573.

[37] Yorke J. A. The maximum principle and controllability of nonlinear equations // SIAM J. Control Optim. — 1972. — Vol. 10. — P. 334—338.

[38] Zhu Q. J. The equivalence of several basic theorems for subdifferentials // Set-Valued Analysis. — 1998. — Vol. 6. — P. 171—185.