С. Г. МИХЛИН

Численная реализация вариационных методов

С. Г. МИХЛИН

ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ ВАРИАЦИОННЫХ МЕТОДОВ

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1966

518
М 69
УДК 519.3

АННОТАЦИЯ

В настоящее время основные методы решения задач математической физики разделяются на два класса — вариационные и сеточные. Предлагаемая книга содержит результаты, полученные автором и его сотрудниками и относящиеся к первой группе методов. В этой книге впервые точно поставлен и полностью решен вопрос об устойчивости вариационных методов и, в частности, важнейшего из них — метода Ритца. Книга является совершенно оригинальной и не имеет подобных в мировой литературе.

Книга рассчитана на практиков-вычислителей, инженеров, физиков и математиков, которым по роду их практической деятельности приходится сталкиваться с применением вариационных методов, а также на аспирантов и студентов старших курсов соответствующих специальностей.

Соломон Григорьевич Михлин

Численная реализация вариациончых методов

М., 1966 г., 432 стр. с илл.

Редактор *А. Ф. Лапко*

Техн. редактор *А. А. Благовещенская* Корректор *О. А. Бутусова*

Сдано в набор 3/XI 1965 г. Подписано к печати 25/I 1966 г. Бумага $60\times90^1/_{16}$. Физ. печ. л. 27.
Условн. печ. л. 27. Уч.-изд. л. 25,37.
Тираж 10500 экз. Т-01441. Цена книги 1 р. 80 к. Заказ № 2009.

Издательство «Наука»
Главная редакция физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Ленинградская типография № 2 имени Евгении Соколовой Главполиграфпрома
Комитета по печати при Совете Министров СССР,
Измайловский проспект, 29.

$\frac{2\text{-}2\text{-}4}{96\text{-}1966}$

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Эта книга является продолжением ранее вышедшей книги автора «Вариационные методы в математической физике» [12], в которой изложена общая теория вариационных методов, даны приложения к наиболее важным задачам математической физики, разработаны методы апостериорной оценки погрешности приближенного решения. В книге [12] приведены численные примеры, показывающие, что вариационные методы во многих случаях приводят к вполне удовлетворительным результатам. Некоторые более общие теоретические результаты, относящиеся к вариационным методам, изложены также в книге автора «Проблема минимума квадратичного функционала» [5].

Как кажется автору, проблемы численной реализации вариационных методов сводятся к следующим: 1) выбор системы координатных функций («координатной системы»); 2) составление системы Ритца; 3) решение системы Ритца; 4) учет влияния погрешностей, допущенных при составлении и решении системы Ритца, на точность приближенного решения.

В практически интересных случаях вторая из названных проблем сводится к вычислению интегралов от известных функций; этот вопрос хорошо исследован [1]), ему посвящено много места в монографиях и учебниках и мы его не касаемся. Третий вопрос — о решении систем Ритца — для линейных задач сводится к решению линейных алгебраических систем, что также хорошо изучено; из относящихся сюда многочисленных сочинений отметим монографии В. Н. Фаддеевой и Д. К. Фаддеева [1] и Дж. Х. Уилкинсона [1]. Для нелинейных задач этот вопрос изучен не столь обстоятельно. Мы не будем поэтому заниматься решением линейных систем Ритца, но уделим некоторое внимание этой проблеме для систем нелинейных.

Основными для настоящей книги являются четвертая и первая из названных выше проблем: учет влияния погрешностей и выбор координатной системы. Эти проблемы, которые мы здесь пытаемся решать применительно к линейным задачам, оказались выдвинутыми в связи с появлением электронных цифровых вычислительных машин.

---

[1]) За исключением случая осциллирующих функций, случая, не лишенного интереса для вариационных методов.

Широкое внедрение ЭЦВМ, имевшее столь большое влияние на всю вычислительную математику, не могло оставить в стороне и тот ее раздел, который связан с вариационными методами, и вызвало появление новых вопросов, ранее не привлекавших к себе достаточного внимания. Электронные машины дали возможность строить значительно более точные приближения, с использованием не трех-четырех, а многих десятков координатных функций. При этом сразу же выявились серьезные трудности, определяемые тем, что процесс счета по методу Ритца мог оказаться неустойчивым при увеличении числа координатных функций. Отметим тут же, что близкие явления отмечались и ранее при вычислениях, связанных с использованием большого числа координатных функций. Так, пропадание знаков при решении системы Ритца высокого порядка было обнаружено Л. М. Качановым [2], который применил процесс Ритца для исследования напряжений в турбинной лопасти. Исследование явлений устойчивости и неустойчивости процесса Ритца и тесно связанный с этим вопрос о рациональном выборе координатных элементов занимает важное место в настоящей книге.

В числе других здесь подробно рассмотрены задачи математической физики для бесконечных областей и вопросы, возникающие при решении интегральных уравнений, в частности, сингулярных; в упомянутых выше книгах автора [5] и [12] этим задачам было уделено сравнительно немного места.

Специальная глава V посвящена рациональному выбору координатных функций для ряда наиболее важных конкретных классов задач математической физики. Эта глава играет роль своеобразного справочника, хотя все утверждения в ней сопровождаются доказательствами или достаточными указаниями на то, как эти доказательства провести.

Одна из глав книги посвящена устойчивости процесса Ритца при вычислении собственных чисел и собственных элементов самосопряженного оператора. В другой главе исследуется влияние ошибки, допущенной при составлении уравнения, на его решение. Сформулированные в абстрактных терминах результаты конкретизируются для ряда задач математической физики и, особенно, для теории оболочек.

Особое место занимают главы IX и X, посвященные нелинейным задачам. В главе IX излагаются некоторые вопросы общей теории вариационных методов для нелинейных задач: сведение к вариационной задаче, сходимость минимизирующей последовательности, ее построение с помощью процесса Ритца. В главе X рассматриваются методы численного решения систем Ритца; эти методы применяются также и к нелинейным сеточным системам.

Часть материала настоящей книги конспективно изложена в обзорной статье автора [21] и в книге автора и Х. Л. Смолицкого [1].

В вариационных методах наиболее трудоемкой частью является составление системы Ритца (или Бубнова — Галёркина), требующее часто выполнения большого числа простых операций над элементарными функциями: дело идет о сложении, умножении, дифференцировании, неопределенном интегрировании, подстановке численных значений аргумента и т. п. В «Приложении», написанном Т. Н. Смирновой [1], [2], приведен ряд весьма интересных соображений о выполнении такого рода операций на электронных машинах.

Автор сделал попытку несколько упорядочить терминологию вариационных методов. В литературе термин «метод» принято относить к двум разным вещам: к приемам замены краевой задачи для данного уравнения вариационной задачей (энергетический метод, метод наименьших квадратов и т. д.) и к приемам приближенного решения вариационной задачи (метод Ритца, метод Бубнова — Галёркина). По существу, не приходится возражать против применения термина «метод» как в том, так и в другом случае, однако это создает известные неудобства. Мы будем поэтому называть «методами» только приемы приведения данной краевой задачи к задаче вариационной; приемы же приближенного решения вариационных задач (или самих краевых задач) будем называть «процессами» и будем в соответствии с этим говорить о процессе Ритца, процессе Бубнова — Галёркина, процессе моментов и т. д. Мы сохраним, однако, название «метод сеток», которое прочно привилось в научном обиходе.

Ниже принята следующая система нумерации теорем и формул. Каждая теорема (а также следствие, формула и т. д.) нумеруются двумя числами: первое из них — номер параграфа, второе — номер теоремы (следствия, формулы и т. д.) внутри данного параграфа.

В последующем часто будут встречаться ссылки на упомянутые выше книги автора [5] и [12]; для упрощения ссылок мы будем обозначать их начальными буквами «ПМ» и «ВМ» соответственно.

В настоящей книге отражены результаты не только автора, но и ряда других лиц, принявших участие в разработке поставленного автором вопроса об устойчивости вычислительных процессов: Г. М. Вайникко, М. А. Велиева, Ю. С. Вержбинской, Л. Н. Довбыш (Гаген-Торн), Б. А. Самокиша, М. Н. Яковлева, Г. Н. Ясковой. Важное место в книге занимают результаты И. В. Гельмана, А. Лангенбаха и Л. М. Качанова, относящиеся к нелинейным задачам, а также относящиеся к тем же задачам результаты автора, З. А. Власовой и Л. Н. Довбыш (Гаген-Торн).

Вычисления были в основном выполнены Н. А. Соловьевой и Т. А. Тушкиной; автор рад выразить им обеим сердечную благодарность. М. К. Гавурин прочитал книгу в рукописи и сделал по ней ряд весьма ценных замечаний. Автор выражает М. К. Гавурину свою искреннюю признательность.

Май 1965 г. *С. Михлин*

ВВЕДЕНИЕ

Основные средства для решения трудных прикладных задач — вариационный метод и метод сеток — пережили своеобразную эволюцию. Найденный В. Ритцем в 1908 г. процесс для построения минимизирующей последовательности дал могучий толчок прикладному анализу. Труднейшие задачи теории уравнений с частными производными, такие, как задачи теории упругости, теперь легко поддавались решению; многие представители прикладных наук, особенно С. П. Тимошенко, дали ряд важных приложений вариационного метода. Большую роль сыграли работы И. Г. Бубнова и Б. Г. Галёркина: разработанный ими процесс во многих случаях упрощает составление уравнений Ритца; кроме того, этот процесс часто оказывался пригодным и тогда, когда данная задача не допускала сведения к вариационной. Сравнительно скоро, однако, выявились и серьезные недостатки вариационного метода: трудность построения координатных функций, удовлетворяющих заданным граничным условиям, при сколько-нибудь сложной форме области, и трудоемкость составления системы Ритца, связанная с тем, что коэффициенты этой системы и ее свободные члены обычно выражаются через интегралы, вычисление которых, особенно при двух или большем числе независимых переменных, может потребовать большой затраты труда. Это последнее обстоятельство не имело, правда, серьезного значения, пока строились сравнительно грубые приближения, представляющие собой линейную комбинацию небольшого числа координатных функций.

Параллельно с вариационным методом развивался и метод сеток. На первых порах теоретическое значение этого метода было, пожалуй, больше, чем прикладное: при ручном счете решение систем достаточно высоких порядков, к чему приводит метод сеток, часто оказывается невыполненным.

Появление ЭЦВМ в корне изменило положение. Выявились важные преимущества метода сеток по сравнению с вариационным: для метода сеток менее важен вид области, и составление сеточной системы связано с выполнением простых и однообразных действий. В настоящее время метод сеток используется для вычисления значительно чаще, чем вариационный.

Следует в то же время признать, что основной недостаток метода сеток — трудоемкость решения систем высоких порядков — продолжает сказываться и в настоящее время. В книге В. Вазова и Дж. Форсайта [1], написанной в 1959 г., приведены примерные данные о количестве машинного времени, потребного для решения обычных задач математической физики методом сеток. Предполагая, что выполнение одного арифметического действия требует 50 микросекунд машинного времени, упомянутые авторы приходят к следующему выводу: чтобы получить решение с большой точностью, для одномерной задачи потребуется 1 час машинного времени, для двумерной — 6 недель, для трехмерной — 100 лет машинного времени. Конечно, с появлением более быстродействующих машин положение улучшится, тем не менее кажется вероятным, что для многомерных (начиная с трехмерных) задач применение метода сеток практически будет затруднено еще в течение длительного времени.

Вариационный метод требует значительно меньшего числа действий и соответственно значительно меньшей затраты машинного времени. По-видимому, следует считать, что к обширному классу задач можно применять как вариационный, так и сеточный метод, но в то же время каждый из этих методов имеет свою «область наивыгоднейшего применения». Для метода сеток это двумерные задачи для областей более сложной формы; для вариационного метода это в основном многомерные задачи, а также, возможно, задачи с различного рода вырождениями, когда разностная аппроксимация производных может привести к большим погрешностям. Возможно, что вариационный метод имеет некоторые преимущества и в случае бесконечных областей.

Разумеется, такое разделение «сфер влияния» обоих методов не является строгим. Вариационный метод удается применять и в случае областей сравнительно сложной формы; в то же время, по подсчетам, приведенным в цитированной выше книге В. Вазова и Дж. Форсайта, решение трехмерной задачи по методу сеток со средней точностью требует одной недели машинного времени. Это технически доступно и в случае необходимости может быть выполнено. К сказанному следует добавить, что «сфера влияния» каждого метода естественным образом расширяется по мере его разработки.

Однако применение вариационного метода для получения достаточно точных приближений требует предварительного выяснения ряда обстоятельств. Из результатов книги ВМ очевидно вытекает, что для получения более точного приближения по Ритцу следует увеличивать число используемых координатных функций. При этом возрастает порядок системы Ритца, но ее матрица и свободные члены неизбежно вычисляются с некоторыми, пусть малыми, погрешностями, поэтому при высоком порядке системы погрешность ее решения может оказаться весьма значительной; более того, эта погрешность может бесконечно возрастать вместе с порядком системы, как бы ни были малы погрешности при составлении самой системы. Практика вычислений показывает, что такие явления на самом деле могут иметь место. Мы уже упоминали в этой связи работу [2] Л. М. Качанова; назовем еще работу В. М. Миткевича [1]. В связи с этим можно говорить об «устойчивости» или «неустойчивости» процесса Ритца (точные формулировки даны ниже, в гл. II). В работах автора [13], [15], [17] выяснено, что устойчивость или неустойчивость процесса Ритца определяется некоторыми свойствами системы координатных элементов; исследования автора были продолжены в работах Г. М. Вайникко [1], Г. Н. Ясковой и М. Н. Яковлева [1], которые, в частности, рассмотрели вопрос об устойчивости метода Бубнова — Галёркина для стационарных задач, и М. А. Велиева [1] — [2], изучившего тот же вопрос для задач нестационарных. Укажем еще работу Л. Н. Довбыш [2], которая исследовала устойчивость процесса Ритца в спектральных задачах.

Общая точка зрения на устойчивость вычислительных процессов дана в статье автора [22], из результатов которой как частные случаи получаются перечисленные выше результаты.

Упомянутые выше работы автора, вместе с одной из его более ранних работ [11] и с работой П. Е. Соболевского [1], позволили высказать ряд соображений о рациональном выборе системы координатных элементов. Этому посвящена статья автора [18].

Второе, весьма важное обстоятельство заключается в следующем. Элементы матрицы и столбца свободных членов системы Ритца обычно выражаются через некоторые интегралы. На их вычисление и затрачивается основная масса труда. Часто эти интегралы вычисляются элементарно. В таких случаях затрата труда резко уменьшается, но вычисление интегралов вручную все же очень трудоемко. Представляется крайне желательной автоматизация таких аналитических выкладок. Первые шаги в этом направлении уже сделаны. Так, в статьях Т. Н. Смирновой [1] — [2] разработан прием проведения на электронных вычислительных машинах ряда элементарных операций над полиномами — алгебраическими и тригонометрическими. В число таких операций входят арифметические действия, дифференцирование, неопределенное и определенное интегрирование, подстановка значений аргумента и др.; для алгебраических полиномов с двумя независимыми переменными в число элементарных операций входят и такие, как вычисление двойного интеграла по области прямоугольника, эллипса или половины эллипса.

Работа Т. Н. Смирновой является одной из большого цикла работ, выполненных по идейному замыслу и под руководством Л. В. Канторовича и посвященных машинному проведению аналитических операций над объектами различной природы; ряд статей в этом направлении опубликован в сборнике «Труды Матем. ин-та им. В. А. Стеклова», вып. 66 (1962).

Если сравнить по содержанию настоящую книгу (в той ее части, которая относится к линейным задачам) с книгой ВМ, то можно отметить следующее. В ВМ рассматриваются вопросы построения и сходимости приближенного решения. Для этих вопросов выбор координатной системы безразличен: при любом ее выборе имеет место сходимость приближенного решения к точному в энергетической метрике. Более того, приближенное решение зависит не от самих координатных элементов, а только от натянутого на них подпространства. Если, например, решать вариационным методом задачу

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)+q(x)u=f(x),$$
$$u(-1)=u(1)=0,$$
$$p(x)\geqslant p_0=\text{const}>0, \qquad q(x)\geqslant 0,$$

то две координатные системы

$$\varphi_n(x)=(1-x)(1+x)^n \qquad (n=1,\ 2,\ 3,\ \ldots)$$

и (P_n — полиномы Лежандра)

$$\psi_n(x)=\sqrt{2n+1}\int\limits_{-1}^{x}P_n(t)\,dt \qquad (n=1,\ 2,\ 3,\ \ldots)$$

приводят к одинаковым приближенным решениям; с точки зрения проблем, изучаемых в ВМ, обе координатные системы совершенно равноценны.

В данной книге изучаются вопросы устойчивости приближенного решения, для которых свойства выбранной координатной системы (а не подпространств, натянутых на координатные элементы) играют решающую роль. Так, в указанном выше примере система $\{\varphi_n(x)\}$ приводит к неустойчивым приближениям, а система $\{\psi_n(x)\}$ — к устойчивым; если для упрощения выкладок отбросить у функции $\psi_n(x)$ множитель $\sqrt{2n+1}$, то приближенные решения потеряют устойчивость.

Сказанное выше относится к линейным задачам и, следовательно, к связанным с ними задачам о минимуме квадратичных функционалов. В области нелинейных задач сложилось иное положение. Почти нет работ, трактующих метод сеток для нелинейных задач; по-видимому, это следует объяснить трудностью решения нелинейных систем уравнений с большим числом неизвестных. Зато число работ, посвященных конкретным применениям вариационного метода к нелинейным задачам, довольно велико. Автор далек от мысли дать сколько-нибудь исчерпывающий обзор соответствующей литературы и хотел бы ограничиться только указанием основных направлений.

Прежде всего следует отметить заложенное С. Н. Бернштейном направление, связанное с доказательством теорем существования задач вариационного исчисления. Мы не будем останавливаться подробнее на этом направлении, далеком от предмета настоящей книги. Ряд работ посвящен построению вариационных принципов для нелинейных задач теории упругости и теории пластичности; эти работы изложены в книгах А. С. Вольмира [1], Л. М. Качанова [3], [5], В. В. Новожилова [2], Р. Хилла [1]. Упомянем еще работы Э. Райснера [1] и К. Э. Галимова [1] — [4], в которых даны новые формы вариационных принципов для нелинейных задач теории упругости, а также работу Д. Дракера [1], в которой сделана попытка систематически изложить все вариационные принципы теории пластичности на основе единого постулата.

Еще одно направление связано с применением вариационных методов для доказательства существования и исследования решений нелинейных задач теории упругости и теории пластичности. В этой связи отметим работы И. И. Воровича [1] — [6] по нелинейной теории оболочек. Используя то обстоятельство, что главные члены уравнений этой теории линейны, И. И. Ворович сводит свою задачу к решению системы уравнений, содержащих нелинейные вполне непрерывные операторы. Применяя топологические методы, И. И. Ворович доказывает компактность множества приближенных решений по Ритцу. Далее доказывается, что каждая предельная точка этого множества есть решение задачи. Ряд теорем, доказанных в цитированных статьях И. И. Воровича, позволяет подойти к оценке погрешности приближенного решения.

Нелинейной теории пластин посвящены работы Н. Ф. Морозова [1] — [5]. Этот автор подметил интересное интегральное тождество, которому удовлетворяют решения уравнения Т. Кармана при условиях жесткого закрепления или свободного опирания пластинки; использование упомянутого тождества сильно упростило исследование и позволило получить теорему существования. В статье [5] Н. Ф. Морозов доказывает, что при известных условиях существуют несимметричные решения задачи об изгибе круглой симметрично загруженной пластины, край которой жестко закреплен. Схема рассуждений такова. Фактически строится симметричное решение; доказывается, что такое решение единственно. Далее строится допустимое смещение, для которого потенциальная энергия изогнутой пластины меньше, чем для симметричного решения. Отсюда следует, что абсолютный минимум достигается на несимметричном решении.

А. Лангенбах [1] — [3] рассмотрел некоторый класс задач о минимуме, содержащий многие вариационные задачи теории пластичности. Функционалы, рассмотренные А. Лангенбахом, имеют вид:

$$F(u)=\int_{\Omega}\left\{\sum_{j=1}^{k}\int_{0}^{\tau_j(u)} g_j(\xi)\,d\xi\right\}dx-\int_{\Omega} f(x)\,u(x)\,dx. \tag{0.1}$$

Здесь Ω — некоторая конечная область m-мерного евклидова пространства, x — переменная точка этой области; g_j — неотрицательные функции от ξ, $\tau_j(u)$ — неотрицательные квадратичные формы от функции u и ее производных до некоторого порядка. Подробнее об условиях, которым подчинены величины, входящие в функционал F, будет сказано в главе IX. При этих условиях для функционала $F(u)$ устанавливается существование обобщенного решения вариационной задачи и сходимость, в соответствующей метрике, процесса Ритца.

Последнее направление, о котором мы хотели бы здесь упомянуть, связано с процессом Ритца и его численной реализацией. Этому направлению полностью посвящена глава X настоящей книги, поэтому здесь мы ограничимся самыми общими указаниями. Как уже было упомянуто выше, сходимость процесса Ритца устанавливалась для отдельных классов нелинейных задач, например, в работах И. И. Воровича и А. Лангенбаха. Результат общего характера приведен в книге И. М. Гельфанда и С. В. Фомина [1], где доказано, что процесс Ритца приводит к минимизирующей последовательности для возрастающего непрерывного функционала. В статье автора [19] выяснено, что требование непрерывности функционала достаточно заменить требованием его полунепрерывности сверху.

Проблема численной реализации вариационных методов для нелинейных задач разработана еще далеко недостаточно. Вопрос об устойчивости процесса Ритца для нелинейных задач далеко еще не разрешен.

Несколько подробнее скажем о более простом вопросе — о приближенном решении нелинейных систем Ритца; они получаются как необходимые условия минимума некоторой функции многих независимых переменных. В последнее время появился ряд статей, в которых разрабатываются как приемы решения систем нелинейных алгебраических или трансцендентных уравнений, так и приемы непосредственной минимизации функций; обзор многих работ этого рода дан в статье Х. А. Спэнга [1].

В главе X мы останавливаемся на трех методах решения нелинейных систем Ритца. Первый из них — так называемый «метод Ньютона», разработанный Л. В. Канторовичем; мы будем пользоваться термином «метод Ньютона — Канторовича».

Другой метод решения основан на идее, высказанной В. С. Кирия [1] и Д. Ф. Давиденко [1]; идея эта заключается в том, что вводится вспомогательный параметр t и задача Коши для системы обыкновенных дифференциальных уравнений с независимой переменной t. В конкретных случаях реализация идеи В. С. Кирия и Д. Ф. Давиденко упирается в ту трудность, что задачу Коши надо решать на фиксированном промежутке $0 \leqslant t \leqslant 1$. В статье Л. Н. Довбыш (Гаген-Торн) и автора [1] показано, что упомянутая задача Коши разрешима на промежутке $0 \leqslant t \leqslant 1$, если исходная система уравнений была системой Ритца для функционала, удовлетворяющего некоторым условиям, которые, грубо говоря, сводятся к тому, что этот функционал имеет полиномиальный рост на бесконечности. В этой же работе установлена тождественность методов Ритца и Бубнова — Галёркина для нелинейных задач.

Л. Н. Довбыш (Гаген-Торн) показала в статье [1], что функционалы вида (0.1) при естественных ограничениях на функции и формы $\tau_j(u)$ удовлетворяют упомянутому условию.

В статье З. А. Власовой [1] доказана применимость метода сведения к задаче Коши и для некоторых классов нелинейных сеточных систем.

Третий метод предложен Л. М. Качановым [4], [6] для задач нелинейной теории упругости и пластичности; этот метод представляет собой некоторый вариант метода последовательных приближений; при этом методе на каждом шагу приходится решать линейную систему с тем же числом неизвестных, что и у данной нелинейной системы.

В заключение отметим еще одно направление в развитии вариационного метода для нелинейных задач. В работах этого последнего направления вариационным методом (который часто несколько неточно именуется методом Галёркина или Бубнова — Галёркина, иногда методом Бубнова) строятся конкретные приближения к решениям конкретных нелинейных задач. Чаще всего приближенное решение строится в виде линейного агрегата двух и даже одной координатной функции. Укажем некоторые из последних работ этого направления[1]).

Задачи нелинейной теории пологих оболочек решаются в статье М. С. Корнишина и Х. М. Муштари [1]. Задача упруго-пластического кручения стержней — в статье А. М. Горлова [1]. Задачам нелинейной теории пластин посвящены работы М. Стиппса [1], Л. И. Календерьяна [1], Х. М. Бергера [1], Н. А. Вайля и Н. М. Ньюмарка [1], В. А. Постнова [1], С. Г. Винокурова [1], В. Д. Клюшникова [1], К. Э. Галимова [5]. Работа Вайля и Ньюмарка интересна тем, что в ней используется сравнительно большое число, а именно 11 координатных функций.

Во всей книге рассматриваются только сепарабельные пространства; этого вполне достаточно для прикладных задач математической физики.

---

[1]) Перечисленные ниже работы цитируются по реф. журн. «Механика».

ГЛАВА I

НЕКОТОРЫЕ КЛАССЫ СИСТЕМ ЭЛЕМЕНТОВ В ГИЛЬБЕРТОВОМ ПРОСТРАНСТВЕ

Настоящая глава является вспомогательной. Ее результаты в подавляющем большинстве известны из литературы; многое вошло в специальные монографии. Нам кажется, однако, удобным собрать вместе те понятия и факты, которые в дальнейшем будут широко использованы; как нам кажется, это облегчит пользование книгой.

§ 1. Минимальные системы[1])

Определение 1.1. Множество элементов гильбертова пространства называется *минимальной в этом пространстве системой*[2]), если вычеркивание любого элемента этого множества сужает натянутое на него подпространство.

Приведем некоторые примеры.

Любое конечное множество линейно независимых элементов минимально: вычеркивание любого элемента такого множества уменьшает на единицу размерность подпространства, натянутого на это множество. Наоборот, линейно зависимые элементы образуют неминимальную систему, так как вычеркивание элемента, линейно зависимого с остальными, не изменяет подпространства, натянутого на множество.

Любая (конечная или бесконечная) система элементов гильбертова пространства, ортогональных и отличных от нулевого, минимальна. Так, например, система функций $\sin k\pi x$ $(k = 1, 2, \ldots)$ минимальна в пространстве $L_2(0, \pi)$.

В том же пространстве система

$$x, \quad \sin \pi x, \quad \sin 2\pi x, \ldots, \sin k\pi x, \ldots$$

---

[1]) См. С. Левин [1], а также С. Качмаж и Г. Штейнгауз [1].

[2]) Определение и основные свойства минимальной системы сразу распространяются на любое банахово пространство. См. С. Качмаж и Г. Штейнгауз [1].

неминимальна: если вычеркнуть элемент x, то пространство, натянутое на остальные элементы, совпадает с самим пространством $L_2(0, \pi)$ и, следовательно, не претерпело сужения.

Приведем важный пример неминимальной системы. В силу известной теоремы Г. Мюнца[1]) последовательность $\{x^{p_k}\}$ $(k=1, 2, \ldots)$, $0 \leqslant p_1 < p_2 < p_3 \ldots$, полна в $L_2[0, 1]$, если показатели $p_k \geqslant 0$ и ряд

$$\sum_{k=2}^{\infty} \frac{1}{p_k}$$

расходится. Но тогда упомянутая последовательность неминимальна в $L_2(0, 1)$: вычеркнув любой ее член, мы не нарушим условий теоремы Г. Мюнца и, следовательно, опять получим полную в $L_2(0, 1)$ систему. Таким образом, подпространство, натянутое на нашу последовательность, не сужается, если вычеркнуть какой-нибудь ее член; по определению, эта последовательность в $L_2(0, 1)$ неминимальна.

В частности, в $L_2[0, 1]$ неминимальна последовательность $\{x^n\}$ $(n=0, 1, 2, \ldots)$.

Нам будет полезно другое определение минимальной системы, равносильное определению, данному выше. Для простоты ограничимся случаем конечной или счетной системы — только этот случай интересен для целей настоящей книги.

Определение 1.2. Пусть

$$u_1, u_2, \ldots, u_n, \ldots \tag{1.1}$$

— элементы гильбертова пространства $\mathfrak{H}$. Обозначим через $\mathfrak{H}_k$ подпространство, натянутое на элементы

$$u_1, u_2, \ldots, u_{k-1}, u_{k+1}, \ldots, u_n, \ldots$$

Система (1.1) *минимальна*, если, каково бы ни было k, элемент u_k не принадлежит подпространству $\mathfrak{H}_k$, и *неминимальна* в противном случае.

Следующая теорема непосредственно вытекает из этого определения.

Теорема 1.1. *Для того чтобы система* (1.1) *была неминимальной, необходимо и достаточно существование номера j, удовлетворяющего такому условию: каково бы ни было число $\varepsilon > 0$, можно найти число N и постоянные $\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_N$ такие, что имеет место неравенство*

$$\left\| u_j - \sum_{k=1,\, k \neq j}^{N} \alpha_k u_k \right\|_{\mathfrak{H}} < \varepsilon. \tag{1.2}$$

Определение 1.3. Система

$$v_1, v_2, \ldots, v_n, \ldots \tag{1.3}$$

---

[1]) См., например, С. Качмаж и Г. Штейнгауз [1].

элементов пространства $\mathfrak{H}$ называется *биортонормированной* с системой (1.1), если

$$(u_j, v_k)_{\mathfrak{H}} = \delta_{jk} = \begin{cases} 0, & j \neq k, \\ 1, & j = k. \end{cases} \tag{1.4}$$

Теорема 1.2. *Для того чтобы система* (1.1) *была минимальна, необходимо и достаточно, чтобы существовала биортонормированная с ней система. Эта последняя определяется единственным образом, если потребовать, чтобы ее элементы принадлежали подпространству, натянутому на систему* (1.1).

Необходимость. Пусть система (1.1) минимальна. Подпространство, указанное в формулировке теоремы, обозначим через $\overline{\mathfrak{H}}$. В системе (1.1) выделим элемент u_j и на оставшиеся элементы натянем подпространство $\mathfrak{H}_j$. Положим $u_j = \xi + \eta$, где $\eta \in \mathfrak{H}_j$, а ξ ортогонально к $\mathfrak{H}_j$; так как система (1.1), по предположению, минимальна, то $u_j \,\overline{\in}\, \mathfrak{H}_j$, и следовательно, $\xi \neq 0$. Любой элемент u пространства $\overline{\mathfrak{H}}$ можно представить в виде $u = \lambda\xi + \zeta$, $\zeta \in \mathfrak{H}_j$. Зададим в $\overline{\mathfrak{H}}$ линейный функционал l_j, положив $l_j(u) = \lambda$. Этот функционал ограничен, так как

$$|l_j(u)| = |\lambda| \leqslant \frac{\|u\|_{\mathfrak{H}}}{\|\xi\|_{\mathfrak{H}}}.$$

По теореме Ф. Риса, в пространстве $\overline{\mathfrak{H}}$ существует единственный элемент v_j такой, что $l_j(u) = (u, v_j)_{\mathfrak{H}}$. Значение λ равно единице для элемента u_j и нулю для элементов u_k, $k \neq j$, поэтому

$$(u_k, v_j) = l_j(u_k) = \begin{cases} 0, & k \neq j, \\ 1, & k = j. \end{cases}$$

Достаточность[1]). Пусть существует система (1.2), биортонормированная с системой (1.1), и пусть эта последняя неминимальна. Зададим достаточно малое положительное число ε и подберем числа N, $\alpha_1, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_N$ так, чтобы выполнялось неравенство (1.2). Тогда

$$\left| \left(u_j - \sum_{k=1,\, k \neq j}^{N} \alpha_k u_k,\ v_j \right)_{\mathfrak{H}} \right| \leqslant \varepsilon \| v_j \|_{\mathfrak{H}}.$$

С другой стороны, из равенства (1.4) вытекает, что

$$\left(u_j - \sum_{k=1,\, k \neq j}^{N} \alpha_k u_k,\ v_j \right)_{\mathfrak{H}} = 1.$$

Последние два соотношения несовместны при достаточно малом ε.

---

[1]) Ср. С. Качмаж и Г. Штейнгауз [1].

§ 2. Сильно минимальные и почти ортонормированные системы

В этом параграфе речь будет идти только о счетных системах элементов.

Рассмотрим счетную систему элементов

$$u_1, \ u_2, \ \ldots, \ u_n, \ \ldots, \tag{2.1}$$

принадлежащих данному гильбертову пространству $\mathfrak{H}$, и матрицу Грама первых n элементов системы (1.1):

$$R_n = \begin{pmatrix} (u_1, u_1)_{\mathfrak{H}} & (u_1, u_2)_{\mathfrak{H}} & \ldots & (u_1, u_n)_{\mathfrak{H}} \\ (u_2, u_1)_{\mathfrak{H}} & (u_2, u_2)_{\mathfrak{H}} & \ldots & (u_2, u_n)_{\mathfrak{H}} \\ \cdot & \cdot & \cdot & \cdot \\ (u_n, u_1)_{\mathfrak{H}} & (u_n, u_2)_{\mathfrak{H}} & \ldots & (u_n, u_n)_{\mathfrak{H}} \end{pmatrix}. \tag{2.2}$$

Эта матрица эрмитова и неотрицательна, поэтому ее собственные числа неотрицательны. Запишем их в порядке возрастания

$$0 \leqslant \lambda_1^{(n)} \leqslant \lambda_2^{(n)} \leqslant \ldots \leqslant \lambda_n^{(n)}. \tag{2.3}$$

Число $\lambda_m^{(n)}$ можно найти как минимум отношения

$$\frac{\sum\limits_{j,\,k=1}^{n} (u_j, u_k)_{\mathfrak{H}} t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2} = \frac{\left\| \sum\limits_{k=1}^{n} t_k u_k \right\|_{\mathfrak{H}}^2}{\sum\limits_{k=1}^{n} |t_k|^2} \tag{2.4}$$

на подпространстве векторов $(t_1, t_2, \ldots, t_n)$, ортогональных к собственным векторам, принадлежащих собственным числам $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots$ $\ldots, \lambda_{m-1}^{(n)}$; в частности,

$$\lambda_1^{(n)} = \min \frac{\left\| \sum\limits_{k=1}^{n} t_k u_k \right\|_{\mathfrak{H}}^2}{\sum\limits_{k=1}^{n} |t_k|^2}; \tag{2.5}$$

минимум берется по всевозможным векторам $(t_1, t_2, \ldots, t_n)$, отличным от нулевого.

Важно отметить, что, как это непосредственно следует из минимаксимального принципа (ВМ, § 40), при фиксированном m и возрастающем n число $\lambda_m^{(n)}$ не возрастает.

Система (2.1) называется *сильно минимальной* [1]) в пространстве $\mathfrak{H}$, если

$$\inf \lambda_1^{(n)} = \lim_{n\to\infty} \lambda_1^{(n)} > 0. \tag{2.6}$$

При этом, очевидно, существует такая положительная постоянная λ_0, что при любом n выполняется неравенство

$$\lambda_1^{(n)} \geqslant \lambda_0. \tag{2.7}$$

Простейший пример сильно минимальной системы дает любая ортонормированная система — для нее $\lambda_k^{(n)} = 1$ при любых n и k, и можно положить $\lambda_0 = 1$.

Теорема 2.1. *Всякая сильно минимальная в некотором пространстве система минимальна в том же пространстве.*

Пусть система (2.1) сильно минимальна в гильбертовом пространстве $\mathfrak{H}$ и пусть она в этом же пространстве неминимальна. Задав произвольное число $\varepsilon > 0$, можно, в силу теоремы 1.1, найти число N и постоянные $\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_N$, при которых для некоторого фиксированного номера j имеет место неравенство (1.2). Тем более верно неравенство

$$\frac{\left\| u_j - \sum\limits_{k=1,\, k\neq j}^{N} \alpha_k u_k \right\|^2}{1 + \sum\limits_{k=1,\, k\neq j}^{N} |\alpha_k|^2} < \varepsilon^2. \tag{2.8}$$

Отношение (2.8) есть частный случай отношения (2.5) при $t_j = 1$ и $t_k = -\alpha_k$, $k \neq j$; из соотношений (2.5) и (2.8) вытекает теперь, что $\lambda_1^{(N)} < \varepsilon^2$, и следовательно, $\inf \lambda_1^{(n)} = 0$; это противоречит неравенству (2.6), которое выполняется, так как система (2.1) сильно минимальна.

Теорема, обратная теореме 2.1, неверна: система может быть в некотором пространстве минимальной, но не сильно минимальной. Так, например, пусть система $\{\omega_n\}$ ортонормирована в $\mathfrak{H}$, тогда система $\left\{\frac{\omega_n}{n}\right\}$, очевидно, минимальна в $\mathfrak{H}$, но она в $\mathfrak{H}$ не сильно минимальна: нетрудно видеть, что в этом случае $\lambda_1^{(n)} = \frac{1}{n}$ и $\inf \lambda_1^{(n)} = 0$.

Сделаем следующее замечание. Если система (2.1) сильно минимальна в $\mathfrak{H}$, то

$$\sum_{k=1}^{n} |\sigma_{kj}^{(n)}|^2 \leqslant \lambda_0^{-2}, \tag{2.9}$$

---

[1]) Термин А. Т. Талдыкина [1]; в книге С. Качмажа и Г. Штейнгауза [1] (стр. 431) такие системы названы «бесселевыми».

где $\sigma_{kj}^{(n)}$ — элементы матрицы R_n^{-1}. Действительно, наибольшее собственное число этой матрицы равно $\frac{1}{\lambda_1^{(n)}}$, поэтому для любого вектора $t=(t_1, t_2, \ldots, t_n)$ справедливо неравенство

$$\| R_n^{-1} t \|^2 \leqslant [\lambda_1^{(n)}]^{-2} \| t \|^2 \leqslant \lambda_0^{-2} \| t \|^2.$$

Полагая здесь $t=(0, 0, \ldots, 0, 1, 0, \ldots, 0)$, где единица занимает j-е место, мы и получим неравенство (2.9).

Введем еще одно определение. Систему (2.1) назовем *почти ортонормированной*[1]) в пространстве $\mathfrak{H}$, если существуют такие постоянные λ_0 и Λ_0, что при любых n и $m \leqslant n$ выполняются неравенства

$$\lambda_0 \leqslant \lambda_m^{(n)} \leqslant \Lambda_0. \tag{2.10}$$

Любая ортонормированная система является и почти ортонормированной: в этом случае $\lambda_0=\Lambda_0=1$.

Всякая почти ортонормированная в некотором пространстве система сильно минимальна и, тем более, минимальна в том же пространстве.

§ 3. Сходные и полусходные операторы

Назовем два самосопряженных положительно определенных оператора *сходными*, если они имеют общую область определения, и *полусходными*, если их энергетические пространства[2]) состоят из одних и тех же элементов.

Будем обозначать через H гильбертово пространство, в котором действуют рассматриваемые операторы; скалярное произведение и норму в H будем обозначать обычными символами $(\ ,\ )$ и $\|\ \|$. Для обозначения энергетического произведения и энергетической нормы будем пользоваться той же символикой, что и в ВМ и ПМ. Если A — некоторый оператор, то его область определения будем обозначать через $D(A)$, область значений — через $R(A)$.

Э. Хайнц [1] доказал, что если положительные операторы A и B имеют общую область определения, так что $D(A)=D(B)$, то и $D(A^\alpha)=D(B^\alpha)$ при любом α из промежутка $0<\alpha<1$. С другой стороны, для положительно определенного оператора множество

[1]) В статье А. Т. Талдыкина [1] такие системы названы «нормальными», а в книге С. Качмажа и Г. Штейнгауза [1] (стр. 440) — «системами Риса — Фишера», а также «базисами Риса».

[2]) Здесь и ниже мы называем «*энергетическим*» пространство H_A, построенное по Фридрихсу, т. е. как замыкание области $D(A)$ определения положительно определенного оператора A в метрике $[u, v]_A=(Au, v)$; см. ВМ, § 46, или ПМ, § 3.

элементов, образующих его энергетическое пространство, совпадает[1]) с $D\left(A^{\frac{1}{2}}\right)$. Отсюда сразу вытекает, что сходные операторы одновременно и полусходны и что операторы A и B полусходны тогда и только тогда, когда операторы $A^{\frac{1}{2}}$ и $B^{\frac{1}{2}}$ сходные.

Теорема 3.1. *Если операторы A и B сходные, то операторы AB^{-1}, BA^{-1}, $A^{-1}B$, $B^{-1}A$ ограничены и, следовательно, существуют положительные постоянные c_1 и c_2 такие, что*

$$c_1 \|Bu\| \leqslant \|Au\| \leqslant c_2 \|Bu\|, \qquad u \in D(A). \tag{3.1}$$

Заметим прежде всего, что $D\left(AB^{-1}\right) = H$. Действительно, $D\left(B^{-1}\right) = R(B) = H$, так как оператор B — самосопряженный и положительно определенный[2]). С другой стороны, $R\left(B^{-1}\right) = D(B) = D(A)$, поэтому если f — произвольный элемент из H, то выражение $AB^{-1}f = A\left(B^{-1}f\right)$ вполне определено.

Докажем теперь, что оператор AB^{-1} замкнут. Пусть $f_n \to f$ и $AB^{-1}f_n \to g$. Положим $B^{-1}f_n = h_n$. Оператор B^{-1} ограничен[3]), поэтому $B^{-1}f_n \to B^{-1}f$. Положим $B^{-1}f = h$, тогда $h_n \to h$ и $Ah_n \to g$. Будучи самосопряженным, оператор A замкнут; отсюда следует, что $h \in D(A)$ и $Ah = g$, или $AB^{-1}f = g$, что и требовалось доказать. Раз оператор AB^{-1} определен на всем пространстве H и замкнут, то он ограничен[4]) в H.

Ограниченность оператора BA^{-1} доказывается аналогично, а операторы $A^{-1}B$ и $B^{-1}A$ ограничены как сопряженные к ограниченным операторам BA^{-1} и AB^{-1} соответственно.

Положим $\|AB^{-1}\| = c_2$, $\|BA^{-1}\| = \frac{1}{c_1}$. Тогда

$$\|AB^{-1}f\| \leqslant c_2 \|f\|, \qquad \|BA^{-1}g\| \leqslant \frac{1}{c_1} \|g\|,$$

где f и g — произвольные элементы пространства H. Положив $f = Bu$, $g = Au$, мы придем к неравенству (2.1).

Следствие 3.1. *Если операторы A и B — полусходные, то операторы*

$$A^{\frac{1}{2}}B^{-\frac{1}{2}}, \quad B^{\frac{1}{2}}A^{-\frac{1}{2}}, \quad A^{-\frac{1}{2}}B^{\frac{1}{2}}, \quad B^{-\frac{1}{2}}A^{\frac{1}{2}} \tag{3.2}$$

---

[1]) См. ПМ, § 5, стр. 24.
[2]) См. ПМ, § 5.
[3]) См. ПМ, § 5.
[4]) См. Л. В. Канторович и Г. П. Акилов [1], теорема 4 (1. XII), стр. 427.

ограничены в H; существуют положительные постоянные c_1 и c_2 такие, что

$$c_1 |u|_B \leqslant |u|_A \leqslant c_2 |u|_B. \tag{3.3}$$

Доказательство. Как было отмечено в начале параграфа, в данном случае операторы $A^{\frac{1}{2}}$ и $B^{\frac{1}{2}}$ — сходные и, следовательно, операторы (3.2) ограничены; для этих операторов неравенство (3.1) принимает вид:

$$C_1 \left\| B^{\frac{1}{2}} u \right\| \leqslant \left\| A^{\frac{1}{2}} u \right\| \leqslant C_2 \left\| B^{\frac{1}{2}} u \right\|. \tag{3.4}$$

Далее, для положительно определенного оператора A справедливо равенство $|u|_A = \left\| A^{\frac{1}{2}} u \right\|$. Действительно, если $u \in D(A)$, то

$$|u|_A = \sqrt{(Au, u)} = \sqrt{\left(A^{\frac{1}{2}} u,\ A^{\frac{1}{2}} u\right)} = \left\| A^{\frac{1}{2}} u \right\|;$$

в общем случае, когда $u \in H_A$, это равенство получается простым предельным переходом. Неравенство (3.3) оказывается теперь только иной записью неравенства (3.4).

Теорема 3.2. *Пусть операторы A и B — положительно определенные и $H_A \subset H_B$. Тогда существует такая постоянная $c > 0$, что*

$$|u|_A \geqslant c |u|_B, \quad u \in H_A. \tag{3.5}$$

По условию теоремы, $D\left(A^{\frac{1}{2}}\right) \subset D\left(B^{\frac{1}{2}}\right)$; повторяя рассуждения теоремы 3.1, найдем, что оператор $B^{\frac{1}{2}} A^{-\frac{1}{2}}$ ограничен. Пусть $\left\| B^{\frac{1}{2}} A^{-\frac{1}{2}} \right\| = \frac{1}{c}$, тогда $\left\| B^{\frac{1}{2}} A^{-\frac{1}{2}} f \right\| \leqslant \frac{1}{c} \|f\|$, $f \in H$. Полагая $A^{-\frac{1}{2}} f = u$, имеем $u \in D\left(A^{\frac{1}{2}}\right)$ или, что то же, $u \in H_A$ и $\left\| B^{\frac{1}{2}} u \right\| \leqslant \frac{1}{c} \left\| A^{\frac{1}{2}} u \right\|$, что равносильно неравенству (3.5).

§ 4. Теорема сравнения[1]

Пусть $\mathfrak{H}_1$ и $\mathfrak{H}_2$ — два гильбертовых пространства. Говорят, что пространство $\mathfrak{H}_1$ *вложено*[2] (или *вкладывается*) в $\mathfrak{H}_2$, если существует линейный оператор V («*оператор вложения*»), который

---

[1]) Результаты настоящего параграфа по существу содержатся в статьях автора [15], [17].

[2]) Подробно о вложении пространств (не только гильбертовых) см. в книге [1] С. Л. Соболева, который ввел это понятие, установил столь важные «теоремы вложения» и дал многие существенные их приложения. См. также В. И. Смирнов [3].

каждому элементу $u \in \mathfrak{H}_1$ приводит в соответствие некоторый элемент $Vu \in \mathfrak{H}_2$, и если этот оператор ограничен:

$$\| Vu \|_2 \leqslant K \| u \|_1; \tag{4.1}$$

символ $\| \ \|_k$ означает норму в пространстве $\mathfrak{H}_k$ $(k = 1, 2)$.

Если $\mathfrak{H}_1$ вкладывается в $\mathfrak{H}_2$, то будем писать $\mathfrak{H}_1 \subset \mathfrak{H}_2$.

Часто встречается тот случай, когда все элементы пространства $\mathfrak{H}_1$ принадлежат также и пространству $\mathfrak{H}_2$, а оператор V есть тождественный оператор. В этом случае неравенство (4.1), определяющее вложение пространств, принимает более простой вид:

$$\| u \|_2 \leqslant K \| u \|_1. \tag{4.1$'$}$$

Для упрощения формулировок мы будем в настоящем параграфе иметь в виду именно этот случай. Все утверждения параграфа остаются в силе и в общем случае, когда оператор вложения отличен от тождественного. Чтобы в этом убедиться, достаточно во всех рассуждениях настоящего параграфа, там, где элемент $\varphi_k \in \mathfrak{H}_1$ рассматривается как элемент пространства $\mathfrak{H}_2$, заменить φ_k на $V\varphi_k$.

Теорема 4.1. *Пусть пространство $\mathfrak{H}_1$ вложено в пространство $\mathfrak{H}_2$ и пусть все элементы системы*

$$\varphi_1, \ \varphi_2, \ \ldots, \ \varphi_n, \ \ldots \tag{4.2}$$

принадлежат пространству $\mathfrak{H}_1$. Если система (4.2) *минимальна в $\mathfrak{H}_2$, то она минимальна и в $\mathfrak{H}_1$.*

Допустим, что система (4.2) неминимальна в $\mathfrak{H}_1$. По теореме 1.1 найдется элемент φ_j, удовлетворяющий неравенству

$$\left\| \varphi_j - \sum_{k=1,\ k \neq j}^{N} \alpha_k \varphi_k \right\|_1 < \varepsilon, \tag{4.3}$$

в котором ε — любое положительное число, а натуральное число N и постоянные $\alpha_1, \ \ldots, \ \alpha_{j-1}, \ \alpha_{j+1}, \ \ldots, \ \alpha_N$ выбраны подходящим образом. В силу соотношения (4.1$'$) имеем:

$$\left\| \varphi_j - \sum_{k=1,\ k \neq j}^{N} \alpha_k \varphi_k \right\|_2 < K\varepsilon$$

и система (4.2) неминимальна в $\mathfrak{H}_2$, вопреки условию теоремы.

Теорема 4.2. *Если в условиях теоремы* 4.1 *система* (4.2) *сильно минимальна в $\mathfrak{H}_2$, то она сильно минимальна и в $\mathfrak{H}_1$.*

Составим матрицы

$$R_n = \begin{pmatrix} (\varphi_1, \varphi_1)_1 & (\varphi_1, \varphi_2)_1 & \ldots & (\varphi_1, \varphi_n)_1 \\ (\varphi_2, \varphi_1)_1 & (\varphi_2, \varphi_2)_1 & \ldots & (\varphi_2, \varphi_n)_1 \\ \cdot & \cdot & \cdot & \cdot \\ (\varphi_n, \varphi_1)_1 & (\varphi_n, \varphi_2)_1 & \ldots & (\varphi_n, \varphi_n)_1 \end{pmatrix} \tag{4.4}$$

и

$$r_n = \begin{pmatrix} (\varphi_1, \varphi_1)_2 & (\varphi_1, \varphi_2)_2 & \dots & (\varphi_1, \varphi_n)_2 \\ (\varphi_2, \varphi_1)_2 & (\varphi_2, \varphi_2)_2 & \dots & (\varphi_2, \varphi_n)_2 \\ \dots & \dots & \dots & \dots \\ (\varphi_n, \varphi_1)_2 & (\varphi_n, \varphi_2)_2 & \dots & (\varphi_n, \varphi_n)_2 \end{pmatrix}. \tag{4.5}$$

Собственные числа матриц (4.4) и (4.5), расположенные в порядке возрастания, обозначим через $\lambda_k^{(n)}$ и $\mu_k^{(n)}$ соответственно, так что

$$\lambda_1^{(n)} \leqslant \lambda_2^{(n)} \leqslant \dots \leqslant \lambda_n^{(n)}, \qquad \mu_1^{(n)} \leqslant \mu_2^{(n)} \leqslant \dots \leqslant \mu_n^{(n)}.$$

Система (4.2) сильно минимальна в $\mathfrak{H}_2$; по определению, существует такая независящая от n постоянная μ_0, что $\mu_1^{(n)} \geqslant \mu_0$. Оценим снизу число $\lambda_1^{(n)}$. Имеем:

$$\lambda_1^{(n)} = \inf \frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2} =$$

$$= \inf \frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k} \cdot \frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2} \geqslant$$

$$\geqslant \inf \frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k} \cdot \inf \frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2};$$

точная нижняя граница берется по множеству всевозможных ненулевых векторов $(t_1, t_2, \dots, t_n)$.

В последнем неравенстве второй множитель справа равен $\mu_1^{(n)}$, а первый оценивается по неравенству (4.1), именно:

$$\frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k} = \frac{\left\| \sum\limits_{k=1}^{n} t_k \varphi_k \right\|_1^2}{\left\| \sum\limits_{k=1}^{n} t_k \varphi_k \right\|_2^2} \geqslant \frac{1}{K^2}.$$

Отсюда

$$\inf \frac{\sum\limits_{j,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k} \geqslant \frac{1}{K^2}.$$

Теперь $\lambda_1^{(n)} \geqslant K^{-2}\mu_1^{(n)}$ и, следовательно,

$$\lambda_1^{(n)} \geqslant K^{-2}\mu_0. \tag{4.6}$$

Теорема доказана.

Теорема 4.3. *Если каждое из пространств $\mathfrak{H}_1$ и $\mathfrak{H}_2$ вкладывается в другое и система* (4.2) *почти ортонормирована в одном из них, то она почти ортонормирована и в другом.*

В данном случае существуют такие положительные постоянные K_1 и K_2, что

$$K_1 \|u\|_2 \leqslant \|u\|_1 \leqslant K_2 \|u\|_2. \tag{4.7}$$

Сохраним обозначения, принятые нами при доказательстве теоремы 4.2. Пусть система (4.2) почти ортонормирована в $\mathfrak{H}_2$, тогда она в этом пространстве и сильно минимальна. По теореме 4.2, система (4.2) сильно минимальна и в $\mathfrak{H}_1$ существует такая постоянная $\lambda_0 > 0$, что $\lambda_k^{(n)} \geqslant \lambda_1^{(n)} \geqslant \lambda_0$. Остается доказать, что собственные числа матрицы (4.4) ограничены сверху постоянной, не зависящей от n. Система (4.2) почти ортонормирована в $\mathfrak{H}_2$ и потому существует такая постоянная M_0, что $\mu_n^{(n)} \leqslant M_0$. Имеем:

$$\lambda_n^{(n)} = \sup \frac{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2} =$$

$$= \sup \frac{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k} \cdot \frac{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2} \leqslant$$

$$\leqslant \sup \frac{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_1 t_j \bar{t}_k}{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k} \cdot \sup \frac{\sum\limits_{j,\,k=1}^{n} (\varphi_j, \varphi_k)_2 t_j \bar{t}_k}{\sum\limits_{k=1}^{n} |t_k|^2};$$

точная верхняя граница берется по множеству всевозможных ненулевых векторов $(t_1, t_2, \ldots, t_n)$. Второй множитель справа равен $\mu_n^{(n)}$, а первый оценивается сверху, по неравенству (4.7), постоянной K_2^2. Теперь $\lambda_k^{(n)} \leqslant \lambda_n^{(n)} \leqslant K_2^2\mu_n^{(n)} \leqslant K_2^2 M_0 = \text{const}$.

Следствие 4.1. *Пусть A и B — положительно определенные операторы, действующие в гильбертовом пространстве H, и пусть все элементы пространства H_A принадлежат также пространству H_B. Если система* (4.2) *минимальна* *(сильно минимальна) в H_B, то она минимальна (сильно минимальна) и в H_A. В частности, если элементы системы* (4.2) *принадлежат H_A и эта система ортонормирована в H_B, то она сильно минимальна в H_A.*

Действительно, в силу теоремы 3.2 пространство H_A вкладывается в H_B, а тогда можно применить теоремы 4.1 и 4.2.

Следствие 4.2. *Полагая в следствии* 4.1 $B=I$, *где I — тождественный оператор, найдем, что если A — положительно определенный оператор, все элементы системы* (4.2) *принадлежат пространству H_A и эта система минимальна (сильно минимальна) в H, то она минимальна (сильно минимальна) и в H_A. В частности, если элементы системы* (4.2) *принадлежат H_A и эта система ортонормирована в H, то она сильно минимальна в H_A.*

Следствие 4.3. *Если A и B — полуисходные операторы и система* (4.2) *почти ортонормирована (в частности, ортонормирована) в одном из пространств H_A или H_B, то она почти ортонормирована во втором пространстве.*

Это следствие сразу вытекает из неравенства (3.3) и теоремы 4.3

§ 5. Некоторые свойства наилучшего приближения[1])

Задача о наилучшем приближении ставится так: дана последовательность

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots \tag{5.1}$$

элементов гильбертова пространства $\mathfrak{H}$, которое мы будем считать сепарабельным, и некоторый элемент u того же пространства; требуется найти постоянные $a_1, a_2, \ldots, a_n$ (число n фиксировано) так, чтобы

$$\left\| u - \sum_{k=1}^{n} a_k \varphi_k \right\| = \min.$$

Как легко видеть, постоянные a_k определяются из системы уравнений

$$\sum_{k=1}^{n} (\varphi_k, \varphi_j)\, a_k = (u, \varphi_j) \qquad (j = 1, 2, \ldots, n). \tag{5.2}$$

Постоянные a_k, удовлетворяющие системе (5.2), вообще говоря, зависят не только от k, но и от n, поэтому мы впредь будем обозначать их через $a_k^{(n)}$. Очевидно, выражение

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k \tag{5.3}$$

---

[1]) Теоремы настоящего параграфа получены в работах С. Левина [1] и А. Т. Талдыкина [1].

есть проекция элемента u на подпространство $\mathfrak{H}_n$, натянутое на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$.

Будем предполагать, что при любом n элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы и что последовательность (5.1) полна в $\mathfrak{H}$. Тогда система (5.2) разрешима при любом n и элемент (5.3) стремится к u при $n \to \infty$. Действительно, система (5.2) разрешима: ее определитель как определитель Грама линейно независимых элементов $\varphi_1, \varphi_2, \ldots, \varphi_n$ отличен от нуля. Далее, решение (5.3) задачи о наилучшем приближении не изменится, если последовательность (5.1) подвергнуть невырожденному линейному преобразованию с треугольной матрицей, так как при таком преобразовании подпространства $\mathfrak{H}_n$ не меняются. Имея это в виду, подвергнем систему (5.1) процессу ортогонализации, и пусть этот процесс приведет нас к ортонормированной системе $\{\omega_n\}$ $(n = 1, 2, \ldots)$, очевидно, полной в $\mathfrak{H}$. Если мы положим

$$u_n = \sum_{k=1}^{n} \alpha_k \omega_k,$$

то система (5.2) примет вид

$$\alpha_k = (u, \omega_k).$$

Отсюда

$$u_n = \sum_{k=1}^{n} (u, \omega_k)\, \omega_k.$$

Система $\{\omega_n\}$ ортонормирована и полна, поэтому

$$u = \sum_{k=1}^{\infty} (u, \omega_k)\, \omega_k$$

и, следовательно, $u_n \underset{n \to \infty}{\longrightarrow} u$.

Теорема 5.1. *Если система* (5.1) *минимальна, то существуют пределы*

$$a_k = \lim_{n \to \infty} a_k^{(n)} \qquad (k = 1, 2, \ldots). \tag{5.4}$$

Если, кроме того, система, биортонормированная к системе (5.1), *ограничена, то стремление к пределу в формуле* (5.4) — *равномерное относительно k.*

Пусть система $\{\psi_k\}$ $(k = 1, 2, \ldots)$ биортогональна к системе (5.1). Умножив скалярно обе части формулы (5.3) на ψ_j $(j = 1, 2, \ldots, n)$, находим:

$$a_j^{(n)} = (u_n, \psi_j);$$

отсюда видно, что существует предел

$$a_j = \lim_{n \to \infty} a_j^{(n)} = (u, \psi_j). \tag{5.5}$$

Если система $\{\psi_j\}$ ограничена, так что $\|\psi_j\| \leqslant C = \text{const}$, то

$$|a_j - a_j^{(n)}| = |(u - u_n, \psi_j)| \leqslant C \| u - u_n \|$$

и сходимость $a_j^{(n)} \to a_j$ равномерна относительно индекса j.

Теорема 5.2. *Если система* (5.1) *сильно минимальна в* $\mathfrak{H}$, *то последовательность*

$$a = (a_1, a_2, \ldots, a_n, \ldots),$$

где a_k *суть пределы* (5.4), *представляет собой элемент пространства* l_2; *если положить*

$$a^{(n)} = \left(a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}, 0, \ldots\right),$$

то $\lim\limits_{n\to\infty} \| a^{(n)} - a \|_{l_2} = 0$.

Рассмотрим матрицу Грама элементов $\varphi_1, \varphi_2, \ldots, \varphi_n$ и пусть $\lambda_1^{(n)}$ — наименьшее собственное число этой матрицы. Система (5.1) сильно минимальна, поэтому $\lambda_1^{(n)} \geqslant \lambda_0$, где λ_0 — некоторая положительная постоянная. По формуле (5.3) имеем:

$$\| u_n \|^2 = \sum_{j,\,k=1}^{n} (\varphi_j, \varphi_k)\, a_j^{(n)} \bar{a}_k^{(n)} \geqslant \lambda_1^{(n)} \sum_{k=1}^{n} | a_k^{(n)} |^2 \geqslant \lambda_0 \sum_{k=1}^{n} | a_k^{(n)} |^2.$$

Отсюда

$$\sum_{k=1}^{n} | a_k^{(n)} |^2 \leqslant \frac{1}{\lambda_0} \| u_n \|^2.$$

Далее, $\| u_n \| \leqslant \| u \|$, так как u_n есть проекция элемента u. Но тогда

$$\sum_{k=1}^{n} | a_k^{(n)} |^2 \leqslant \frac{1}{\lambda_0} \| u \|^2. \tag{5.6}$$

Тем более,

$$\sum_{k=1}^{p} | a_k^{(n)} |^2 \leqslant \frac{1}{\lambda_0} \| u \|^2, \qquad p \leqslant n.$$

Полагая в этой формуле $n \to \infty$, а затем $p \to \infty$, найдем, что

$$\sum_{k=1}^{\infty} | a_k |^2 \leqslant \frac{1}{\lambda_0} \| u \|^2$$

и, следовательно, $a \in l_2$.

Пусть теперь $m > n$ и $a_k^{(n)} = 0$, если $k < n$. Тогда

$$u_m - u_n = \sum_{k=1}^{m} \left(a_k^{(m)} - a_k^{(n)}\right) \varphi_k.$$

Повторив предшествующие рассуждения, придем к неравенству

$$\sum_{k=1}^{p} \left| a_k^{(m)} - a_k^{(n)} \right|^2 \leqslant \frac{1}{\lambda_0} \| u_m - u_n \|^2, \qquad p \leqslant m.$$

Полагая здесь $m \to \infty$, а затем $p \to \infty$, получим окончательно:

$$\| a - a^{(n)} \|_{l_2}^2 \leqslant \frac{1}{\lambda_0} \| u - u_n \|^2 \underset{n \to \infty}{\longrightarrow} 0.$$

ГЛАВА II

ОБ УСТОЙЧИВОСТИ ПРОЦЕССОВ РИТЦА И БУБНОВА — ГАЛЁРКИНА ДЛЯ СТАЦИОНАРНЫХ ЗАДАЧ

§ 6. Замечания о процессе Ритца

Наиболее общую вариационную задачу, соответствующую линейным стационарным задачам математической физики, и процесс Ритца ее приближенного решения можно сформулировать, по-видимому, следующим образом[1]).

Пусть $A[u]$ — однородный квадратичный функционал, область определения которого $D(A)$ линейна и плотна в гильбертовом пространстве H; пусть квадратичному функционалу $A[u]$ соответствует билинейный функционал $A[u, v]$ с той же областью определения, так что $A[u] = A[u, u]$. Функционал $A[u]$ считаем положительным:

$$A[u] > 0, \qquad u \neq 0.$$

Пусть, далее, $l(u)$ — линейный (т. е. аддитивный и однородный) функционал такой, что его область определения $D(l) \supset D(A)$.

Ставится задача о минимуме функционала

$$F(u) = A[u] - l(u) - \overline{l(u)}. \tag{6.1}$$

Эта задача в общем случае не имеет решения, и мы ее видоизменяем следующим образом. На множестве $D(A)$ вводим новое скалярное («энергетическое») произведение

$$[u, v] = [u, v]_A = A[u, v] \tag{6.2}$$

и новую («энергетическую») норму

$$|u| = |u|_A = \sqrt{[u, u]_A} = \sqrt{A[u]} \tag{6.3}$$

и замыкаем множество $D(A)$ в этой норме.

В результате замыкания получаем новое гильбертово пространство, которое обозначим через H_A и будем называть *энергетическим пространством*.

---

[1]) См. ПМ, §§ 2—8.

Если линейный функционал $l(u)$ неограничен в метрике (6.3), то функционал (6.1) неограничен снизу, и задача о минимуме этого функционала лишена смысла. Пусть теперь функционал $l(u)$ ограничен в метрике (6.3). Его область определения $D(l)$ содержит множество $D(A)$, плотное в H_A, и потому эта область сама плотна в H_A; определенный на плотном множестве ограниченный в пространстве H_A функционал $l(u)$ может быть расширен по непрерывности на все пространство H_A. Но тогда и функционал (6.1), который теперь можно представить в виде

$$F(u) = |u|_A^2 - l(u) - \overline{l(u)}, \tag{6.4}$$

также может быть расширен на все пространство H_A. Поставим теперь задачу *о минимуме функционала* (6.4) *в пространстве* H_A. Эта задача имеет решение, и притом единственное: по теореме Ф. Риса существует единственный элемент $u_0 \in H_A$ такой, что $l(u) = [u, u_0]_A$. Теперь

$$F(u) = |u - u_0|_A^2 - |u_0|_A^2, \tag{6.5}$$

и решением новой вариационной задачи является элемент u_0. Этот элемент рассматривается как обобщенное решение задачи о минимуме функционала (6.1).

Будем в дальнейшем предполагать, что функционал $l(u)$ ограничен в пространстве H_A и, следовательно, существует решение u_0 задачи о минимуме функционала (6.4).

Допустим, что пространство H_A сепарабельно.

Желая найти приближенное значение элемента u_0, мы выбираем *координатную систему*. Так мы называем последовательность элементов $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$, удовлетворяющую следующим трем требованиям: 1) $\varphi_n \in H_A$ $(n = 1, 2, \ldots)$, 2) при любом n элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы, 3) система элементов $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ полна в H_A. Элементы координатной системы обычно называют *координатными элементами* (см. ВМ, § 14). Приближенное решение ищем в виде[1])

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k \tag{6.6}$$

и постоянные $a_k^{(n)}$ («коэффициенты Ритца») подбираем из условия $F(u_n) = \min$; из формулы (6.5) видно, что это условие равносильно такому:

$$|u_n - u_0|_A = \left| u_0 - \sum_{k=1}^{n} a_k^{(n)} \varphi_k \right|_A = \min. \tag{6.7}$$

---

[1]) В ПМ и ВМ было принято обозначение коэффициентов Ритца через a_k. Для целей настоящей книги важно подчеркнуть зависимость коэффициентов Ритца от n, почему мы и вводим для них обозначение $a_k^{(n)}$.

Из формулы (6.7) вытекает следующее. Пусть u_0 — решение задачи о минимуме функционала (6.4). *Построение элемента* (6.6) — *n-го приближения по Ритцу к u_0 — равносильно построению такой линейной комбинации элементов $\varphi_1, \varphi_2, \ldots, \varphi_n$, которая в метрике H_A является наилучшим приближением к u_0.*

Условие $F(u_n) = \min$ легко приводит к системе Ритца, служащей для определения коэффициентов Ритца $a_k^{(n)}$; в данном случае эта система имеет вид:

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]\, a_k^{(n)} = l\varphi_j \qquad (j = 1, 2, \ldots, n). \tag{6.8}$$

Рассмотрим, как реализуется описанная выше схема в основных вариационных методах.

В случае энергетического метода $A[u] = (Au, u)$, где A — положительный в гильбертовом пространстве H оператор, и $l(u) = (u, f)$, где f — заданный элемент пространства H. Если оператор A — положительно определенный, то функционал $l(u) = (u, f)$ ограничен в H_A и задача о минимуме функционала (6.4) имеет решение. Если же оператор A только положителен, то упомянутая задача разрешима лишь для таких f, для которых скалярное произведение (u, f) ограничено в H_A.

В только что описанной форме энергетический метод соответствует задачам математической физики с однородными краевыми условиями. Рассмотрим некоторые примеры неоднородных краевых условий.

Пусть Ω — конечная или бесконечная область m-мерного евклидова пространства, и S — ее граница.

Задача Дирихле

$$\Delta v = 0, \quad v|_{S_1} = g(x), \qquad x \in S \tag{6.9}$$

сводится к вариационной задаче[1]) о минимуме интеграла

$$\int_{\Omega} |\operatorname{grad} v|^2\, dx \tag{6.10}$$

на множестве функций, имеющих квадратично суммируемые первые производные и удовлетворяющих краевому условию (6.9); при этом предполагается, что существует хотя бы одна функция $\psi(x)$ такая, что $\psi|_S = g(x)$ и

$$\int_{\Omega} |\operatorname{grad} \psi|^2\, dx < \infty. \tag{6.11}$$

---

[1]) См., например, ВМ, стр. 116.

Область определения функционала (6.10) в сформулированной сейчас вариационной задаче, очевидно, нелинейна, поэтому мы введем в рассмотрение функцию $u(x) = v(x) - \psi(x)$. Тогда

$$u(x)|_S = 0 \tag{6.12}$$

и

$$\int_\Omega |\operatorname{grad} v|^2\,dx = \int_\Omega |\operatorname{grad} u|^2\,dx - \int_\Omega \operatorname{grad} u \cdot \operatorname{grad}\psi\,dx - \\ - \int_\Omega \operatorname{grad}\psi \cdot \operatorname{grad} u\,dx + \int_\Omega |\operatorname{grad}\psi|^2\,dx.$$

Последний интеграл есть величина постоянная, и сформулированная выше вариационная задача равносильна задаче о минимуме функционала

$$F(u) = \int_\Omega |\operatorname{grad} u|^2\,dx - \\ - \int_\Omega \operatorname{grad} u \cdot \operatorname{grad}\psi\,dx - \int_\Omega \operatorname{grad}\psi \cdot \operatorname{grad} u\,dx \tag{6.13}$$

на линейном уже множестве функций, удовлетворяющих условию (6.12) и имеющих квадратично суммируемые в Ω первые производные. В данном случае

$$A[u] = |u|_A^2 = \int_\Omega |\operatorname{grad} u|^2\,dx, \tag{6.14}$$

$$l(u) = \int_\Omega \operatorname{grad} u \cdot \operatorname{grad}\psi\,dx; \tag{6.15}$$

функционал (6.15) ограничен в норме (6.14), так как по неравенству Буняковского

$$\left|\int_\Omega \operatorname{grad} u \cdot \operatorname{grad}\psi\,dx\right| \leqslant \left\{\int_\Omega |\operatorname{grad} u|^2\,dx\right\}^{\frac{1}{2}} \cdot \left\{\int_\Omega |\operatorname{grad}\psi|^2\,dx\right\}^{\frac{1}{2}} = \\ = \left\{\int_\Omega |\operatorname{grad}\psi|^2\,dx\right\}^{\frac{1}{2}} |u|_A,$$

и задача о минимуме функционала (6.13) имеет решение.

Рассмотрим еще задачу Неймана с неоднородным краевым условием.

Для упрощений рассуждений поставим эту задачу не для уравнения Лапласа, а для уравнения $-\Delta u + u = 0$. Мы рассмотрим, следовательно, задачу

$$-\Delta u + u = 0, \qquad \left.\frac{\partial u}{\partial \nu}\right|_S = h(x), \qquad x \in S. \tag{6.16}$$

Область Ω будем считать конечной, ее границу S — состоящей из конечного числа кусков достаточно гладких поверхностей. Легко доказать[1]), что задача (6.16) равносильна задаче о минимуме функционала

$$\int_{\Omega} \{|\operatorname{grad} u|^2 + |u|^2\}\, dx - \int_{S} u\bar{h}\, dS - \int_{S} \bar{u}h\, dS \qquad (6.17)$$

на линейном множестве функций, имеющих квадратично суммируемые первые производные. В данном случае это множество образует пространство H_A; метрика в этом пространстве определяется формулой

$$|u|_A^2 = \int_{\Omega} \{|\operatorname{grad} u|^2 + |u|^2\}\, dx. \qquad (6.18)$$

Наша вариационная задача разрешима, если функционал

$$l(u) = \int_{S} hu\, dS \qquad (6.19)$$

ограничен в норме (6.18). Докажем, что такая ограниченность имеет место, если $h \in L_2(S)$, так что

$$K = \int_{S} |h|^2\, dS < \infty.$$

Действительно, в этом случае, по неравенству Буняковского,

$$\left|\int_{S} hu\, dS\right|^2 \leqslant K \int_{S} |u|^2\, dS.$$

Далее, из теорем вложения С. Л. Соболева[2])

$$\int_{S} |u|^2\, dS \leqslant K_1 \int_{\Omega} \{|\operatorname{grad} u|^2 + |u|^2\}\, dx = K_1 |u|_A^2, \qquad K_1 = \text{const}.$$

Отсюда

$$\left|\int_{S} hu\, dS\right| \leqslant \sqrt{KK_1}\, |u|_A$$

и функционал (6.19) на самом деле ограничен в метрике (6.18).

Рассмотрим теперь другие вариационные методы. В методе наименьших квадратов решение уравнения $Au = f$, $f \in H$, заменяется отысканием минимума величины

$$\|Au - f\|^2 \qquad (6.20)$$

---

[1]) Ср. ВМ, § 18.

[2]) См. С. Л. Соболев [1], § 8, п. 2, стр. 64—65, а также В. И. Смирнов [3], п. 114, стр. 359, теорема 2.

или, что равносильно, функционала

$$(Au,\ Au) - (Au,\ f) - (f,\ Au) \tag{6.20$'$}$$

в области $D(A)$ определения линейного (не обязательно симметричного) оператора A. В данном случае $A[u] = (Au,\ Au) = \| Au \|^2$; если оператор A замкнут и имеет ограниченный обратный (что мы будем предполагать), то пространство H_A состоит из элементов области $D(A)$ и норма в этом пространстве задается формулой

$$|u|_A = \| Au \|. \tag{6.21}$$

Далее, функционал $l(u) = (Au,\ f)$ ограничен в H_A, так как

$$|l(u)| = |(Au,\ f)| \leqslant \| Au \| \| f \| = \| f \| \cdot |u|_A.$$

Из теоремы Ф. Риса вытекает существование элемента $u_0 \in D(A)$ такого, что

$$(Au,\ f) = [u,\ u_0]_A = (Au,\ Au_0).$$

Теперь, как легко видеть,

$$\| Au - f \|^2 = |u - u_0|_A^2 + \| f \|^2 - \| Au_0 \|^2,$$

и минимум функционала (6 19) равен $\| f \|^2 - \| Au_0 \|^2$; применение процесса Ритца равносильно наилучшему приближению в метрике (6.21) линейными комбинациями координатных функций к элементу u_0.

Решением задачи о минимуме функционала (6.20) является элемент u_0; приближенные по Ритцу решения этой задачи сходятся к u_0 в метрике (6.21). Так как обратный оператор A^{-1} ограничен, то сходимость имеет место и в метрике пространства H; если при этом оператор A^{-1} определен на всем пространстве H, то элемент u_0 удовлетворяет исходному уравнению $Au = f$.

Метод ортогональных проекций состоит[1]) в том, что минимизируется величина $\left\| V - \sum_{k=1}^{n} a_k \psi_k \right\|^2$, где V — данный элемент некоторого гильбертова пространства $\mathfrak{H}$, а координатные элементы $\psi_1,\ \psi_2,\ \ldots$ принадлежат некоторому подпространству $\mathfrak{H}_2$ пространства $\mathfrak{H}$.

При этом $V = v_0 + w_0$, где v_0 — искомый элемент в методе ортогональных проекций, принадлежащий подпространству $\mathfrak{H}_1$, ортогональному к $\mathfrak{H}_2$, а $w_0 \in \mathfrak{H}_2$.

Элемент

$$w_n = \sum_{k=1}^{n} a_k \psi_k,$$

[1]) См. ВМ, § 51.

обращающий в минимум величину

$$\|V - w_n\|^2 = \left\| V - \sum_{k=1}^{n} a_k \psi_k \right\|^2,$$

является наилучшим приближением к элементу w_0.

Тем самым метод ортогональных проекций непосредственно включается в схему, изложенную в начале настоящего параграфа.

Действительно, в данном случае можно считать, что функционал (6.4) имеет вид $\|V - w\|^2$, где $w \in \mathfrak{H}_2$. Энергетическая метрика совпадает с метрикой пространства $\mathfrak{H}$, а энергетическое пространство — с $\mathfrak{H}_2$. Минимизируя упомянутую выше величину

$$\left\| V - \sum_{k=1}^{n} a_k \psi_k \right\|^2 = \|V - w_n\|^2,$$

мы тем самым строим приближение по Ритцу к элементу w_0, при этом элементы $\psi_1, \psi_2, \ldots$ служат координатными; в то же время, как мы только что отметили, w_n есть наилучшее приближение к w_0.

В заключение рассмотрим метод Трефца. Для простоты ограничимся задачей (6.9). По методу Трефца[1]) мы вводим в рассмотрение гильбертово пространство G_2^1 функций, гармонических в Ω и имеющих в этой области квадратично суммируемые первые производные; норму в этом пространстве задаем формулой

$$|u|^2 = \Lambda(u) = \int_{\Omega} |\operatorname{grad} u|^2 \, dx; \tag{6.22}$$

вводя такую норму, мы тем самым уславливаемся не различать функции, которые отличаются одна от другой на постоянное слагаемое. В определенном так пространстве выбираем координатную систему $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x), \ldots,$ задаем натуральное число n и определяем постоянные $a_1, a_2, \ldots, a_n$ из условия

$$\Lambda(u_0 - u_n) = \min, \quad u_n = \sum_{k=1}^{n} a_k \varphi_k, \tag{6.23}$$

где u_0 — решение задачи (6.9). В данном случае $A[u] = \Lambda(u)$, пространство H_A есть построенное выше пространство G_2^1 гармонических функций, $|u|_A = |u| = \sqrt{\Lambda(u)}$; метод Трефца по существу заключается в том, что задача о минимуме функционала $\Lambda(u_0 - u)$ в пространстве G_2^1 приближенно решается процессом Ритца или, что то же, в том, что строится наилучшее приближение к искомой функции u_0 в метрике (6.22).

---

1) См., например, ВМ, § 55.

§ 7. Предельные свойства коэффициентов Ритца

Рассмотрим задачу о минимуме функционала (6.1), и пусть эта задача имеет обобщенное решение u_0. Пусть выбрана координатная система

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, \tag{7.1}$$

удовлетворяющая трем условиям, перечисленным в § 6, и пусть построена последовательность приближенных по Ритцу решений

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k. \tag{7.2}$$

Как было показано в предшествующем параграфе, элементы u_n суть наилучшие приближения к искомому решению u_0 в пространстве H_A; из теорем 5.1 и 5.2 теперь вытекают следующие теоремы о предельных свойствах коэффициентов Ритца.

Теорема 7.1. *Если координатная система* (7.1) *минимальна в H_A, то существуют пределы коэффициентов Ритца*

$$a_k = \lim_{n\to\infty} a_k^{(n)} \qquad (k = 1, 2, \ldots). \tag{7.3}$$

Если, кроме того, система, биортонормированная к системе (7.1) *в H_A, ограничена в метрике этого пространства, то стремление к пределу в формуле* (7.3) — *равномерное относительно k.*

Теорема 7.2. *Если координатная система* (7.1) *сильно минимальна в H_A, то последовательность*

$$a = (a_1, a_2, \ldots, a_n, \ldots),$$

где a_k суть пределы (7.3), *есть элемент пространства l_2; если положить*

$$a^{(n)} = \left(a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}, 0, \ldots\right),$$

то

$$\lim_{n\to\infty} \| a^{(n)} - a \|_{l_2} = 0.$$

Если координатная система не минимальна в H_A, то пределы коэффициентов $a_k^{(n)}$ могут не существовать; в счете это обнаруживается тем, что при фиксированном k значения коэффициента $a_k^{(n)}$ иногда начинают резко меняться при изменении n.

Приведем два примера.

Пример 7.1. Будем решать по методу Ритца задачу

$$-\frac{d^2u}{dx^2} = \frac{1}{1+x}, \qquad u(0) = u'(1) = 0, \tag{7.4}$$

равносильную задаче о минимуме интеграла [1]):

$$\int_0^1 \left(u'^2 - 2\frac{1}{1+x}u\right)dx, \qquad u(0)=0. \tag{7.5}$$

В нашем примере $H = L_2(0, 1)$. Пространство H_A состоит из функций, которые абсолютно непрерывны на отрезке $0 \leqslant x \leqslant 1$, обращаются в нуль на его левом конце и имеют первые производные, суммируемые с квадратом на упомянутом отрезке; скалярное произведение и норма в H_A определены формулами

$$[u, v] = \int_0^1 u'(x)\, v'(x)\, dx, \qquad |u|^2 = \int_0^1 u'^2(x)\, dx. \tag{7.6}$$

В качестве координатной системы выберем последовательность x^k ($k = 1, 2, \ldots$); нетрудно видеть, что она удовлетворяет требованиям 1) — 3) § 6. Вычислим первые восемь приближений по Ритцу. Система Ритца для восьмого приближения имеет вид:

$$\left.\begin{aligned}
&a_1^{(8)} + a_2^{(8)} + a_3^{(8)} + a_4^{(8)} + a_5^{(8)} + a_6^{(8)} + a_7^{(8)} + a_8^{(8)} = 1 - \ln 2,\\
&a_1^{(8)} + \tfrac{4}{3}a_2^{(8)} + \tfrac{3}{2}a_3^{(8)} + \tfrac{8}{5}a_4^{(8)} + \tfrac{5}{3}a_5^{(8)} + \tfrac{12}{7}a_6^{(8)} + \tfrac{7}{4}a_7^{(8)} + \tfrac{16}{9}a_8^{(8)} = \ln 2 - \tfrac{1}{2},\\
&a_1^{(8)} + \tfrac{3}{2}a_2^{(8)} + \tfrac{9}{5}a_3^{(8)} + 2a_4^{(8)} + \tfrac{15}{7}a_5^{(8)} + \tfrac{9}{4}a_6^{(8)} + \tfrac{7}{3}a_7^{(8)} + \tfrac{12}{5}a_8^{(8)} = \tfrac{5}{6} - \ln 2,\\
&a_1^{(8)} + \tfrac{8}{5}a_2^{(8)} + 2a_3^{(8)} + \tfrac{16}{7}a_4^{(8)} + \tfrac{5}{2}a_5^{(8)} + \tfrac{8}{3}a_6^{(8)} + \tfrac{14}{5}a_7^{(8)} + \tfrac{32}{11}a_8^{(8)} = \ln 2 - \tfrac{7}{12},\\
&a_1^{(8)} + \tfrac{5}{3}a_2^{(8)} + \tfrac{15}{7}a_3^{(8)} + \tfrac{5}{2}a_4^{(8)} + \tfrac{25}{9}a_5^{(8)} + 3a_6^{(8)} + \tfrac{35}{11}a_7^{(8)} + \tfrac{10}{3}a_8^{(8)} = \tfrac{47}{60} - \ln 2,\\
&a_1^{(8)} + \tfrac{12}{7}a_2^{(8)} + \tfrac{9}{4}a_3^{(8)} + \tfrac{8}{3}a_4^{(8)} + 3a_5^{(8)} + \tfrac{36}{11}a_6^{(8)} + \tfrac{7}{2}a_7^{(8)} + \tfrac{48}{13}a_8^{(8)} = \ln 2 - \tfrac{37}{60},\\
&a_1^{(8)} + \tfrac{7}{4}a_2^{(8)} + \tfrac{7}{3}a_3^{(8)} + \tfrac{14}{5}a_4^{(8)} + \tfrac{35}{11}a_5^{(8)} + \tfrac{7}{2}a_6^{(8)} + \tfrac{49}{13}a_7^{(8)} + 4a_8^{(8)} = \tfrac{319}{420} - \ln 2,\\
&a_1^{(8)} + \tfrac{16}{9}a_2^{(8)} + \tfrac{12}{5}a_3^{(8)} + \tfrac{32}{11}a_4^{(8)} + \tfrac{10}{3}a_5^{(8)} + \tfrac{48}{13}a_6^{(8)} + 4a_7^{(8)} + \tfrac{64}{15}a_8^{(8)} = \ln 2 - \tfrac{533}{840};
\end{aligned}\right\} \tag{7.7}$$

системы Ритца для предшествующих приближений получаются из системы (7.7) усечением (см. ВМ, замечание на стр. 93). Решая усеченные системы, а также саму систему (7.7), получим следующую таблицу точных значений коэффициентов Ритца (табл. 7.1).

---

[1]) Решение задачи (7.4) вещественно, поэтому мы ограничиваемся рассмотрением вещественных функций. Это замечание следует иметь в виду и в последующем.

Таблица 7.1

n	$a_1^{(n)}$	$a_2^{(n)}$	$a_3^{(n)}$	$a_4^{(n)}$	$a_5^{(n)}$	$a_6^{(n)}$	$a_7^{(n)}$	$a_8^{(n)}$
1	$1-\ln 2$							
2	$5\frac{1}{2}-7\ln 2$	$-4\frac{1}{2}+6\ln 2$						
3	$26\frac{1}{3}-37\ln 2$	$-67+96\ln 2$	$41\frac{2}{3}-60\ln 2$					
4	$133\frac{1}{12}-191\ln 2$	$-707\frac{1}{2}+1\,020\ln 2$	$1\,109\frac{1}{6}-1\,600\ln 2$	$-533\frac{3}{4}+770\ln 2$				
5	$694\frac{8}{15}-1\,001\ln 2$	$9\,120\ln 2-6\,322$	$17\,952\frac{2}{3}-25\,900\ln 2$	$29\,120\ln 2-20\,184\frac{1}{2}$	$7\,860\frac{3}{10}-11\,340\ln 2$			
6	$3\,698\frac{19}{30}-5\,335\ln 2$	$74\,130\ln 2-51\,383\frac{1}{2}$	$228\,239\frac{2}{3}-329\,280\ln 2$	$635\,880\ln 2-440\,758\frac{1}{2}$	$386\,376\frac{9}{10}-557\,424\ln 2$	$182\,028\ln 2-126\,172\frac{1}{5}$		
7	$19\,972\frac{12}{35}-28\,813\ln 2$	$567\,168\ln 2-393\,131\frac{2}{5}$	$2\,506\,559-3\,616\,200\ln 2$	$10\,496\,640\ln 2-7\,275\,716\frac{1}{2}$	$10\,638\,813\frac{9}{10}-15\,348\,564\ln 2$	$11\,028\,864\ln 2-7\,644\,626$	$2\,148\,129\frac{23}{35}-3\,099\,096\ln 2$	
8	$108\,951\frac{181}{280}-157\,183\ln 2$	$4\,161\,528\ln 2-2\,884\,551\frac{9}{10}$	$24\,929\,343\frac{1}{2}-35\,965\,440\ln 2$	$145\,285\,140\ln 2-100\,703\,985\frac{1}{4}$	$216\,181\,005\frac{3}{20}-311\,883\,264\ln 2$	$366\,870\,504\ln 2-254\,295\,255\frac{1}{2}$	$154\,836\,614\frac{41}{70}-223\,382\,016\ln 2$	$55\,070\,730\ln 2-38\,172\,121\frac{13}{56}$

Подставив в табл. 7.1 значение

$$\ln 2 = 0,69314718055994 5,$$

получим следующие значения коэффициентов Ритца, в которых выписанные знаки верны (табл. 7.2). Из табл. 7.2 видно, что хотя

Таблица 7.2

n	$a_1^{(n)}$	$a_2^{(n)}$	$a_3^{(n)}$	$a_4^{(n)}$	$a_5^{(n)}$	$a_6^{(n)}$	$a_7^{(n)}$	$a_8^{(n)}$
1	0,3069							
2	0,6480	—0,3411						
3	0,6869	—0,4579	0,0788					
4	0,6922	—0,4899	0,1312	—0,0267				
5	0,6930	—0,4977	0,1547	—0,0541	0,0110			
6	0,6931	—0,4995	0,1631	—0,0708	0,0260	—0,0050		
7	0,6931	—0,4999	0,1657	—0,0786	0,0378	—0,0136	0,0025	
8	0,6931	—0,5000	0,1664	—0,0817	0,0446	—0,0218	0,0075	—0,0013

с возрастанием номера приближения коэффициентов Ритца проявляют известную тенденцию к стабилизации, но эта тенденция проявляется неравномерно: в коэффициенте $a_1^{(n)}$ первые четыре знака стабилизируются, начиная с $n = 5$, в коэффициенте $a_2^{(n)}$ — с $n = 7$, в остальных коэффициентах виден довольно заметный разброс. По-видимому, в рассматриваемом примере коэффициенты $a_k^{(n)}$ либо не стремятся к пределам при $n \to \infty$, либо стремление во всяком случае неравномерное. Это явление можно объяснить тем, что выбранная нами координатная система неминимальна в пространстве H_A, т. е. в метрике (7.6). Последнее обстоятельство можно установить так: из формулы (7.6) видно, что минимальность системы функций в H_A в нашем примере равносильна минимальности системы их производных в метрике $L_2(0, 1)$, а в этом пространстве kx^{k-1} $(k = 1, 2, \ldots)$ неминимальна в силу теоремы Мюнца.

По данным табл. 7.2 легко восстановить и сами приближенные решения; мы на этом не останавливаемся ввиду полной очевидности этого действия.

Несмотря на отмеченный выше разброс коэффициентов, полученные нами приближенные решения довольно точны. Так, при $x = 1$ получаем:

$$u_1(1) = 0,3069; \quad u_2(1) = 0,3069; \quad u_3(1) = 0,3071; \quad u_4(1) = 0,3068.$$

Точное значение

$$u_0(x) = x(1 + \ln 2) - (x + 1)\ln(x + 1),$$

откуда $u_0(1) = 1 - \ln 2 = 0,3069$; $\quad u_0(1) - u_4(1) = 0,0001$.

Близость приближенных решений к точному становится отчетливой, если составить энергетические нормы разностей $u_0 - u_k$. Имеем (ВМ, стр. 94, формула (17)): $|u_0 - u_n|_A^2 = |u_0|_A^2 - |u_n|_A^2$. Величина $|u_0|_A^2$ непосредственно вычисляется по формуле (7.6):

$$|u_0|_A^2 = 2(1 - \ln 2) - \ln^2 2 = 0{,}1333.$$

Что касается величин $|u_n|_A^2$, то их проще вычислять не непосредственно по формуле (7.6), а по следующей формуле (ВМ, стр. 93, формула (12), и стр. 283, формула (8)):

$$|u_n|_A^2 = \sum_{k=1}^{n} a_k^{(n)} (f, \varphi_k), \tag{7.8}$$

где (f, φ_k) суть свободные члены системы Ритца. Это дает нам, например,

$$|u_1|_A^2 = 0{,}0942; \quad |u_2|_A^2 = 0{,}1330; \quad |u_3|_A^2 = 0{,}1332; \quad |u_4|_A^2 = 0{,}1332$$

и, следовательно,

$$|u_0 - u_1|_A = 0{,}20; \qquad |u_0 - u_2|_A = 0{,}02;$$
$$|u_0 - u_3|_A = 0{,}01; \qquad |u_0 - u_4|_A = 0{,}01.$$

Из формулы (7.6) и неравенства Буняковского вытекает равномерная оценка погрешности $|u_0(x) - u_n(x)|$:

$$|u_0(x) - u_n(x)| \leqslant |u_0 - u_n|_A.$$

Отсюда, в частности, $|u_0(1) - u_4(1)| \leqslant 0{,}01$. Фактическая погрешность, как мы видели, значительно ниже.

Пример 7.2. Приведем еще один пример, где разброс коэффициентов еще более заметен. Этот пример разобран в книге автора [4], § 38.

Рассмотрим мембрану, имеющую форму равнобедренного прямоугольного треугольника и подверженную действию равномерной нормальной нагрузки. Допустим, что мембрана закреплена по катетам, а гипотенуза остается свободной. Направим оси координат x и y по катетам. При подходящем выборе единиц измерения задача приводится к интегрированию уравнения

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 1$$

в области треугольника $x \geqslant 0$, $y \geqslant 0$, $x + y - 1 \leqslant 0$, при краевых условиях

$$u = 0 \text{ на прямых } x = 0, \quad y = 0,$$

$$\frac{\partial u}{\partial \nu} = 0 \text{ на прямой } x + y - 1 = 0;$$

как обычно, ν означает нормаль к контуру. Эта последняя задача в свою очередь сводится к задаче о минимуме интеграла:

$$F(u)=\int_0^1\left\{\int_0^x\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2-2u\right]dy\right\}dx \tag{7.9}$$

при краевых условиях

$$u(x,\ 0)=u(0,\ y)=0. \tag{7.10}$$

Пространство H_A в данном случае состоит из функций, имеющих в области мембраны квадратично суммируемые обобщенные первые производные и удовлетворяющих (в смысле С. Л. Соболева [1]) условиям (7.10); энергетическое произведение и энергетическая норма определяются формулами

$$\left.\begin{aligned}[u,\ v]_A&=\int_0^1\left\{\int_0^x\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dy\right\}dx,\\ |u|_A^2&=\int_0^1\left\{\int_0^x\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dy\right\}dx.\end{aligned}\right\} \tag{7.11}$$

В качестве координатной возьмем систему

$$x^k y^m \qquad (k,\ m=1,\ 2,\ 3,\ \ldots); \tag{7.12}$$

нетрудно видеть, что она удовлетворяет трем требованиям, перечисленным в § 6. В то же время эта система неминимальна в H_A. Для доказательства положим $H_A=H_2$ и рассмотрим пространство H_1 функций, определенных почти всюду в квадрате $0\leqslant x,\ y\leqslant 1$, имеющих в этом квадрате обобщенные квадратично суммируемые первые производные и удовлетворяющих в смысле С. Л. Соболева условиям (7.10); скалярное произведение и норму в H_1 зададим формулами

$$(u,\ v)_1=\int_0^1\int_0^1\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dx\,dy, \tag{7.13}$$

$$\|u\|_1^2=\int_0^1\int_0^1\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dx\,dy. \tag{7.14}$$

Обозначим через V оператор, который каждой функции $u(x,\ y)$, заданной в квадрате $0\leqslant x,\ y\leqslant 1$, приводит в соответствие ту же функцию, но определенную только в треугольнике $x\geqslant 0$, $y\geqslant 0$, $x+y-1\leqslant 0$. Очевидно, что V отображает H_1 на H_A и что $|Vu|_A\leqslant\|u\|_1$. Сравнивая это с неравенством (4.1), видим, что оператор V осуществляет вложение пространства H_1 в пространство H_A.

Докажем теперь, что система (7.12) неминимальна в H_1; из теоремы 4.1 будет вытекать тогда, что эта система неминимальна и в H_A.

Рассмотрим в H_1 подпространство $H_1^{(m)}$, натянутое на элементы xy^m, x^2y^m, ... Очевидно,

$$H_1 = \bigcup_{m=1}^{\infty} H_1^{(m)}.$$

Как было отмечено в § 1, из теоремы Мюнца вытекает, что из последовательности nx^{n-1} $(n = 1, 2, \ldots)$ можно вычеркнуть любой член, не нарушая этим ее полноты в $L_2(0, 1)$. Отсюда вытекает, что каков бы ни был номер k_0, по данному $\varepsilon > 0$ можно найти натуральное число N и постоянные $\alpha_1, \alpha_2, \ldots, \alpha_{k_0-1}, \alpha_{k_0+1}, \ldots, \alpha_N$ такие, чтобы имело место неравенство

$$\left\| k_0 x^{k_0-1} - \sum_{n=1,\, n\neq k_0}^{N} \alpha_n n x^n \right\|_{L_2}^2 = \int_0^1 \left[k_0 x^{k_0-1} - \sum_{n=1,\, n\neq k_0}^{N} \alpha_n n x^n \right]^2 dx < \varepsilon^2.$$

Обозначив для краткости

$$f(x) = x^{k_0} - \sum_{n=1,\, n\neq k_0}^{N} \alpha_n x^n,$$

имеем, следовательно,

$$\| f'(x) \|_{L_2}^2 = \int_0^1 [f'(x)]^2 dx < \varepsilon^2.$$

Оценим величину

$$\| f(x) \|_{L_2}^2 = \int_0^1 [f(x)]^2 dx.$$

Так как $f(0) = 0$, то

$$f(x) = \int_0^x f'(t)\, dt.$$

Отсюда, по неравенству Буняковского,

$$[f(x)]^2 \leqslant x \int_0^x [f'(t)]^2 dt \leqslant \| f'(x) \|_{L_2}^2 < \varepsilon^2$$

и, следовательно,

$$\| f(x) \|_{L_2}^2 < \varepsilon^2.$$

Теперь нетрудно оценить величину

$$\left\| x^{k_0} y^m - \sum_{n=1,\, n\neq k_0}^{N} \alpha_n x^n y^m \right\|_1 = \| f(x)\, y^m \|_1.$$

Имеем:

$$\| f(x)\, y^m \|_1^2 = \frac{1}{2m+1} \| f'(x) \|_{L_2}^2 + \frac{\mathring{m}^2}{2m-1} \| f(x) \|_{L_2}^2 \leqslant C\varepsilon^2, \qquad C = \text{const}.$$

Последнее неравенство показывает, что последовательность xy^m, x^2y^m, ... неминимальна в $H_1^{(m)}$. Отсюда следует, что система (7.12) неминимальна в H_1. Как было отмечено выше, отсюда в свою очередь вытекает, что система (7.12) неминимальна в H_A.

Из симметрии задачи ясно, что координатные функции x^ky^m и x^my^k войдут в приближенное решение с одинаковыми коэффициентами, поэтому систему (7.12) можно заменить такой: $(xy)^k(x^m + y^m)$ $(k = 1, 2, 3, \ldots;\ m = 0, 1, 2, 3, \ldots)$. Если положить

$$u_4 = a_1^{(4)}xy + a_2^{(4)}xy(x+y) + a_3^{(4)}x^2y^2 + a_4^{(4)}xy(x^2+y^2),$$

то получим следующую систему Ритца:

$$\begin{aligned}
\frac{1}{6}\, a_1^{(4)} + \frac{1}{6}\, a_2^{(4)} + \frac{1}{30}\, a_3^{(4)} + \frac{1}{10}\, a_4^{(4)} &= \frac{1}{24}, \\
\frac{1}{6}\, a_1^{(4)} + \frac{8}{45}\, a_2^{(4)} + \frac{4}{105}\, a_3^{(4)} + \frac{23}{210}\, a_4^{(4)} &= \frac{1}{30}, \\
\frac{1}{30}\, a_1^{(4)} + \frac{4}{105}\, a_2^{(4)} + \frac{1}{105}\, a_3^{(4)} + \frac{19}{840}\, a_4^{(4)} &= \frac{1}{180}, \\
\frac{1}{10}\, a_1^{(4)} + \frac{23}{210}\, a_2^{(4)} + \frac{19}{840}\, a_3^{(4)} + \frac{1}{14}\, a_4^{(4)} &= \frac{1}{60};
\end{aligned}$$

отсюда усечением получаются системы Ритца для коэффициентов $a_k^{(n)}$ $(n = 1, 2, 3)$. Приводим таблицу точных значений коэффициентов Ритца (табл. 7.3).

Таблица 7.3

n	$a_1^{(n)}$	$a_2^{(n)}$	$a_3^{(n)}$	$a_4^{(n)}$	n	$a_1^{(n)}$	$a_2^{(n)}$	$a_3^{(n)}$	$a_4^{(n)}$
1	$\frac{1}{4}$				3	$1\frac{2}{9}$	$-1\frac{1}{6}$	$\frac{35}{36}$	
2	1	$-\frac{3}{4}$			4	$1\frac{6}{19}$	$-1\frac{3}{4}$	$1\frac{8}{13}$	$\frac{14}{39}$

В таблице 7.4 даны значения тех же коэффициентов с пятью знаками после запятой

Как видно из табл. 7.4, в данном примере разброс коэффициентов Ритца оказался довольно значительным: эти коэффициенты резко меняются при переходе от одного приближения к другому.

Таблица 7.4

n	$a_1^{(n)}$	$a_2^{(n)}$	$a_3^{(n)}$	$a_4^{(n)}$
1	0,25000			
2	1,00000	—0,75000		
3	1,22222	—1,16667	0,97222	
4	1,46154	—1,75000	1,61538	0,35897

В отличие от примера 7.1 настоящего параграфа, здесь мы не располагаем точным решением задачи, и потому трудно судить о погрешности приближенных решений. Мы можем, однако, вычислить энергетические нормы разностей $u_n - u_m$ (n, $m = 1, 2, 3, 4$) и по ним составить суждения о близости между собой самих приближенных решений. Воспользуемся формулой (ВМ, стр. 94, формула (16))

$$|u_m - u_n|_A^2 = |u_m|_A^2 - |u_n|_A^2, \qquad m > n,$$

а также формулой (7.8) настоящего параграфа. Эта последняя формула дает:

$$|u_1|_A^2 = \frac{1}{96} = 0{,}01042, \qquad |u_2|_A^2 = \frac{1}{60} = 0{,}01667,$$

$$|u_3|_A^2 = \frac{113}{6480} = 0{,}01744, \qquad |u_4|_A^2 = \frac{41}{2340} = 0{,}01752.$$

Отсюда видно, что приближенные решения по Ритцу различаются по энергетической норме не столь значительно, несмотря на резкое различие в коэффициентах. Так,

$$|u_3 - u_2|_A = \sqrt{|u_3|_A^2 - |u_2|_A^2} = \sqrt{0{,}00077} < 0{,}028,$$

$$|u_4 - u_3|_A = \sqrt{|u_4|_A^2 - |u_3|_A^2} = \sqrt{0{,}00008} < 0{,}009.$$

В отличие от примера 7.1, судить о степени равномерной близости приближенных решений на основе последних неравенств невозможно

По поводу сказанного в настоящем параграфе можно сделать следующее замечание. Сам по себе «разброс» коэффициентов не является дефектом приближенного решения: если оно вычислено без серьезных погрешностей, то при достаточно высоком номере приближения оно будет сколь угодно близко к точному решению, независимо от того, имеет место разброс или нет. Тем не менее, наличие разброса нежелательно: он затрудняет сравнение приближений с разными номерами, в то же время он служит указанием на то, что координатная система не сильно минимальна (даже, может быть, неминимальна) в энергетической метрике; как будет показано ниже, это влечет за собой неустойчивость процесса Ритца по отношению к малым погрешностям промежуточных вычислений.

§ 8. Примеры, подводящие к понятию об устойчивости

Обратимся к системе (7.7) и заменим ее правые части их десятичными приближениями с четырьмя знаками после запятой. Мы получим тогда «неточную» систему Ритца:

$$\left.\begin{aligned}
&b_1^{(8)}+b_2^{(8)}+b_3^{(8)}+b_4^{(8)}+b_5^{(8)}+b_6^{(8)}+b_7^{(8)}+b_8^{(8)}=0{,}3069,\\
&b_1^{(8)}+\tfrac{4}{3}b_2^{(8)}+\tfrac{3}{2}b_3^{(8)}+\tfrac{8}{5}b_4^{(8)}+\tfrac{5}{3}b_5^{(8)}+\tfrac{12}{7}b_6^{(8)}+\tfrac{7}{4}b_7^{(8)}+\tfrac{16}{9}b_8^{(8)}=0{,}1931,\\
&b_1^{(8)}+\tfrac{3}{2}b_2^{(8)}+\tfrac{9}{5}b_3^{(8)}+2b_4^{(8)}+\tfrac{15}{7}b_5^{(8)}+\tfrac{9}{4}b_6^{(8)}+\tfrac{7}{3}b_7^{(8)}+\tfrac{12}{5}b_8^{(8)}=0{,}1402,\\
&b_1^{(8)}+\tfrac{8}{5}b_2^{(8)}+2b_3^{(8)}+\tfrac{16}{7}b_4^{(8)}+\tfrac{5}{2}b_5^{(8)}+\tfrac{8}{3}b_6^{(8)}+\tfrac{14}{5}b_7^{(8)}+\tfrac{32}{11}b_8^{(8)}=0{,}1098,\\
&b_1^{(8)}+\tfrac{5}{3}b_2^{(8)}+\tfrac{15}{7}b_3^{(8)}+\tfrac{5}{2}b_4^{(8)}+\tfrac{25}{9}b_5^{(8)}+3b_6^{(8)}+\tfrac{35}{11}b_7^{(8)}+\tfrac{10}{3}b_8^{(8)}=0{,}0902,\\
&b_1^{(8)}+\tfrac{12}{7}b_2^{(8)}+\tfrac{9}{4}b_3^{(8)}+\tfrac{8}{3}b_4^{(8)}+3b_5^{(8)}+\tfrac{36}{11}b_6^{(8)}+\tfrac{7}{2}b_7^{(8)}+\tfrac{48}{13}b_8^{(8)}=0{,}0765,\\
&b_1^{(8)}+\tfrac{7}{4}b_2^{(8)}+\tfrac{7}{3}b_3^{(8)}+\tfrac{14}{5}b_4^{(8)}+\tfrac{35}{11}b_5^{(8)}+\tfrac{7}{2}b_6^{(8)}+\tfrac{49}{13}b_7^{(8)}+4b_8^{(8)}=0{,}0664,\\
&b_1^{(8)}+\tfrac{16}{9}b_2^{(8)}+\tfrac{12}{5}b_3^{(8)}+\tfrac{32}{11}b_4^{(8)}+\tfrac{10}{3}b_5^{(8)}+\tfrac{48}{13}b_6^{(8)}+4b_7^{(8)}+\tfrac{64}{15}b_8^{(8)}=0{,}0586.
\end{aligned}\right\}\quad(8.1)$$

В этой «неточной» системе матрица не содержит погрешности, а погрешности свободных членов меньше половины единицы последнего десятичного знака.

Система (8.1), а также соответствующие ей усеченные системы, были решены точно, без ошибок округления. Результаты приведены в табл. 8.1.

Сравним эту таблицу с табл. 7.2, которую мы будем рассматривать как точную, так как в ней все выписанные знаки верны. Мы получим тогда таблицу погрешностей:

$$\eta_k^{(n)}=\left|b_k^{(n)}-a_k^{(n)}\right|$$

(табл. 8.2).

Рассмотрим отношение $\eta_k^{(n)}:\delta$, где $\delta=0{,}00005$ есть верхняя граница погрешностей, допущенных нами при вычислении свободных членов системы (8.1). Мы видим, что это отношение быстро растет с увеличением n; оно равно 6 при $n=2$ и достигает значений порядка 10^7 при $n=8$.

При переходе от уравнений (7.7) к уравнениям (8.1) при решении системы (8.1), а также усеченных систем мы не вносили погрешностей в матрицу и избежали погрешностей округления. Если бы такие погрешности были внесены (например, если бы коэффициенты уравнений (8.1) были заменены их десятичными приближениями, а промежуточные выкладки при решении системы производились бы

Таблица 8.1

n	$b_1^{(n)}$	$b_2^{(n)}$	$b_3^{(n)}$	$b_4^{(n)}$	$b_5^{(n)}$	$b_6^{(n)}$	$b_7^{(n)}$	$b_8^{(n)}$
1	0,3069							
2	0,6483	—0,3414						
3	0,6883	—0,4614	0,0800					
4	0,6974	—0,5160	0,1710	—0,0455				
5	0,7127	—0,6690	0,6300	—0,5810	0,2142			
6	0,7633	—1,4280	4,1720	—7,6650	6,5898	—2,1252		
7	0,9141	—4,5948	25,2840	—71,0010	101,5938	—71,7948	19,9056	
8	1,6266	—24,5448	204,8340	—819,1260	1747,4688	—2046,8448	1242,5556	—305,6625

Таблица 8.2

n	$\eta_1^{(n)}$	$\eta_2^{(n)}$	$\eta_3^{(n)}$	$\eta_4^{(n)}$	$\eta_5^{(n)}$	$\eta_6^{(n)}$	$\eta_7^{(n)}$	$\eta_8^{(n)}$
1	0,0000							
2	0,0003	0,0003						
3	0,0014	0,0035	0,0012					
4	0,0052	0,0261	0,0398	0,0188				
5	0,0197	0,1713	0,4753	0,5269	0,2032			
6	0,0702	0,9285	4,0089	7,5942	6,5638	2,1202		
7	0,2210	4,0949	25,1183	70,9224	101,5560	71,7812	19,9031	
8	0,9335	24,0448	204,6676	819,0443	1747,4242	2046,8230	1242,5481	305,6612

с меньшей точностью), то погрешности в решении системы (8.1) возросли бы еще больше.

Создается парадоксальное положение. Чтобы увеличить точность приближенного решения по Ритцу, необходимо увеличить число n координатных элементов в приближенном решении — это следует из того, что энергетическая норма $|u_0 - u_n|_A$, где u_0 — точное решение вариационной задачи, а u_n — n-е приближенное по Ритцу решение той же задачи, монотонно убывает с возрастанием n (см. ВМ, стр. 94). Но при увеличении n могут резко возрасти погрешности при решении системы Ритца, а эти погрешности могут уменьшить или вовсе уничтожить полезный эффект от увеличения числа координатных элементов.

Такие неприятные явления обнаруживаются, однако, не всегда. Простой и довольно важный случай, когда такое резкое возрастание погрешностей не происходит, — это случай координатной системы, ортонормированной в энергетической метрике. Действительно, если $[\varphi_k, \varphi_j]_A = \delta_{kj}$, то система Ритца принимает вид:

$$a_j = (f, \varphi_j) \qquad (j = 1, 2, \ldots, n),$$

погрешность каждого из коэффициентов Ритца равна погрешности свободного члена с тем же номером; при фиксированном номере j погрешность коэффициента $a_j^{(n)}$ очевидным образом не зависит от n, поскольку самый коэффициент $a_j^{(n)} = a_j = (f, \varphi_j)$ от n не зависит.

Но координатная система не должна быть обязательно ортонормированной по энергии для того, чтобы погрешности не нарастали при возрастании n. Чтобы пояснить это утверждение, рассмотрим следующий пример.

Оператор задачи

$$Au = -\frac{d}{dx}\left[(2+x)\frac{du}{dx}\right] = 1, \qquad u(-1) = u(1) = 0, \qquad (8.2)$$

положительно определен в пространстве $L_2(-1, 1)$. Действительно,

$$(Au, u) = -\int_{-1}^{1} \overline{u(x)}\,\frac{d}{dx}\left[(2+x)\frac{du}{dx}\right]dx = -(2+x)\,\overline{u(x)}\,\frac{du}{dx}\Big|_{-1}^{1} +$$

$$+\int_{-1}^{1}(2+x)\left|\frac{du}{dx}\right|^2 dx = \int_{-1}^{1}(2+x)\left|\frac{du}{dx}\right|^2 dx \geqslant \int_{-1}^{1}\left|\frac{du}{dx}\right|^2 dx. \qquad (8.3)$$

Введем в рассмотрение оператор B, определяемый формулами

$$Bu = -\frac{d^2u}{dx^2}, \qquad u(-1) = u(1) = 0. \qquad (8.4)$$

Его энергетическое произведение и норма таковы:

$$[u, v]_B = \int_{-1}^{1} \frac{du}{dx} \frac{d\bar{v}}{dx}\, dx, \qquad |u|_B^2 = \int_{-1}^{1} \left| \frac{du}{dx} \right|^2 dx. \tag{8.5}$$

Оператор B имеет дискретный спектр; наименьшее собственное число этого оператора, как это легко вычислить, равно $\frac{\pi^2}{4}$, поэтому

$$|u|_B^2 = \int_{-1}^{1} \left| \frac{du}{dx} \right|^2 dx \geqslant \frac{\pi^2}{4} \| u \|^2.$$

Теперь из неравенства (8.3) следует:

$$(Au, u) \geqslant \frac{\pi^2}{4} \| u \|^2, \tag{8.6}$$

и положительная определенность оператора A доказана.

Энергетическое произведение и энергетическая норма для оператора A определяются формулами

$$[u, v]_A = \int_{-1}^{1} (2+x) \frac{du}{dx} \frac{d\bar{v}}{dx}\, dx, \quad |u|_A^2 = \int_{-1}^{1} (2+x) \left| \frac{du}{dx} \right|^2 dx. \tag{8.7}$$

Отсюда и из формулы (8.5) вытекает неравенство

$$|u|_B \leqslant |u|_A \leqslant \sqrt{3}\, |u|_B, \tag{8.8}$$

которое показывает, что операторы A и B — полусходные, нетрудно, впрочем, убедиться, что они на самом деле сходные.

Раз A — положительно определенный оператор, то задачу (8.2) можно решать процессом Ритца. В качестве координатных функций можно взять интегралы от нормированных полиномов Лежандра:

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}} \int_{-1}^{x} P_n(t)\, dt \qquad (n = 1, 2, \ldots). \tag{8.9}$$

Действительно, эти функции удовлетворяют краевым условиям задачи: очевидно, что $\varphi_n(-1) = 0$; кроме того,

$$\varphi_n(1) = \sqrt{\frac{2n+1}{2}} \int_{-1}^{1} P_n(t)\, dt = 0,$$

так как полиномы Лежандра $P_n(t)$ при $n > 0$ ортогональны в метрике $L_2(-1, 1)$ к полиному Лежандра нулевого порядка $P_0(t) \equiv 1$.

Теперь ясно, что функции (8.9) принадлежат энергетическому пространству H_A. Линейная независимость функций $\varphi_1(x), \varphi_2(x), \ldots$ $\ldots, \varphi_n(x)$ при любом n также очевидна. Докажем полноту системы (8.9) в метрике (8.7).

Неравенство (8.8) означает, что метрики (8.7) и (8.5) эквивалентны, и полноту системы (8.9) достаточно доказать в метрике (8.5).

Энергетическое пространство H_A задачи (8.2) состоит из функций, каждая из которых: 1) абсолютно непрерывна на отрезке $-1 \leqslant x \leqslant 1$; 2) обращается в нуль на концах этого отрезка; 3) имеет первую производную, на том же отрезке суммируемую с квадратом. Если функция $u(x)$ принадлежит пространству H_A, то производная $u'(x)$ удовлетворяет очевидному соотношению

$$\int_{-1}^{1} u'(x)\,dx = u(1) - u(-1) = 0,$$

и множество производных от функций из пространства H_A можно рассматривать как ортогональное к единице подпространство $\tilde{L}_2(-1, 1)$ пространства $L_2(-1, 1)$.

В этом подпространстве полна ортогональная к единице система полиномов

$$\varphi_n'(x) = \sqrt{\frac{2n+1}{2}}\,P_n(x) \qquad (n = 1, 2, \ldots). \tag{8.10}$$

Далее, из формул (8.5) следует, что

$$[u, v]_B = (u', v'),$$

и полнота системы (8.9) в метрике (8.5) сразу вытекает из отмеченной только что полноты системы (8.10) в подпространстве $\tilde{L}_2(-1, 1)$.

Система полиномов (8.9) в метрике (8.7) не ортонормирована — она ортонормирована в метрике (8.5); из неравенства (8.8) и из следствия 4.3 вытекает, что система (8.9) почти ортонормирована в метрике (8.7).

Вычислим матрицу системы Ритца и столбец ее свободных членов. По формуле (8.7)

$$[\varphi_k, \varphi_m] = \frac{1}{2}\sqrt{(2k+1)(2m+1)} \int_{-1}^{1} (x+2) P_k(x) P_m(x)\,dx =$$

$$= \frac{1}{2}\sqrt{(2k+1)(2m+1)} \left[\int_{-1}^{1} x P_k(x) P_m(x)\,dx + \frac{4}{2k+1}\,\delta_{km} \right].$$

Далее, по рекуррентной формуле для полиномов Лежандра

$$xP_k(x) = \frac{k+1}{2k+1} P_{k+1}(x) + \frac{k}{2k+1} P_{k-1}(x)$$

и, следовательно,

$$[\varphi_k, \varphi_m] = \frac{1}{2} \sqrt{(2k+1)(2m+1)} \left[\frac{2(k+1)}{(2k+1)(2k+3)} \delta_{k+1, m} + \right.$$
$$\left. + \frac{2k}{(2k+1)(2k-1)} \delta_{k-1, m} + \frac{4}{2k+1} \delta_{km} \right].$$

Таким образом, при $k \geqslant m$

$$[\varphi_k, \varphi_m] = \begin{cases} 0, & k-m > 1, \\ 2, & k = m, \\ \dfrac{k}{\sqrt{4k^2-1}}, & k = m+1, \end{cases}$$

при $k < m$ энергетическое произведение вычисляется по формуле $[\varphi_k, \varphi_m] = [\varphi_m, \varphi_k]$. Матрица Ритца имеет вид:

$$\left\| \begin{matrix} 2 & \frac{2}{\sqrt{15}} & 0 & 0 & 0 & \dots \\ \frac{2}{\sqrt{15}} & 2 & \frac{3}{\sqrt{35}} & 0 & 0 & \dots \\ 0 & \frac{3}{\sqrt{35}} & 2 & \frac{4}{3\sqrt{7}} & 0 & \dots \\ 0 & 0 & \frac{4}{3\sqrt{7}} & 2 & \frac{5}{3\sqrt{11}} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{matrix} \right\|$$

Для свободных членов системы Ритца имеем:

$$(1, \varphi_k) = \int_{-1}^{1} \varphi_k(x)\, dx = x\varphi_k(x) \Big|_{-1}^{1} - \int_{-1}^{1} x\varphi_k'(x) =$$
$$= -\sqrt{\frac{2k+1}{2}} \int_{-1}^{1} xP_k(x)\, dx$$

и, следовательно,

$$(1, \varphi_k) = \begin{cases} -\sqrt{\dfrac{2}{3}}, & k = 1, \\ 0, & k > 1. \end{cases}$$

Ограничимся системой Ритца восьмого порядка:

$$\left.\begin{aligned}
2a_1^{(8)}+\frac{2}{\sqrt{15}}a_2^{(8)}&=-\sqrt{\frac{2}{3}},\\
\frac{2}{\sqrt{15}}a_1^{(8)}+2a_2^{(8)}+\frac{3}{\sqrt{35}}a_3^{(8)}&=0,\\
\frac{3}{\sqrt{35}}a_2^{(8)}+2a_3^{(8)}+\frac{4}{3\sqrt{7}}a_4^{(8)}&=0,\\
\frac{4}{3\sqrt{7}}a_3^{(8)}+2a_4^{(8)}+\frac{5}{3\sqrt{11}}a_5^{(8)}&=0,\\
\frac{5}{3\sqrt{11}}a_4^{(8)}+2a_5^{(8)}+\frac{6}{\sqrt{143}}a_6^{(8)}&=0,\\
\frac{6}{\sqrt{143}}a_5^{(8)}+2a_6^{(8)}+\frac{7}{\sqrt{195}}a_7^{(8)}&=0,\\
\frac{7}{\sqrt{195}}a_6^{(8)}+2a_7^{(8)}+\frac{8}{\sqrt{255}}a_8^{(8)}&=0,\\
\frac{8}{\sqrt{255}}a_7^{(8)}+2a_8^{(8)}&=0;
\end{aligned}\right\}\qquad(8.11)$$

системы Ритца меньших порядков получаются отсюда усечением.

В табл. 8.3 приведены точные значения коэффициентов Ритца $a_k^{(n)}$ для $1\leqslant n\leqslant 8$; в табл. 8.4 приведены значения тех же коэффициентов с 4-мя верными знаками после запятой.

Отметим, что, в отличие от примеров § 7, здесь явление «разброса» коэффициентов $a_k^{(n)}$ не наблюдается: при увеличении n выписанные в таблице десятичные знаки быстро стабилизируются, и притом равномерно относительно индекса k. Это объясняется тем, что в данном случае координатная система, будучи почти ортогональной в энергетической метрике, тем более сильно минимальна в ней.

От точной системы Ритца (8.11) перейдем теперь к неточной, заменив ее коэффициенты и свободные члены их десятичными приближениями с четырьмя знаками после запятой:

$$\left.\begin{aligned}
2b_1^{(8)}+0{,}5164_2^{(8)}&=-0{,}8165,\\
0{,}5164b_1^{(8)}+2b_2^{(8)}+0{,}5071b_3^{(8)}&=0,\\
0{,}5071b_2^{(8)}+2b_3^{(8)}+0{,}5040b_4^{(8)}&=0,\\
0{,}5040b_3^{(8)}+2b_4^{(8)}+0{,}5025b_5^{(8)}&=0,\\
0{,}5025b_4^{(8)}+2b_5^{(8)}+0{,}5017b_6^{(8)}&=0,\\
0{,}5017b_5^{(8)}+2b_6^{(8)}+0{,}5013b_7^{(8)}&=0,\\
0{,}5013b_6^{(8)}+2b_7^{(8)}+0{,}5010b_8^{(8)}&=0,\\
0{,}5010b_7^{(8)}+2b_8^{(8)}&=0.
\end{aligned}\right\}\qquad(8.12)$$

Таблица 8.3

n	$\frac{a_1^{(n)}}{\sqrt{6}}$	$\frac{a_2^{(n)}}{\sqrt{10}}$	$\frac{a_3^{(n)}}{\sqrt{14}}$	$\frac{a_4^{(n)}}{\sqrt{2}}$	$\frac{a_5^{(n)}}{\sqrt{22}}$	$\frac{a_6^{(n)}}{\sqrt{26}}$	$\frac{a_7^{(n)}}{\sqrt{30}}$	$\frac{a_8^{(n)}}{\sqrt{34}}$
1	$-\frac{1}{6}$							
2	$-\frac{5}{28}$	$\frac{1}{28}$						
3	$-\frac{131}{730}$	$\frac{14}{365}$	$-\frac{3}{365}$					
4	$-\frac{1\,099}{6\,122}$	$\frac{118}{3\,061}$	$-\frac{27}{3\,061}$	$\frac{18}{3\,061}$				
5	$-\frac{5\,009}{27\,902}$	$\frac{538}{13\,951}$	$-\frac{53}{5\,979}$	$\frac{88}{13\,951}$	$-\frac{20}{41\,853}$			
6	$-\frac{60\,721}{338\,238}$	$\frac{2\,174}{56\,373}$	$-\frac{4\,499}{507\,357}$	$\frac{1\,072}{169\,119}$	$-\frac{260}{507\,357}$	$\frac{20}{169\,119}$		
7	$-\frac{10\,193\,457}{56\,781\,246}$	$\frac{1\,094\,874}{28\,390\,623}$	$-\frac{251\,761}{28\,390\,623}$	$\frac{60\,008}{9\,463\,541}$	$-\frac{14\,620}{28\,390\,623}$	$\frac{1\,200}{9\,463\,541}$	$-\frac{280}{9\,463\,541}$	
8	$-\frac{53\,876\,779}{300\,113\,162}$	$\frac{5\,786\,878}{150\,056\,581}$	$-\frac{1\,330\,667}{150\,056\,581}$	$\frac{951\,528}{150\,056\,581}$	$-\frac{77\,300}{150\,056\,581}$	$\frac{19\,120}{150\,056\,581}$	$-\frac{4\,760}{150\,056\,581}$	$\frac{1\,120}{150\,056\,581}$

Таблица 8.4

n	$a_1^{(n)}$	$a_2^{(n)}$	$a_3^{(n)}$	$a_4^{(n)}$	$a_5^{(n)}$	$a_6^{(n)}$	$a_7^{(n)}$	$a_8^{(n)}$
1	—0,4082							
2	—0,4374	0,1129						
3	—0,4395	0,1213	—0,0308					
4	—0,4397	0,1219	—0,0330	0,0083				
5	—0,4397	0,1219	—0,0332	0,0089	—0,0022			
6	—0,4397	0,1220	—0,0332	0,0090	—0,0024	0,0006		
7	—0,4397	0,1220	—0,0332	0,0090	—0,0024	0,0006	—0,0002	
8	—0,4397	0,1220	—0,0332	0,0090	—0,0024	0,0006	—0,0002	0,0001

Чтобы по возможности избежать погрешностей округления, промежуточные вычисления производились с достаточно большим числом десятичных знаков — с семью знаками для систем порядка $\leqslant 4$ и с 12 знаками для систем порядков от 5 до 8. В окончательных результатах, приведенных в табл. 8.5, сохранены четыре знака после запятой.

Таблица 8.5

n	$b_1^{(n)}$	$b_2^{(n)}$	$b_3^{(n)}$	$b_4^{(n)}$	$b_5^{(n)}$	$b_6^{(n)}$	$b_7^{(n)}$	$b_8^{(n)}$
1	—0,4083							
2	—0,4374	0,1129						
3	—0,4396	0,1213	—0,0308					
4	—0,4397	0,1219	—0,0330	0,0083				
5	—0,4398	0,1220	—0,0332	0,0089	—0,0022			
6	—0,4398	0,1220	—0,0332	0,0089	—0,0024	0,0006		
7	—0,4398	0,1220	—0,0332	0,0089	—0,0024	0,0007	—0,0002	
8	—0,4398	0,1220	—0,0332	0,0089	—0,0024	0,0007	—0,0002	0,0001

Сравнивая таблицы 8.4 и 8.5, видим, что погрешность в решениях системы Ритца не превосходит единицы четвертого десятичного знака и, следовательно, не более, чем вдвое, превосходит погрешность в коэффициентах и свободных членах системы Ритца.

§ 9. Об устойчивости процесса Ритца[1])

Будем рассматривать задачу о минимуме функционала

$$F(u) = A[u] - lu - \overline{lu}, \tag{9.1}$$

[1]) См. статьи автора [15] и [17].

где $A[u]$ — положительный квадратичный функционал, lu — линейный функционал. Как всегда, введем энергетическое пространство H_A с энергетическим произведением и нормой

$$[u, v] = [u, v]_A = A[u, v], \qquad |u|^2 = |u|_A^2 = A[u]. \tag{9.2}$$

Здесь $A[u, v]$ — билинейный функционал, соответствующий квадратичному функционалу $A[u]$. Будем считать, что функционал lu ограничен в метрике (9.2); пусть известно, что

$$|l|_A \leqslant C_0 = \text{const}. \tag{9.3}$$

Тогда существует элемент $u_0 \in H_A$, реализующий минимум функционала $F(u)$ в энергетическом пространстве; этот элемент определяется формулой[1]) $lu = [u, u_0]$, вытекающей из теоремы Ф. Риса об общем виде ограниченного линейного функционала в гильбертовом пространстве. Заметим, что при этом

$$|l|_A = |u_0|_A$$

и, следовательно,

$$|u_0|_A \leqslant C. \tag{9.4}$$

Будем строить приближенное решение нашей вариационной задачи с помощью процесса Ритца: выберем координатную последовательность $\{\varphi_n\}$ $(n = 1, 2, \ldots)$, удовлетворяющую трем условиям § 6, и примем за приближенное решение по Ритцу элемент, задаваемый формулой (6.6), в которой коэффициенты $a_k^{(n)}$ определяются из системы Ритца (6.8).

Как мы видели в предшествующем параграфе, в разных случаях решения системы Ритца по-разному реагируют на малые погрешности, которые мы допускаем при составлении этой системы, а именно, при вычислении ее матрицы и столбца свободных членов: в одних случаях погрешность решения быстро возрастает с ростом порядка системы, в других случаях погрешность решения такой тенденции не проявляет. Эти соображения приводят к понятию об устойчивости процесса Ритца; к формулировке этого понятия мы и переходим.

Обозначим через R_n матрицу Ритца n-го порядка:

$$R_n = \begin{pmatrix} [\varphi_1, \varphi_1]_A & [\varphi_2, \varphi_1]_A & \ldots & [\varphi_n, \varphi_1]_A \\ [\varphi_1, \varphi_2]_A & [\varphi_2, \varphi_2]_A & \ldots & [\varphi_n, \varphi_2]_A \\ \cdot & \cdot & \cdot & \cdot \\ [\varphi_1, \varphi_n]_A & [\varphi_2, \varphi_n]_A & \ldots & [\varphi_n, \varphi_n]_A \end{pmatrix},$$

через $f^{(n)}$ — столбец свободных членов системы Ритца:

$$f^{(n)} = (l\varphi_1, l\varphi_2, \ldots, l\varphi_n)',$$

[1]) См. § 6.

через $a^{(n)}$ обозначим столбец, составленный из коэффициентов Ритца:

$$a^{(n)} = \left(a_1^{(n)},\ a_2^{(n)},\ \ldots,\ a_n^{(n)}\right)'.$$

В этих обозначениях система Ритца записывается так:

$$R_n a^{(n)} = f^{(n)}. \tag{9.5}$$

Допустим, что энергетические произведения $[\varphi_k, \varphi_m]$ вычислены с малыми погрешностями $\gamma_{km} = \gamma_{mk}$, а скалярные произведения $f_m = (f, \varphi_m)$ — с малыми же погрешностями δ_m. Обозначим через Γ_n матрицу n-го порядка с элементами γ_{jk} $(j, k = 1, 2, \ldots, n)$ и через $\delta^{(n)}$ — вектор-столбец с составляющими $\delta_1, \delta_2, \ldots, \delta_n$. Вместо системы (9.5) мы получим систему

$$(R_n + \Gamma_n)\, b^{(n)} = f^{(n)} + \delta^{(n)}, \tag{9.6}$$

где $b^{(n)}$ — вектор «неточных» значений коэффициентов Ритца.

Матрицы R_n, R_n^{-1}, Γ_n порождают в n-мерном унитарном пространстве линейные операторы, которые мы будем обозначать теми же символами R_n, R_n^{-1}, Γ_n; символами $\|R_n\|$, $\|R_n^{-1}\|$, $\|\Gamma_n\|$ будем обозначать нормы этих операторов. Заметим, что

$$\|R_n^{-1}\| = \frac{1}{\lambda_1^{(n)}}, \tag{9.7}$$

где $\lambda_1^{(n)}$ — наименьшее собственное число матрицы R_n; отметим также, что

$$\|\Gamma_n\| \leqslant \left\{ \sum_{j,\ k=1}^{n} |\gamma_{jk}|^2 \right\}^{\frac{1}{2}}. \tag{9.8}$$

Погрешности γ_{jk} и δ_j иногда обусловлены тем, что величины $[\varphi_j, \varphi_k]$ и $l\varphi_j$ заменяются их десятичными приближениями с определенным числом знаков после запятой. Пусть при этом сохраняется k знаков после запятой, тогда (ограничиваемся случаем вещественного гильбертова пространства) $|\gamma_{jk}| < \frac{1}{2} \cdot 10^{-k}$ и $|\delta_j| < \frac{1}{2} \cdot 10^{-k}$. Отсюда

$$\|\Gamma_n\| < \frac{n}{2} \cdot 10^{-k}$$

и

$$\|\delta_n\| = \left\{ \sum_{j=1}^{n} \delta_j^2 \right\}^{\frac{1}{2}} < \frac{\sqrt{n}}{2} \cdot 10^{-k}.$$

Будем говорить, что процесс Ритца устойчив, если существуют такие независящие от n постоянные p, q, r, что при $\|\Gamma_n\| \leqslant r$ и любых $\delta^{(n)}$ система (9.6) разрешима и справедливо неравенство

$$\|b^{(n)} - a^{(n)}\| \leqslant p\|\Gamma_n\| + q\|\delta^{(n)}\|. \tag{9.9}$$

В противном случае будем говорить, что процесс Ритца неустойчив.

Это определение предполагает, что как «точная» система Ритца (9.5), так и «неточная» система (9.6) решены идеально точно, без погрешности. Тем самым наше определение устойчивости не учитывает ошибок округления, возникающих при решении системы.

Устойчивость или неустойчивость процесса Ритца полностью определяется свойствами координатной системы, как это видно из следующей теоремы.

Теорема 9.1. *Для того чтобы процесс Ритца был устойчив, необходимо и достаточно, чтобы координатная система была сильно минимальна в соответствующем энергетическом пространстве.*

Необходимость. Пусть процесс Ритца устойчив. Допустим, что $\Gamma_n \equiv 0$, так что при составлении системы Ритца погрешность была допущена только при вычислении свободных членов. Тогда

$$\|b^{(n)} - a^{(n)}\| \leqslant q\|\delta^{(n)}\|. \tag{9.10}$$

Пусть $\lambda_1^{(n)}$ означает наименьшее собственное число матрицы R_n, а $x_1^{(n)}$ — соответствующий ему нормированный собственный вектор.

В уравнении (9.6) положим $\delta^{(n)} = x_1^{(n)}$. Уравнение (9.6) принимает вид:

$$R_n b^{(n)} = f^{(n)} + x_1^{(n)};$$

отсюда $b^{(n)} = a^{(n)} + \dfrac{1}{\lambda_1^{(n)}} x_1^{(n)}$ и

$$\|b^{(n)} - a^{(n)}\| = \frac{1}{\lambda_1^{(n)}} = \frac{1}{\lambda_1^{(n)}}\|\delta^{(n)}\|.$$

Теперь из неравенства (9.10) следует, что $\lambda_1^{(n)} \geqslant \dfrac{1}{q}$, и координатная система оказывается сильно минимальной.

Достаточность. Пусть координатная система сильно минимальна, так что $\lambda_1^{(n)} \geqslant \lambda_0 = \mathrm{const} > 0$. Положим $r = \beta\lambda_0$, где β — постоянная, $0 < \beta < 1$. Будем считать, что $\|\Gamma_n\| \leqslant \beta\lambda_0$. Тогда система (9.6) однозначно разрешима.

Из уравнений (9.5) и (9.6) имеем:

$$a^{(n)} = R_n^{-1} f^{(n)}, \quad b^{(n)} = (R_n + \Gamma_n)^{-1}(f^{(n)} + \delta^{(n)}).$$

Но

$$(R_n + \Gamma_n)^{-1} = \left[R_n\left(I_n + R_n^{-1}\Gamma_n\right)\right]^{-1} = \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} R_n^{-1},$$

где I_n — единичная матрица порядка n. Отсюда

$$b^{(n)} = \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} R_n^{-1}\left(f^{(n)} + \delta^{(n)}\right) = \\ = \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} a^{(n)} + \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} R_n^{-1}\delta^{(n)}$$

и, следовательно,

$$\left\| b^{(n)} - a^{(n)} \right\| \leqslant \left\| \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} - I_n \right\| \left\| a^{(n)} \right\| + \\ + \left\| \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} \right\| \left\| R_n^{-1} \right\| \left\| \delta^{(n)} \right\|.$$

Имеем:

$$\left\| R_n^{-1}\Gamma_n \right\| \leqslant \left\| R_n^{-1} \right\| \left\| \Gamma_n \right\| = \frac{\|\Gamma_n\|}{\lambda_1^{(n)}} \leqslant \frac{\|\Gamma_n\|}{\lambda_0} \leqslant \beta < 1,$$

поэтому матрицу $\left(I_n + R_n^{-1}\Gamma_n\right)^{-1}$ можно разложить в степенной ряд:

$$\left(I_n + R_n^{-1}\Gamma_n\right)^{-1} = I_n - R_n^{-1}\Gamma_n + \left(R_n^{-1}\Gamma_n\right)^2 - \ldots$$

и

$$\left(I_n + R_n^{-1}\Gamma_n\right)^{-1} - I_n = -R_n^{-1}\Gamma_n + \left(R_n^{-1}\Gamma_n\right)^2 - \ldots$$

Отсюда

$$\left\| \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} \right\| \leqslant 1 + \left\| R_n^{-1}\Gamma_n \right\| + \left\| \left(R_n^{-1}\Gamma_n\right)^2 \right\| + \ldots \leqslant \\ \leqslant 1 + \frac{\|\Gamma_n\|}{\lambda_0} + \frac{\|\Gamma_n\|^2}{\lambda_0^2} + \ldots = \frac{1}{1 - \lambda_0^{-1}\|\Gamma_n\|} \leqslant \frac{1}{1-\beta}$$

и

$$\left\| \left(I_n + R_n^{-1}\Gamma_n\right)^{-1} - I_n \right\| \leqslant \left\| R_n^{-1}\Gamma_n \right\| + \left\| \left(R_n^{-1}\Gamma_n\right)^2 \right\| + \ldots \leqslant \\ \leqslant \frac{\lambda_0^{-1}\|\Gamma_n\|}{1 - \lambda_0^{-1}\|\Gamma_n\|} \leqslant \frac{\lambda_0^{-1}\|\Gamma_n\|}{1-\beta}.$$

Теперь

$$\left\| b^{(n)} - a^{(n)} \right\| \leqslant \frac{\lambda_0^{-1}\|\Gamma_n\| \left\| a^{(n)} \right\| + \lambda_0^{-1} \left\| \delta^{(n)} \right\|}{1-\beta}. \tag{9.11}$$

Оценим величину $\left\| a^{(n)} \right\|$. Имеем

$$|u_n|^2 = \sum_{j,\,k=1}^{n} [\varphi_k, \varphi_j]\, a_k^{(n)} a_j^{(n)} \geqslant \lambda_1^{(n)} \sum_{k=1}^{n} \left| a_k^{(n)} \right|^2 = \\ = \lambda_1^{(n)} \left\| a^{(n)} \right\|^2 \geqslant \lambda_0 \left\| a^{(n)} \right\|^2.$$

Отсюда

$$\left\| a^{(n)} \right\| \leqslant \lambda_0^{-\frac{1}{2}} |u_n|.$$

Но $|u_n| \leqslant |u_0|$ (см. ВМ, стр. 94, формула (15)), поэтому

$$\| a^{(n)} \| \leqslant \lambda_0^{-\frac{1}{2}} |u_0|. \tag{9.12}$$

Подставив это в неравенство (9.11), получим:

$$\| b^{(n)} - a^{(n)} \| \leqslant \frac{\lambda_0^{-\frac{3}{2}} |u_0| \, \| \Gamma_n \| + \lambda_0^{-1} \| \delta^{(n)} \|}{1-\beta}, \tag{9.13}$$

что совпадает с неравенством (9.9) при

$$p = \frac{\lambda_0^{-\frac{3}{2}} |u_0|}{1-\beta}, \qquad q = \frac{\lambda_0^{-1}}{1-\beta}.$$

Замечание. Величина $|u_0|$ обычно неизвестна; для фактической оценки погрешности можно пользоваться формулой

$$\| b^{(n)} - a^{(n)} \| \leqslant \frac{\lambda_0^{-\frac{3}{2}} C \| \Gamma_n \| + \lambda_0^{-1} \| \delta^{(n)} \|}{1-\beta}, \tag{9.13'}$$

вытекающей из неравенств (9.13) и (9.4).

Как это видно из хода доказательства, оценки (9.13) и (9.13′) бесспорно верны только в том случае, когда обе системы Ритца — и «точная» система (9.5), и «неточная» система (9.6) — решены абсолютно точно, без ошибок округления. Если же такие ошибки были допущены, то норма погрешности $b^{(n)} - a^{(n)}$ может превзойти правую часть каждой из формул (9.13) или (9.13′).

Применим полученные результаты ко второму примеру § 8. Использованная в этом примере координатная система (8.9), как это уже было отмечено, в соответствующем энергетическом пространстве почти ортонормирована; тем более она в нем сильно минимальна. Найдем границу λ_0, для чего воспользуемся некоторыми результатами, попутно полученными при доказательстве теоремы 4.2. Система (8.9) ортонормирована и, следовательно, сильно минимальна в H_B, где оператор B определен формулой (8.4). Отождествим пространства H_A и H_B соответственно с пространствами $\mathfrak{H}_1$ и $\mathfrak{H}_2$ § 4. Система (8.9) ортонормирована в $\mathfrak{H}_2$, поэтому все собственные числа матрицы r_n (формула (4.5)) равны единице, и можно положить $\mu_0 = 1$, где μ_0 — нижняя граница собственных чисел матрицы r_n. Далее, неравенство (4.1) в нашем случае имеет вид (см. формулу (8.8)):

$$\| u \|_2 \leqslant \| u \|_1,$$

так что $K = 1$. Оценка (4.6) дает теперь $\lambda_1^{(n)} \geqslant 1$ и можно, следовательно, принять $\lambda_0 = 1$. При переходе от точной системы Ритца (8.11) к неточной системе (8.12) ошибка в каждом элементе матрицы или свободных членов не превосходит $\frac{1}{2}\cdot 10^{-4}$; таким образом, $|\gamma_{jk}| \leqslant \frac{1}{2}\cdot 10^{-4}$ и $|\delta_k| \leqslant \frac{1}{2}\cdot 10^{-4}$. Отсюда

$$\|\Gamma_8\| \leqslant 4\cdot 10^{-4}, \qquad \|\delta^{(8)}\| \leqslant \sqrt{2}\cdot 10^{-4},$$

и можно принять $\beta = 5\cdot 10^{-4}$.

Для оценки величины $C \geqslant |u_0|$, входящей в формулу (9.13′), воспользуемся неравенством $|u_0| \leqslant \frac{\|f\|}{\gamma}$, где f — свободный член уравнения, а γ — постоянная неравенства $(Au, u) \geqslant \gamma^2 \|u\|^2$, характеризующего положительную определенность оператора данной задачи. Как показывает формула (8.7), можно принять $\gamma = \frac{\pi}{2}$. Далее, $f(x) \equiv 1$ и, следовательно,

$$\|f\| = \left\{\int_{-1}^{1} dx\right\}^{\frac{1}{2}} = \sqrt{2},$$

можно поэтому положить $C = \frac{\|f\|}{\gamma} = \frac{2\sqrt{2}}{\pi}$. Для упрощения счета заменим величину $\frac{2\sqrt{2}}{\pi}$ несколько большей величиной $C = 1$. Теперь по формуле (9.13′) получаем оценку для нормы вектора погрешности:

$$\|b^{(8)} - a^{(8)}\| \leqslant \frac{1\cdot 1\cdot 4\cdot 10^{-4} + \sqrt{2}\cdot 10^{-4}}{1 - 5\cdot 10^{-4}} \approx 5{,}4\cdot 10^{-4}.$$

Сравнение таблиц 8.4 и 8.5 дает следующее значение этой нормы:

$$\|b^{(8)} - a^{(8)}\| = \sqrt{3}\cdot 10^{-4} < 1{,}8\cdot 10^{-4},$$

что меньше только что полученной теоретической верхней границы нормы, но совпадает с этой границей по порядку.

§ 10. Об устойчивости приближенного решения

Решение $b^{(n)}$ «неточной» системы Ритца (9.6) приводит к «неточному» приближенному решению по Ритцу

$$v_n = \sum_{k=1}^{n} b_k^{(n)} \varphi_k. \tag{10.1}$$

Приближенное решение по Ритцу

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k, \tag{10.2}$$

где $a^{(n)}=\left(a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}\right)$ есть решение точной системы Ритца (системы (9.5)), назовем устойчивым, если существуют независящие от n постоянные p_1, q_1, r_1 такие, что при $\|\Gamma_n\| \leqslant r_1$ и любых $\delta^{(n)}$ справедливо неравенство

$$|u_n - v_n| \leqslant p_1 \|\Gamma_n\| + q_1 \|\delta^{(n)}\|. \tag{10.3}$$

Теорема 10.1. *Для того чтобы приближенное решение по Ритцу было устойчиво, необходимо и достаточно, чтобы координатная система была сильно минимальной в соответствующем энергетическом пространстве.*

Необходимость. Пусть приближенное решение по Ритцу устойчиво. Положим $\Gamma_n=0$, тогда

$$|v_n - u_n| \leqslant q_1 \|\delta^{(n)}\|, \tag{10.4}$$

и это неравенство должно иметь место при любом векторе погрешностей $\delta^{(n)}$. Возьмем $\delta^{(n)}=x_1^{(n)}$; смысл обозначений тот же, что и в теореме 9.1. В этом случае, как мы видели в § 9, $b^{(n)}=a^{(n)}+\frac{1}{\lambda_1^{(n)}} x_1^{(n)}$.

Но тогда

$$v_n = u_n + \frac{1}{\lambda_1^{(n)}} \sum_{k=1}^{n} x_{1k}^{(n)} \varphi_k,$$

где $x_{1k}^{(n)}$ $(k=1, 2, \ldots, n)$ — составляющие вектора $x_1^{(n)}$. Отсюда

$$|v_n - u_n| = \frac{1}{\lambda_1^{(n)}} \left| \sum_{k=1}^{n} x_{1k}^{(n)} \varphi_k \right| = \frac{\|\delta^{(n)}\|}{\lambda_1^{(n)}} \left| \sum_{k=1}^{n} x_{1k}^{(n)} \varphi_k \right|.$$

Далее,

$$\left| \sum_{k=1}^{n} x_{1k}^{(n)} \varphi_k \right|^2 = \left[\sum_{k=1}^{n} x_{1k}^{(n)} \varphi_k, \sum_{j=1}^{n} x_{1j}^{(n)} \varphi_j \right] = \sum_{j,k=1}^{n} [\varphi_k, \varphi_j]\, x_{1k}^{(n)} \overline{x_{1j}^{(n)}} =$$
$$= \left(R_n x_1^{(n)}, x_1^{(n)}\right) = \left(\lambda_1^{(n)} x_1^{(n)}, x_1^{(n)}\right) = \lambda_1^{(n)}$$

и, следовательно,

$$|v_n - u_n| = \frac{\|\delta^{(n)}\|}{\sqrt{\lambda_1^{(n)}}}.$$

Теперь из неравенства (10.4) следует, что $\lambda_1^{(n)} \geqslant q_1^{-2}$, и координатная система оказывается сильно минимальной в энергетическом пространстве.

Достаточность[1]). Пусть координатная система сильно минимальна в энергетическом пространстве.

---

[1]) См. Г. Н. Яскова и М. Н. Яковлев [1].

Имеем:

$$v_n - u_n = \sum_{k=1}^{n} \left(b_k^{(n)} - a_k^{(n)}\right) \varphi_k.$$

Отсюда

$$|v_n - u_n|^2 = \left[\sum_{k=1}^{n} \left(b_k^{(n)} - a_k^{(n)}\right) \varphi_k, \ \sum_{j=1}^{n} \left(b_j^{(n)} - a_j^{(n)}\right) \varphi_j\right] =$$

$$= \sum_{j,\,k=1}^{n} [\varphi_k, \ \varphi_j] \left(b_k^{(n)} - a_k^{(n)}\right)\left(\overline{b}_j^{(n)} - \overline{a}_j^{(n)}\right) =$$

$$= \left(R_n\left(b^{(n)} - a^{(n)}\right), \ b^{(n)} - a^{(n)}\right) \leqslant \| R_n b^{(n)} - R_n a^{(n)} \| \, \| b^{(n)} - a^{(n)} \|. \quad (10.5)$$

Второй множитель справа оценивается формулой (9.13). Оценим первый множитель.

Из уравнений (9.5) и (9.6) следует:

$$R_n a^{(n)} = f^{(n)}, \quad R_n b^{(n)} = f^{(n)} + \delta^{(n)} - \Gamma_n b^{(n)}.$$

Отсюда

$$R_n b^{(n)} - R_n a^{(n)} = \delta^{(n)} - \Gamma_n b^{(n)} = \delta^{(n)} + \Gamma_n a^{(n)} - \Gamma_n \left(b^{(n)} - a^{(n)}\right) \quad (10.6)$$

или

$$(R_n + \Gamma_n)\left(b^{(n)} - a^{(n)}\right) = \Gamma_n a^{(n)} + \delta^{(n)}.$$

Последнее уравнение можно представить в виде:

$$\left(I_n + \Gamma_n R_n^{-1}\right) R_n \left(b^{(n)} - a^{(n)}\right) = \Gamma_n a^{(n)} + \delta^{(n)}.$$

Отсюда

$$R_n b^{(n)} - R_n a^{(n)} = \left(I_n + \Gamma_n R_n^{-1}\right)^{-1} \left(\Gamma_n a^{(n)} + \delta^{(n)}\right).$$

Потребуем, чтобы $\| \Gamma_n \| \leqslant \beta \lambda_0$, $0 < \beta < 1$, тогда:

$$\| \Gamma_n R_n^{-1} \| \leqslant \beta, \qquad \left\| \left(I_n + \Gamma_n R_n^{-1}\right)^{-1} \right\| \leqslant \frac{1}{1-\beta},$$

$$\left\| R_n \left(b^{(n)} - a^{(n)}\right) \right\| \leqslant \frac{1}{1-\beta} \left[\| \Gamma_n \| \, \| a^{(n)} \| + \| \delta^{(n)} \|\right].$$

Но по неравенству (9.12), $\| a^{(n)} \| \leqslant \lambda_0^{-\frac{1}{2}} |u_0|$ и, следовательно,

$$\left\| R_n \left(b^{(n)} - a^{(n)}\right) \right\| \leqslant \frac{1}{1-\beta} \left[\lambda_0^{-\frac{1}{2}} |u_0| \, \| \Gamma_n \| + \| \delta^{(n)} \|\right].$$

Теперь, по неравенствам (9.5) и (9.7),

$$|v_n - u_n|^2 \leqslant \frac{1}{1-\beta} \left[\lambda_0^{-\frac{1}{2}} |u_0| \, \| \Gamma_n \| + \| \delta^{(n)} \|\right] \left(p \| \Gamma_n \| + q \| \delta^{(n)} \|\right),$$

и если мы положим

$$p_1 = \max\left(p, \ \frac{\lambda_0^{-\frac{1}{2}}}{1-\beta}|u_0|\right), \quad q_1 = \max\left(q, \ \frac{1}{1-\beta}\right), \tag{10.7}$$

то получим неравенство

$$|v_n - u_n| \leqslant p_1 \|\Gamma_n\| + q_1 \|\delta^{(n)}\|,$$

которое и требовалось доказать.

Практически оценку погрешности $|v_n - u_n|$ можно производить по формуле, вытекающей из формул (10.5) и (10.6). Именно,

$$|v_n - u_n|^2 \leqslant \|b^{(n)} - a^{(n)}\| \, \|\delta^{(n)} - \Gamma_n b^{(n)}\|$$

или

$$|v_n - u_n| \leqslant \sqrt{\|b^{(n)} - a^{(n)}\| \left[\|\delta^{(n)}\| + \|\Gamma_n\| \, \|b^{(n)}\|\right]} =$$
$$= \sqrt{\|\eta^{(n)}\| \left[\|\delta^{(n)}\| + \|\Gamma_n\| \, \|b^{(n)}\|\right]}. \tag{10.8}$$

Применим эту формулу к примерам § 8.

В обоих примерах при переходе от «точной» системы Ритца к «неточной» были сохранены четыре знака после запятой, поэтому $|\gamma_{jk}| < \frac{1}{2} \cdot 10^{-4}$ и $|\delta_k| < \frac{1}{2} \cdot 10^{-4}$. Ограничиваясь системой Ритца 4-го порядка, в обоих случаях имеем:

$$\|\Gamma_4\| \leqslant 2 \cdot 10^{-4}, \qquad \|\delta^{(4)}\| \leqslant 10^{-4}.$$

С помощью таблиц 8.1 и 8 2 находим значения $\|b^{(4)}\|$ и $\|\eta^{(4)}\| = \|b^{(4)} - a^{(4)}\|$ для первого примера:

$$\|b^{(4)}\| = 0{,}8854, \qquad \|\eta^{(4)}\| = 0{,}0514.$$

Теперь по формуле (10.8)

$$|v_4 - u_4| \leqslant \sqrt{0{,}0514\left[10^{-4} + 2 \cdot 10^{-4} \cdot 0{,}8854\right]} =$$
$$= 0{,}00377 = 3{,}77 \cdot 10^{-3}.$$

Для второго примера с помощью табл. 8.5 находим значение

$$\|b^{(4)}\| = 0{,}4576.$$

Таблицы 8.4 и 8.5, в которых сохранены только 4 знака после запятой, дают значение $\|\eta^{(4)}\| = 0$. Используя более точные значения величин $a_k^{(n)}$ и $b_k^{(n)}$, можно получить значение $\|\eta^{(4)}\| = 2{,}9 \cdot 10^{-5}$. Теперь

$$|v_4 - u_4| \leqslant \sqrt{2{,}9 \cdot 10^{-5}\left[10^{-4} + 2 \cdot 10^{-4} \cdot 0{,}4576\right]} < 7{,}4 \cdot 10^{-5}.$$

Если бы мы исходили не из фактической величины $\|\eta^{(4)}\|$, а из ее теоретической оценки, даваемой формулой (9.13) или (9.13′), то мы получили бы:

$$|v_4 - u_4| \leqslant \sqrt{3 \cdot 10^{-4}\left[10^{-4} + 2 \cdot 0{,}4576 \cdot 10^{-4}\right]} \approx 2{,}4 \cdot 10^{-4}.$$

Как мы видим, применение сильно минимальной координатной системы во втором примере дало погрешность значительно меньшую, чем в первом примере, где координатная система была просто неминимальной.

Примечательно такое обстоятельство: во втором примере, т. е. в случае сильно минимальной координатной системы, оценки величин $\|b^{(n)} - a^{(n)}\|$ и $|v_n - u_n|$, т. е. норм погрешностей вектора коэффициентов Ритца и приближенного решения по Ритцу, имеют один и тот же порядок малости.

В то же время в первом примере порядок малости погрешности приближенного решения выше, чем порядок малости вектора коэффициентов [1]). Причина этого явления делается ясной из анализа формулы (10.8). Порядок малости второго множителя в подкоренном выражении определяется величинами $\|\delta^{(n)}\|$ и $\|\Gamma_n\|$. Если координатная система сильно минимальна в энергетическом пространстве, то первый множитель под корнем имеет тот же порядок малости, что и $\|\Gamma_n\|$ и $\|\delta^{(n)}\|$, а тогда и самый корень имеет тот же порядок малости, что и $\|\eta^{(n)}\| = \|b^{(n)} - a^{(n)}\|$; теперь из формулы (10.8) ясно, что оценки для $|v_n - u_n|$ и $\|b^{(n)} - a^{(n)}\|$ имеют один и тот же порядок малости. Если же координатная система не сильно минимальна, то $\|\eta^{(n)}\|$ имеет более низкий порядок малости, чем $\|\Gamma_n\|$ и $\|\delta^{(n)}\|$, а из той же формулы (10.8) ясно, что порядок малости величины $|v_n - u_n|$ выше, чем порядок малости $\|\eta^{(n)}\|$.

В то же время очевидно, что если координатная система не сильно минимальна, то порядок малости нормы $|v_n - u_n|$ ниже, чем порядок малости величин $\|\Gamma_n\|$ и $\|\delta^{(n)}\|$ и, следовательно, ниже, чем в случае сильно минимальной координатной системы.

§ 11. Число обусловленности матрицы Ритца

Введенное выше понятие устойчивости и основанные на нем оценки погрешности, допущенной при вычислении коэффициентов Ритца или приближенного решения по Ритцу, основаны на допущении, что система Ритца решается абсолютно точно, так что погрешности возникают только при составлении этой системы. Однако хорошо известно, что

---

[1]) Явления, аналогичные описанным выше, были обнаружены также в интересном числовом эксперименте, выполненном В. М. Миткевичем [1].

на практике решение алгебраических систем неизбежно связано с ошибками округления, влияние которых может быть весьма значительным. Известно также, что влияние ошибок округления тем заметнее, чем больше так называемое «число обусловленности» матрицы системы [1]). Поскольку матрица Ритца R_n — положительно определенная, ее обусловленность удобно характеризовать так называемым P-числом:

$$\rho(R_n) = \frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}, \tag{11.1}$$

где $\lambda_1^{(n)}$ — наименьшее, а $\lambda_n^{(n)}$ — наибольшее собственное число матрицы R_n.

Важно потребовать, чтобы с ростом n число $\rho(R_n)$ не возрастало неограниченно, — тогда можно ожидать, что ошибки округления также будут оставаться ограниченными. Нетрудно выяснить условия ограниченности P-числа матрицы Ритца При возрастании n число $\lambda_1^{(n)}$ не возрастает, а $\lambda_n^{(n)}$ не убывает, поэтому $\rho(R_n)$ не убывает [2]). При этом если $\lambda_1^{(n)} \to 0$ или $\lambda_n^{(n)} \to \infty$, то $\rho(R_n) \to \infty$ и потому для ограниченности P-числа матрицы R_n необходимо и достаточно, чтобы ее собственные числа были заключены между двумя положительными числами, которые не зависят от n. Иначе говоря, P-число матрицы Ритца ограничено независимо от n тогда и только тогда, когда координатная система почти ортонормирована в соответствующем энергетическом пространстве; если при этом $\lambda_0 \leqslant \lambda_k^{(n)} \leqslant \Lambda_0$, где λ_0 и Λ_0 — положительные постоянные, то

$$\rho(R_n) \leqslant \frac{\Lambda_0}{\lambda_0}. \tag{11.2}$$

[1]) См. Д. К. Фаддеев и В. Н. Фаддеева [1], § 15 и Дж. Х. Уилкинсон [1].

[2]) Приведем доказательство этого утверждения. По известному свойству собственных чисел симметричной матрицы

$$\lambda_1^{(n+1)} = \min_t \left\{ \sum_{j,\,k=1}^{n+1} [\varphi_j, \varphi_k]\, t_j \bar{t}_k : \sum_{k=1}^{n+1} |t_k|^2 \right\} \tag{*}$$

и

$$\lambda_{n+1}^{(n+1)} = \max_t \left\{ \sum_{j,\,k=1}^{n+1} [\varphi_j, \varphi_k]\, t_j \bar{t}_k : \sum_{k=1}^{n+1} |t_k|^2 \right\}; \tag{**}$$

здесь минимум или максимум берется по всевозможным векторам $t = (t_1, t_2, \ldots, t_n, t_{n+1})$, отличным от нулевого. Очевидно, $\lambda_1^{(n)}$ можно получить как минимум отношения (*) при дополнительном условии $t_{n+1} = 0$, а $\lambda_n^{(n)}$ — как максимум отношения (**) при том же условии $t_{n+1} = 0$. Но условный минимум не меньше безусловного, а условный максимум — соответственно не больше, поэтому $\lambda_1^{(n)} \geqslant \lambda_1^{(n+1)}$ и $\lambda_n^{(n)} \leqslant \lambda_{n+1}^{(n+1)}$.

Пусть, например, оператор B — полусходный с оператором A рассматриваемой задачи, так что

$$K_1 |u|_B \leqslant |u|_A \leqslant K_2 |u|_B, \tag{11.3}$$

где K_1 и K_2 — положительные постоянные, и пусть координатная система ортонормирована в пространстве H_B. В рассуждениях теорем 4.2 и 4.3 следует положить тогда $\mu_0 = M_0 = 1$, $K = K_1^{-1}$, и названные теоремы дают

$$K_1^2 \leqslant \lambda_k^{(n)} \leqslant K_2^2;$$

отсюда

$$\rho(R_n) \leqslant \frac{K_2^2}{K_1^2}. \tag{11.4}$$

Так, во втором примере § 8, как это видно из неравенства (8.7), $K_1 = 1$, $K_2 = \sqrt{3}$ и при любом n

$$\rho(R_n) \leqslant 3.$$

§ 12. Решение системы Ритца итерациями

Вопрос, разбираемый в настоящем параграфе, не связан непосредственно с понятием устойчивости; однако, как мы увидим ниже, он тесно связан с условиями, налагаемыми на координатную систему в связи с требованиями устойчивости.

Система Ритца

$$R_n a^{(n)} = f^{(n)} \tag{12.1}$$

при любом скалярном множителе $\alpha \neq 0$ равносильна следующей системе:

$$a^{(n)} = (I_n - \alpha R_n)\, a^{(n)} + \alpha f^{(n)}, \tag{12.2}$$

где I_n — единичная матрица порядка n.

Если $\lambda_k^{(n)}$ $(k = 1, 2, \ldots, n)$ суть собственные числа матрицы R_n, то собственные числа матрицы $I_n - \alpha R_n$ равны $1 - \alpha\lambda_k^{(n)}$. Выберем α так, чтобы было

$$-1 < 1 - \alpha\lambda_k^{(n)} < 1 \qquad (k = 1, 2, \ldots, n). \tag{12.3}$$

Тогда итерации, проводимые по формуле

$$[a^{(n)}]_{m+1} = (I_n - \alpha R_n)\,[a^{(n)}]_m + \alpha f^{(n)}, \tag{12.4}$$

сходятся к решению системы (12.1) при любом начальном приближении $[a^{(n)}]_0$.

Число α должно быть положительным — в противном случае $1 - \alpha\lambda_k^{(n)} > 1$, и итерации разойдутся. Если же $\alpha > 0$, то величина

$1-\alpha\lambda_k^{(n)}$ убывает с возрастанием k, и достаточно потребовать, чтобы было

$$1-\alpha\lambda_1^{(n)}<1, \qquad 1-\alpha\lambda_n^{(n)}>-1, \tag{12.5}$$

или

$$0<\alpha<\frac{2}{\lambda_n^{(n)}}. \tag{12.6}$$

Наивыгоднейшим будет такое α, при котором каждое из чисел (12.5) наиболее удалено от своей границы, иначе говоря, такое α, при котором

$$\max\left(\left|1-\alpha\lambda_1^{(n)}\right|, \left|1-\alpha\lambda_n^{(n)}\right|\right)=\min. \tag{12.7}$$

Обозначим для краткости $\lambda_1^{(n)}=m$, $\lambda_n^{(n)}=M$, так что $0<m\leqslant \leqslant M<\infty$. В плоскости координат α, β (рис. 1) рассмотрим линии

$$\beta=|1-\alpha m|, \qquad \beta=|1-\alpha M|. \tag{12.8}$$

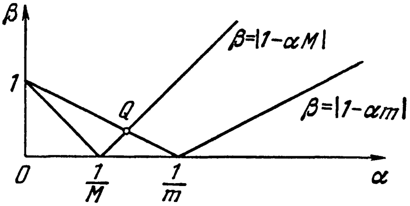

Рис. 1.

Условие (12.7) означает, что нужно найти такое значение α, при котором бо́льшая из ординат линий (12.8) будет минимальной. Из рис. 1 ясно, что таким значением α является абсцисса точки Q, лежащей на пересечении прямых $\beta=1-\alpha m$ и $\beta=\alpha M-1$. Отсюда $\alpha=\dfrac{2}{m+M}$, или [1])

$$\alpha=\frac{2}{\lambda_1^{(n)}+\lambda_n^{(n)}}; \tag{12.9}$$

очевидно, это значение лежит в промежутке (12.6). При таком α абсолютные величины чисел (12.5) равны между собой и их общее значение равно:

$$\frac{\lambda_n^{(n)}-\lambda_1^{(n)}}{\lambda_n^{(n)}+\lambda_1^{(n)}}=\frac{\rho(R_n)-1}{\rho(R_n)+1}; \tag{12.10}$$

итерации (12.4) сходятся не медленнее, чем геометрическая прогрессия со знаменателем (12.10).

С возрастанием n число (12.10) возрастает. Если координатная система — не почти ортонормированная в энергетической метрике, то это число стремится к единице при $n\to\infty$, и при больших n итерации будут сходиться медленно. Если же координатная система почти

[1]) Формулы (12.3) и (12.9) являются частными случаями формул, полученных И. П. Натансоном [1] для операторных уравнений весьма общего вида.

ортонормирована в энергетической метрике, так что $\lambda_0 \leqslant \lambda_k^{(n)} \leqslant \Lambda_0$, то $\rho(R_n) \leqslant \frac{\Lambda_0}{\lambda_0}$ (неравенство (11.2)), и

$$\frac{\rho(R_n)-1}{\rho(R_n)+1} \leqslant \frac{\Lambda_0-\lambda_0}{\Lambda_0+\lambda_0}.$$

Заметим, что формула (12.9) требует знания собственных чисел матрицы Ритца и потому практически неудобна. Однако, если координатная система в энергетической метрике почти ортонормирована, а числа λ_0 и Λ_0 известны, то можно взять

$$\alpha = \frac{2}{\lambda_0+\Lambda_0}. \tag{12.11}$$

При этом

$$0 < 1-\alpha\lambda_1^{(n)} = \frac{\Lambda_0+\lambda_0-2\lambda_1^{(n)}}{\Lambda_0+\lambda_0} < \frac{\Lambda_0-\lambda_0}{\Lambda_0+\lambda_0},$$

$$0 > 1-\alpha\lambda_n^{(n)} = \frac{\Lambda_0+\lambda_0-2\lambda_n^{(n)}}{\Lambda_0+\lambda_0} > -\frac{\Lambda_0-\lambda_0}{\Lambda_0+\lambda_0};$$

спектр матрицы $I_n - \alpha R_n$ заключен в промежутке $\left(-\frac{\Lambda_0-\lambda_0}{\Lambda_0+\lambda_0}, \frac{\Lambda_0-\lambda_0}{\Lambda_0+\lambda_0}\right)$, и при любом n итерации сходятся не медленнее, чем прогрессия со знаменателем

$$\frac{\Lambda_0-\lambda_0}{\Lambda_0+\lambda_0}. \tag{12.12}$$

Пусть, в частности, координатная система ортонормирована в энергетической метрике оператора B, полусходного с данным оператором A, так что имеет место неравенство (11.3). Тогда, как мы видели, можно положить $\lambda_0 = K_1^2$, $\Lambda_0 = K_2^2$. Формула (12.11) в этом случае дает

$$\alpha = \frac{2}{K_1^2+K_2^2}, \tag{12.13}$$

и итерации сходятся, как прогрессия со знаменателем

$$\frac{K_2^2-K_1^2}{K_2^2+K_1^2}. \tag{12.14}$$

Так, для системы (8.9) или близкой к ней системы (8.10) формула (12.13) дает[1])

$$\alpha = \frac{2}{1+3} = 0{,}5.$$

[1]) Значения $K_1 = 1$ и $K_2 = \sqrt{3}$ для этой задачи даны в конце § 11.

Система (12.2) настоящего параграфа (полученная преобразованием системы (8.11)) в данном случае принимает вид (мы ограничиваемся здесь усеченной системой четвертого порядка)

$$\left.\begin{aligned}
b_1 &= 0{,}2b_1 - 0{,}20656b_2 - 0{,}32660,\\
b_2 &= -0{,}20656b_1 + 0{,}2b_2 - 0{,}20284b_3,\\
b_3 &= -0{,}20284b_2 + 0{,}2b_3 - 0{,}20160b_4,\\
b_4 &= -0{,}20160b_3 + 0{,}2b_4.
\end{aligned}\right\} \tag{12.15}$$

Итерации для этой системы сходятся как прогрессия со знаменателем 0,5.

Ниже приведено решение системы (12.15), полученное методом итераций. За начальное приближение были взяты свободные члены системы; первые четыре знака стабилизировались, начиная с 12-й итерации:

$$b_1 = -0{,}4397; \quad b_2 = 0{,}1219; \quad b_3 = -0{,}0330; \quad b_4 = 0{,}0083.$$

Как мы видим, метод итераций дал те же значения, которые были получены в § 8 (см. табл. 8.5).

В заключение заметим следующее. Может случиться, что хотя координатная система и почти ортонормирована, но одно из двух чисел Λ_0 и λ_0 (или даже оба эти числа) на самом деле нам неизвестно. Тогда воспользоваться формулой (12.11) невозможно. Однако если число Λ_0 известно, то воспользоваться итерациями все же возможно. Действительно, любое значение α из промежутка $0 < \alpha \leqslant \frac{2}{\Lambda_0}$ удовлетворяет неравенствам (12.6) и, следовательно, обеспечивает сходимость процесса итераций.

§ 13. Обобщение понятия об устойчивости[1])

Определение устойчивости, данное в § 9, допускает простое и естественное обобщение.

Пусть некоторый вычислительный процесс заключается в решении последовательности уравнений

$$A_n x^{(n)} = y^{(n)} \qquad (n = 1,\ 2,\ \ldots), \tag{13.1}$$

где A_n — оператор, действующий из некоторого банахова пространства X_n в другое банахово пространство Y_n. Будем считать, что для всех n обратный оператор A_n^{-1} существует и определен на всем пространстве Y_n.

[1]) См. статью автора [22].

Наряду с уравнениями (13.1) рассмотрим последовательность уравнений

$$(A_n+\Gamma_n)\,z^{(n)}=y^{(n)}+\delta^{(n)}. \tag{13.2}$$

Будем говорить, что данный вычислительный процесс устойчив, если существуют такие независящие от n постоянные p, q, r, что при $\|\Gamma_n\|\leqslant r$ и любых $\delta^{(n)}$ уравнение (13.2) разрешимо и справедливо неравенство

$$\|\eta^{(n)}\|\leqslant p\,\|\Gamma_n\|+q\,\|\delta^{(n)}\|, \tag{13.3}$$

где $\eta^{(n)}=z^{(n)}-x^{(n)}$.

Теорема 13.1. *Для устойчивости вычислительного процесса* (13.1) *необходимо и достаточно существование таких независящих от n постоянных C_1 и C_2, чтобы выполнялись следующие соотношения*: 1) $\|A_n^{-1}\|\leqslant C_1$; 2) *какова бы ни была последовательность операторов B_n с единичной нормой, действующих соответственно из пространства X_n в пространство Y_n, имеет место неравенство* $\|A_n^{-1}B_nA_n^{-1}y^{(n)}\|=\|A_n^{-1}B_nx^{(n)}\|\leqslant C_2$.

Необходимость. 1) Пусть $\Gamma_n=0$. Тогда из неравенства (13.3) следует $\|\eta^{(n)}\|\leqslant q\,\|\delta^{(n)}\|$. Но в данном случае $A_n\eta^{(n)}=A_nz^{(n)}-A_nx^{(n)}=\delta^{(n)}$ и, следовательно, $\|A_n\eta^{(n)}\|\geqslant\frac{1}{q}\|\eta^{(n)}\|$. Последнее неравенство означает, что $\|A_n^{-1}\|\leqslant q$, и можно положить $C_1=q$.

2) Пусть теперь $\delta^{(n)}=0$ и $\Gamma_n=\varepsilon B_n$, где $\|B_n\|=1$ и $\varepsilon=\|\Gamma_n\|\leqslant r$. В данном случае

$$(A_n+\Gamma_n)\,z^{(n)}=y^{(n)};$$

отсюда

$$z^{(n)}+A_n^{-1}\Gamma_nz^{(n)}=x^{(n)},$$

или

$$z^{(n)}-x^{(n)}+A_n^{-1}\Gamma_n\left(z^{(n)}-x^{(n)}\right)=-A_n^{-1}\Gamma_nx^{(n)}. \tag{13.4}$$

Далее,

$$\|A_n^{-1}\Gamma_nx^{(n)}\|=\varepsilon\,\|A_n^{-1}B_nx^{(n)}\|=\|\Gamma_n\|\,\|A_n^{-1}B_nx^{(n)}\|$$

и

$$\|A_n^{-1}\Gamma_n\left(z^{(n)}-x^{(n)}\right)\|=$$
$$=\varepsilon\,\|A_n^{-1}B_n\left(z^{(n)}-x^{(n)}\right)\|\leqslant\varepsilon\,\|A_n^{-1}\|\,\|z^{(n)}-x^{(n)}\|\leqslant\varepsilon q\,\|z^{(n)}-x^{(n)}\|.$$

Теперь из уравнения (13.4) следует:

$$\|\eta^{(n)}\|=\|z^{(n)}-x^{(n)}\|\geqslant\frac{\|\Gamma_n\|\,\|A_n^{-1}B_nx^{(n)}\|}{1+\varepsilon q}.$$

Сравнивая это с неравенством (13.3), которое в рассматриваемом случае принимает вид $\|\eta^{(n)}\| \leqslant p\,\|\Gamma_n\|$, находим, что необходимо

$$\|A_n^{-1}B_nx^{(n)}\| \leqslant p\,(1+\varepsilon q);$$

так как ε произвольно мало, то необходимо

$$\|A_n^{-1}B_nx^{(n)}\| \leqslant p,$$

и можно положить $C_2 = p$.

Достаточность. Пусть выполнены условия теоремы 13.1.

Положим $r = \frac{\beta}{C_1}$, где β — постоянная, $0 < \beta < 1$, и потребуем, чтобы было $\|\Gamma_n\| \leqslant r$.

Составим уравнения

$$(A_n + \Gamma_n)\,z_1^{(n)} = y^{(n)}, \qquad (A_n + \Gamma_n)\,z_2^{(n)} = \delta^{(n)}. \tag{13.5}$$

Тогда $z^{(n)} = z_1^{(n)} + z_2^{(n)}$ и $\eta^{(n)} = \left(z_1^{(n)} - x^{(n)}\right) + z_2^{(n)}$. Если $\Gamma_n = 0$, то $z_1^{(n)} = x^{(n)}$ и $z_2^{(n)} = \eta^{(n)} = A_n^{-1}\delta^{(n)}$; при этом

$$\|\eta^{(n)}\| \leqslant \|A_n^{-1}\|\,\|\delta^{(n)}\| \leqslant C_1\,\|\delta^{(n)}\|,$$

и теорема доказана. Если же $\Gamma_n \neq 0$, то положим $\Gamma_n = \|\Gamma_n\|\,B_n$, $\|B_n\| = 1$. Совсем просто оценивается $\|z_2^{(n)}\|$:

$$z_2^{(n)} = (I_n + A_n^{-1}\Gamma_n)^{-1}\,A_n^{-1}\,\delta^{(n)},$$

где I_n — тождественный оператор в пространстве X_n. Отсюда

$$\|z_2^{(n)}\| \leqslant C_1\,\|(I_n + A_n^{-1}\Gamma_n)^{-1}\|\,\|\delta^{(n)}\|.$$

Нетрудно видеть, что

$$\|(I_n + A_n^{-1}\Gamma_n)^{-1}\| \leqslant \frac{1}{1-\beta},$$

поэтому

$$\|z_2^{(n)}\| \leqslant \frac{C_1}{1-\beta}\,\|\delta^{(n)}\|. \tag{13.6}$$

Первое из уравнений (13.5) преобразуем к виду

$$(I_n + A_n^{-1}\Gamma_n)z_1^n = x^{(n)};$$

отсюда

$$(I_n + A_n^{-1}\Gamma_n)\left(z_1^{(n)} - x^{(n)}\right) = -\,A_n^{-1}\Gamma_nx^{(n)}.$$

Это дает нам

$$z_1^{(n)} - x^{(n)} = -(I_n + A_n^{-1}\Gamma_n)^{-1}A_n^{-1}\Gamma_nx^{(n)}$$

и

$$\|z_1^{(n)} - x^{(n)}\| \leqslant \frac{\|A_n^{-1}\Gamma_nx^{(n)}\|}{1-\beta} = \frac{\|A_n^{-1}B_nx^{(n)}\|}{1-\beta}\,\|\Gamma_n\| \leqslant \frac{C_2}{1-\beta}\,\|\Gamma_n\|. \tag{13.7}$$

Из формул (13.6) и (13.7) вытекает, что неравенство (13.3) выполняется при значениях постоянных

$$p=\frac{C_2}{1-\beta}, \qquad q=\frac{C_1}{1-\beta}.$$

Замечание 1. Пусть выполнено условие 1) теоремы 13.1: $\|A_n^{-1}\|\leqslant C_1$. Тогда, если $\|x^{(n)}\|\leqslant C_3=\text{const}$, то условие 2) этой теоремы также выполнено: $\|A_n^{-1}B_nx^{(n)}\|\leqslant\|A_n^{-1}\|\|x^{(n)}\|\leqslant C_1C_3$. Таким образом, для устойчивости процесса (13.1) достаточно, чтобы величины $\|A_n^{-1}\|$ и $\|A_n^{-1}y^{(n)}\|=\|x^{(n)}\|$ были ограничены в совокупности.

Замечание 2. Если нормы $\|A_n^{-1}\|$ положительно ограничены снизу, то из условия 2) теоремы 13.1 вытекает ограниченность величин $\|x^{(n)}\|$ в совокупности. Для доказательства допустим противное и пусть $\|x^{(n_k)}\|\to\infty$. Выберем элемент $t^{(n_k)}$ так, чтобы $\|t^{(n_k)}\|=1$ и $\|A_{n_k}^{-1}t^{(n_k)}\|>\frac{1}{2}\|A_{n_k}^{-1}\|$. Построим линейный функционал $f_{n_k}(x)$ такой, чтобы $f_{n_k}(x^{(n_k)})=\|x^{(n_k)}\|$ и $\|f_{n_k}\|=1$. Положим $B_{n_k}x=$ $=f_{n_k}(x)\,t^{(n_k)}$. Тогда

$$\|B_{n_k}x\|=|f_{n_k}(x)|\leqslant\|x\|;$$

знак равенства достигается при $x=x^{(n_k)}$ и, следовательно, $\|B_{n_k}\|=1$. При этом

$$A_{n_k}^{-1}B_{n_k}x^{(n_k)}=\|x^{(n_k)}\|\,A_{n_k}^{-1}t^{(n_k)}$$

и

$$\|A_{n_k}^{-1}B_{n_k}x^{(n_k)}\|>\frac{1}{2}\|A_{n_k}^{-1}\|\|x^{(n_k)}\|\to\infty,$$

вопреки условию 2).

Результаты, отмеченные в замечаниях 1 и 2, позволяют сформулировать следующую теорему.

Теорема 13.2. *Если нормы* $\|A_n^{-1}\|$ *положительно ограничены снизу в совокупности, то для устойчивости вычислительного процесса* (13.1) *необходимо и достаточно, чтобы величины* $\|A_n^{-1}\|$ *и* $\|x^{(n)}\|=\|A_n^{-1}y^{(n)}\|$ *были в совокупности ограничены сверху.*

В случае процесса Ритца $A_n=R_n$ и нормы $\|R_n^{-1}\|$ положительно ограничены снизу. В самом деле, $\|R_n^{-1}\|=[\lambda_1^{(n)}]^{-1}$; с возрастанием n собственное число $\lambda_1^{(n)}$ не возрастает и потому $[\lambda_1^{(n)}]^{-1}\geqslant[\lambda_1^{(1)}]^{-1}=$ $=|\varphi_1|^{-2}$. Теперь к процессу Ритца применима теорема 13.2.

Ограниченность норм $\|R_n^{-1}\|$ означает сильную минимальность в пространстве H_A координатной системы, из которой как следствие вытекает ограниченность норм $\|x^{(n)}\| = \|a^{(n)}\|$ (см. формулу (9.12)), и теорема 9.1 оказывается следствием из теоремы 13.2.

Замечание 3. Пусть X_n суть подпространства некоторого банахова пространства X.

Назовем вычислительный процесс (13.1) сходящимся, если существует предел $x_0 = \lim\limits_{n\to\infty} x^{(n)}$. В этом случае величины $\|x^{(n)}\|$ ограничены в совокупности и из теоремы 13.1 вытекает

Теорема 13.3. *Если процесс* (13.1) *сходящийся, то для его устойчивости необходимо и достаточно, чтобы нормы* $\|A_n^{-1}\|$ *были ограничены в совокупности.*

Из теоремы 13.3 вытекает устойчивость приближенного по Ритцу решения в некотором новом смысле. В данном случае $X = H_A$, X_n есть подпространство H_A, натянутое на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$; под Y_n следует понимать n-мерное унитарное пространство.

Обозначим через S_n оператор, который каждому элементу $u_n \in X_n$ приводит в соответствие вектор $a^{(n)}$ его коэффициентов: если

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k$$

и $a^{(n)} = \left(a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}\right)$, то $S_n u_n = a^{(n)}$. Пусть элемент u_n построен с помощью процесса Ритца. В этом случае по данному вектору $f^{(n)}$ вектор $a^{(n)}$ определяется уравнением (9.5):

$$R_n a^{(n)} = f^{(n)}.$$

Отсюда

$$R_n S_n u_n = f^{(n)} \tag{13.8}$$

и, следовательно, здесь естественно ввести обозначение

$$A_n = R_n S_n. \tag{13.9}$$

Далее,

$$A_n^{-1} f^{(n)} = u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k. \tag{13.10}$$

Воспользуемся формулой

$$|u_n|_A^2 = \left(R_n a^{(n)}, a^{(n)}\right),$$

которая по существу совпадает с формулой (10.5). Положив в ней $a^{(n)} = R_n^{-1} f^{(n)}$, получим:

$$|A_n^{-1} f^{(n)}|_A^2 = |u_n|_A^2 = \left(f^{(n)}, R_n^{-1} f^{(n)}\right). \tag{13.11}$$

Из уравнения (13.8) следует, что $u_n = S_n^{-1} R_n^{-1} f^{(n)}$ и

$$|u_n|_A^2 = \left[S_n^{-1} R_n^{-1} f^{(n)}, S_n^{-1} R_n^{-1} f^{(n)}\right]_A = \left(\left(S_n^{-1}\right)^* S_n^{-1} R_n^{-1} f^{(n)}, R_n^{-1} f^{(n)}\right)_{E_n}.$$

Сравнив это с формулой (13.11) и учтя, что $R_n^{-1}f^{(n)}$ есть произвольный вектор, придем к соотношению

$$R_n = (S_n^{-1})^* S_n^{-1}. \tag{13.12}$$

Отсюда

$$A_n = (S_n^{-1})^* \tag{13.13}$$

и

$$\|A_n^{-1}\| = \|S_n^*\| = \|S_n\| = \sqrt{\|R_n^{-1}\|} = \frac{1}{\sqrt{\lambda_1^{(n)}}}. \tag{13.14}$$

В метрике H_A приближенные решения u_n сходятся к точному решению u_0, поэтому процесс построения приближенных решений — сходящийся. По теореме 13.3, для его устойчивости необходимо и достаточно, чтобы нормы $\|A_n^{-1}\|$ были ограничены в совокупности; как показывает формула (13.14), для этого в свою очередь необходимо и достаточно, чтобы координатная система была сильно минимальной в H_A.

Итак, мы доказали, что если координатная система сильно минимальна в H_A, то приближенное решение по Ритцу устойчиво в следующем смысле: существуют такие три положительные постоянные p', q', r', что если оператор A_n (формула (13.9)) вычислен с малой погрешностью $\tilde{\Gamma}_n$, $\|\tilde{\Gamma}_n\| \leqslant r'$, а вектор $f^{(n)}$ — с погрешностью $\delta^{(n)}$, то энергетическая норма погрешности $v_n - u_n$ приближенного решения u_n не превосходит величины $p'\|\tilde{\Gamma}_n\| + q'\|\delta^{(n)}\|$.

Докажем, что отсюда вытекает утверждение теоремы 10.1. Пусть матрица R_n вычислена с малой ошибкой Γ_n, тогда $A_n = (S_n^{-1})^*$ окажется вычисленной с некоторой ошибкой $\tilde{\Gamma}_n$. Заметим, что при этом оператор S_n^{-1} вычисляется без ошибки, так как

$$S_n^{-1} a_n = u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k,$$

а координатные элементы предполагаются заданными точно; ошибками же арифметических действий в последней формуле мы пренебрежем. Ошибки Γ_n и $\tilde{\Gamma}_n$ в силу формулы (13.12) связаны соотношением

$$R_n + \Gamma_n = \left[(S_n^{-1})^* + \tilde{\Gamma}_n\right] S_n^{-1}$$

или $\tilde{\Gamma}_n = \Gamma_n S_n$. Отсюда

$$\|\tilde{\Gamma}_n\| \leqslant \|\Gamma_n\|\,\|S_n\| = \frac{1}{\sqrt{\lambda_1^{(n)}}}\|\Gamma_n\| < \lambda_0^{-\frac{1}{2}}\|\Gamma_n\|, \tag{13.15}$$

где λ_0 — положительная нижняя граница чисел $\lambda_1^{(n)}$. Теперь, если координатная система сильно минимальна в H_A, то, по доказанному выше, $|v_n - u_n| \leqslant p' \|\tilde{\Gamma}_n\| + q' \|\delta^{(n)}\|$. По неравенству (13.15),

$$|v_n - u_n| \leqslant p_1 \|\Gamma_n\| + q_1 \|\delta^{(n)}\|, \qquad p_1 = p'\lambda_0^{-\frac{1}{2}}, \qquad q_1 = q',$$

что и требовалось доказать.

§ 14. Об устойчивости процесса Бубнова — Галёркина для стационарных задач

Напомним общую схему процесса Бубнова — Галёркина[1]). Пусть в уравнении

$$Au - f = 0 \tag{14.1}$$

оператор A имеет вид

$$A = A_0 + K, \tag{14.2}$$

где A_0 положительно определен в данном сепарабельном гильбертовом пространстве H, $D(K) \supset D(A_0)$ и произведение $T = A_0^{-1}K$ можно расширить до оператора, вполне непрерывного в энергетическом пространстве $H_0 = H_{A_0}$. Уравнение (14.1) равносильно уравнению

$$u + Tu - A_0^{-1}f = 0. \tag{14.3}$$

Точнее говоря, всякое решение уравнения (14.1) удовлетворяет уравнению (14.3) и всякое решение уравнения (14.3), принадлежащее области $D(A_0)$ задания оператора A_0, удовлетворяет уравнению (14.1); если же уравнение (14.3) имеет решение (оно необходимо принадлежит пространству H_0), не принадлежащее области $D(A_0)$, то мы рассматриваем его как обобщенное решение уравнения (14.1).

По методу Бубнова — Галёркина выбираем координатную систему

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, \tag{14.4}$$

подчиненную, условиям, аналогичным условиям § 6: 1) $\varphi_n \in H_0$ $(n = 1, 2, \ldots)$; 2) при любом n элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы; 3) последовательность (14.4) полна в H_0.

Приближенное решение уравнения (14.1) или, что то же, уравнения (14.3) ищем в виде

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k; \tag{14.5}$$

коэффициенты $a_k^{(n)}$ определяются из условия, чтобы после подста-

---

[1]) См. ВМ, §§ 7?, 78, 82.

новки u_n вместо u в левую часть уравнения (14.3) последняя стала ортогональной в метрике пространства H_0 к элементам $\varphi_1, \varphi_2, \ldots, \varphi_n$. Система уравнений Бубнова — Галёркина имеет вид:

$$\sum_{k=1}^{n} \left\{ [\varphi_k, \varphi_j] + [T\varphi_k, \varphi_j] \right\} a_k^{(n)} = [A_0^{-1} f, \varphi_j] \quad (j = 1, 2, \ldots, n)$$

или, что то же,

$$\sum_{k=1}^{n} \left\{ [\varphi_k, \varphi_j] + [T\varphi_k, \varphi_j] \right\} a_k^{(n)} = (f, \varphi_j) \quad (j = 1, 2, \ldots, n). \qquad (14.6)$$

Здесь квадратные скобки означают энергетическое произведение в пространстве H_0. Если $D(K) \supset H_0$, то $[T\varphi_k, \varphi_j] = [A_0^{-1} K\varphi_k, \varphi_j] = (K\varphi_k, \varphi_j)$, и систему (14.6) можно представить в виде:

$$\sum_{k=1}^{n} \left\{ [\varphi_k, \varphi_j] + (K\varphi_k, \varphi_j) \right\} a_k^{(n)} = (f, \varphi_j) \quad (j = 1, 2, \ldots, n). \qquad (14.6')$$

Эта форма имеет то важное преимущество, что она не требует фактического вычисления обратного оператора A_0^{-1}. Наконец, если $\varphi_k \in D(A_0)$, то $[\varphi_k, \varphi_j] = (A_0\varphi_k, \varphi_j)$, и мы получаем обычную форму уравнений Бубнова — Галёркина:

$$\sum_{k=1}^{n} \left\{ (A_0\varphi_k, \varphi_j) + (K\varphi_k, \varphi_j) \right\} a_k^{(n)} = (f, \varphi_j) \quad (j=1, 2, \ldots, n). \qquad (14.6'')$$

Уравнения (14.6″) получаются, если в левую часть уравнения (14.1) подставить u_n вместо u и потребовать, чтобы полученное выражение было ортогонально (в метрике данного пространства H) к элементам $\varphi_1, \varphi_2, \ldots, \varphi_n$. Известно (ВМ, § 78), что если уравнение (14.1) имеет только одно решение, то приближенные решения (14.5) сходятся к точному решению в метрике пространства H_0.

Теорема 14.1. *Если координатная система* (14.4) *сильно минимальна в пространстве* H_0 *и уравнение* (14.1) *имеет только одно решение, то процесс Бубнова — Галёркина* (14.6) *устойчив* [1]).

Обозначим через R_n и B_n матрицы n-го порядка с элементами $[\varphi_k, \varphi_j]$ и $[T\varphi_k, \varphi_j]$ $(j, k = 1, 2, \ldots, n)$ соответственно, через $a^{(n)}$ — вектор с составляющими $a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}$ и через $f^{(n)}$ — вектор с составляющими $(f, \varphi_1), (f, \varphi_2), \ldots, (f, \varphi_n)$. Систему (14.6) можно теперь записать в виде одного векторного уравнения:

$$(R_n + B_n)\, a^{(n)} = f^{(n)}. \qquad (14.7)$$

---

[1]) См. статью Г. Н. Ясковой и М. Н. Яковлева [1], где исследован более общий вычислительный процесс, а именно так называемый процесс Галёркина — Петрова.

В соответствии с замечанием 1 § 13, достаточно доказать, что величины $\|(R_n+B_n)^{-1}\|$ и $\|a^{(n)}\|$ ограничены в совокупности. Докажем сначала ограниченность величин $\|a^{(n)}\|$. Как и в § 9, мы найдем, что $\|a^{(n)}\| \leqslant \lambda_0^{-\frac{1}{2}} |u_n|$, где λ_0 — положительная нижняя граница собственных чисел матриц R_n; такая граница существует, так как в H_0 координатная система сильно минимальна. Далее, если u_0 есть решение уравнения (14.1), то $|u_n - u_0| \to 0$. Отсюда $|u_n| \to |u_0|$ и потому величины $|u_n|$ ограничены: существует такая постоянная c_0, что $|u_n| \leqslant c_0$. Теперь $\|a^{(n)}\| \leqslant \lambda_0^{-\frac{1}{2}} c_0 = \text{const}$, что и требовалось доказать.

Займемся доказательством ограниченности норм $\|(R_n+B_n)^{-1}\|$. Рассмотрим сначала случай, когда координатная система ортонормирована в H_0. Для этого случая изменим некоторые обозначения. Координатные элементы будем обозначать через ω_n $(n=1, 2, \ldots)$; при этом

$$[\omega_k, \omega_j] = \delta_{kj}.$$

В этом случае матрица $R_n = I_n$, где I_n — единичная матрица порядка n. Далее, матрицу элементов $[T\omega_k, \omega_j]$ $(j=1, 2, \ldots, n)$ обозначим через Δ_n. Уравнение (14.7) тогда принимает вид:

$$(I_n + \Delta_n)\, a^{(n)} = g^{(n)}, \tag{14.7'}$$

где $g^{(n)}$ обозначает вектор с составляющими (f, ω_1), (f, ω_2), ..., (f, ω_n). Уравнение (14.1) строго равносильно бесконечной системе[1])

$$a_j + \sum_{k=1}^{\infty} [T\omega_k, \omega_j]\, a_k = (f, \omega_j) \qquad (j=1, 2, \ldots). \tag{14.8}$$

Если обозначить через Δ бесконечную матрицу элементов $[T\omega_k, \omega_j]$, через a — последовательность $(a_1, a_2, \ldots)$, через g — последовательность (f, ω_1), (f, ω_2), ..., то систему (14.8) можно записать в виде

$$(I+\Delta)\, a = g, \tag{14.8'}$$

где I — единичная бесконечная матрица.

Заметим, что последовательности a и g суть элементы пространства l_2, так как $a_j = [u_0, \omega_j]$ суть коэффициенты Фурье решения u_0 по ортонормированной в H_0 системе (14.4), а $(f, \omega_j) = [A_0^{-1} f, \omega_j]$ суть коэффициенты Фурье элемента $A_0^{-1} f$ по той же системе.

Матрица Δ порождает в пространстве l_2 вполне непрерывный оператор. Будучи равносильной уравнению (14.3), бесконечная

[1]) См. книгу автора [4], § 52.

система (14.8′) имеет в l_2 единственное решение. А тогда, как известно, при достаточно больших n система (14.7) также имеет единственное решение. Это означает, что существует матрица $(I+\Delta)^{-1}$ и, при достаточно больших n, существуют матрицы $(I_n+\Delta_n)^{-1}$.

Известно также, что матрица $G=(I+\Delta)^{-1}-I$ порождает в l_2 вполне непрерывный оператор. Обозначим $G_n=(I_n+\Delta_n)^{-1}-I_n$. Дополним матрицу G_n нулями до бесконечной матрицы, которую обозначим через $\widetilde{G}_n$. Тогда [1])

$$\|\widetilde{G}_n-G\| \xrightarrow[n\to\infty]{} 0. \tag{14.9}$$

Докажем что

$$\|\widetilde{G}_n\|_{l_2}=\|G_n\|_{E_n}; \tag{14.10}$$

норму в левой части равенства следует понимать как норму оператора в l_2, а в правой части — как норму оператора в унитарном n-мерном пространстве E_n. Имеем:

$$\|\widetilde{G}_n\|=\sup_{\tilde{x}\in l_2}\frac{\|\widetilde{G}_n\tilde{x}\|}{\|\tilde{x}\|}. \tag{14.11}$$

Обозначим через g_{jk} элементы матрицы G_n; тогда, если

$$\tilde{x}=(x_1,\ x_2,\ \ldots,\ x_n,\ x_{n+1},\ \ldots),$$

то j-я составляющая вектора $\widetilde{G}_n(\tilde{x})$ определяется формулой

$$[\widetilde{G}_n(\tilde{x})]_j=\begin{cases}\sum\limits_{k=1}^{n} g_{jk}x_k, & j\leqslant n,\\ 0, & j>n.\end{cases}$$

Отсюда видно, что

$$\|\widetilde{G}_n(\tilde{x})\|_{l_2}=\|\widetilde{G}_n(\tilde{x}^{(n)})\|_{l_2}=\|G_n(x^{(n)})\|_{E_n}, \tag{14.12}$$

где $x^{(n)}\in E_n$ есть вектор $(x_1,\ x_2,\ \ldots,\ x_n)$, а

$$\tilde{x}^{(n)}=(x_1,\ x_2,\ \ldots,\ x_n,\ 0,\ 0,\ \ldots). \tag{14.13}$$

Далее, очевидно, что

$$\frac{\|\widetilde{G}_n(\tilde{x})\|_{l_2}}{\|\tilde{x}\|_{l_2}}\leqslant\frac{\|\widetilde{G}_n(\tilde{x}^{(n)})\|_{l_2}}{\|\tilde{x}^{(n)}\|_{l_2}}=\frac{\|G_n(x^{(n)})\|_{E_n}}{\|x^{(n)}\|_{E_n}}.$$

[1]) См. Л. В. Канторович и Г. П. Акилов [1], стр. 318.

Обозначая через $\widetilde{E}_n \subset l_2$ множество векторов вида (14.13), имеем из последнего неравенства:

$$\sup_{x \in l_2} \frac{\| \widetilde{G}_n(\widetilde{x}) \|_{l_2}}{\| \widetilde{x} \|_{l_2}} \leqslant \sup_{\widetilde{x}^{(n)} \in \widetilde{E}_n} \frac{\| \widetilde{G}_n(\widetilde{x}^{(n)}) \|_{l_2}}{\| \widetilde{x}^{(n)} \|_{l_2}} = \sup_{E_n} \frac{\| G_n(x^{(n)}) \|_{E_n}}{\| x^{(n)} \|_{E_n}}.$$

Но множество l_2 шире, чем $\widetilde{E}_n$, поэтому первый член в последнем соотношении не меньше второго, и мы получаем:

$$\sup_{x \in l_2} \frac{\| \widetilde{G}_n(\widetilde{x}) \|_{l_2}}{\| \widetilde{x} \|_{l_2}} = \sup_{x \in E_n} \frac{\| G_n(x^{(n)}) \|_{E_n}}{\| x^{(n)} \|_{E_n}},$$

что тождественно с соотношением (14.10).

Из формулы (14.9) следует, что $\| \widetilde{G}_n \|_{l_2} \to \| G \|_{l_2}$, или, в силу равенства (14.10), $\| G_n \|_{E_n} \to \| G \|_{l_2}$. Отсюда следует, что величины $\| G_n \|_{E_n}$ ограничены в совокупности; вместе с ним ограничены и величины $\| (I_n + \Delta_n)^{-1} \| = \| G_n + I_n \| \leqslant \| G_n \| + 1$.

Обратимся теперь к общему случаю. Пусть координатная система (14.4) не ортонормирована, а только сильно минимальна в H_0. Подвергнем систему (14.4) ортогонализации в метрике H_0, и пусть этот процесс приводит нас к ортонормированной системе

$$\omega_1, \ \omega_2, \ \ldots, \ \omega_n, \ \ldots, \quad [\omega_j, \ \omega_k] = \delta_{jk}.$$

Как известно, элемент φ_k линейно выражается через $\omega_1, \omega_2, \ldots, \omega_k$, поэтому

$$\varphi_k = \sum_{s=1}^{k} [\varphi_k, \ \omega_s] \, \omega_s.$$

Обозначим через C_n левую треугольную матрицу элементов $[\varphi_k, \omega_s]$ $(k = 1, 2, \ldots, n; \ s = 1, 2, \ldots, k)$. Нетрудно видеть, что

$$R_n = C_n C_n^*, \qquad B_n = C_n \Delta_n C_n^*, \tag{14.14}$$

где звездочка обозначает сопряженную матрицу, а через Δ_n обозначена матрица элементов $[T\omega_k, \omega_j]$ $(k, j = 1, 2, \ldots, n)$. Отсюда

$$R_n + B_n = C_n (I_n + \Delta_n) C_n^*$$

и

$$(R_n + B_n)^{-1} = (C_n^*)^{-1} (I_n + \Delta_n)^{-1} C_n^{-1}. \tag{14.15}$$

Из первого равенства (14.14) следует:

$$\| C_n^{-1} \| = \| (C_n^*)^{-1} \| = \sqrt{\| R_n^{-1} \|} = \frac{1}{\sqrt{\lambda_1^{(n)}}} \leqslant \frac{1}{\sqrt{\lambda_0}}$$

и соотношение (14.15) дает:

$$\|(R_n+B_n)^{-1}\| \leqslant \frac{1}{\lambda_0}\|(I_n+\Delta_n)^{-1}\|,$$

что ограничено в силу доказанного выше. Этим завершается доказательство теоремы 14.1.

Буквально, так же, как это было сделано в § 10 для приближенного решения по Ритцу, можно ввести понятие об устойчивости приближенного решения по Бубнову — Галёркину. Мы не станем приводить определения, считая его очевидным.

Теорема 14.2. *Для того чтобы приближенное решение по Бубнову — Галёркину было устойчиво, необходимо и достаточно, чтобы координатная система была сильно минимальна в* H_0 [1]).

Обозначим через Γ_n матрицу погрешностей, через $\delta^{(n)}$ — вектор погрешностей свободных членов, через $b^{(n)}$ — вектор «неточных» коэффициентов Бубнова — Галёркина. Тогда

$$(R_n+B_n+\Gamma_n)\,b^{(n)}=f^{(n)}+\delta^{(n)}. \qquad (14.16)$$

Обозначим еще через v_n «неточное» приближение по Бубнову — Галёркину

$$v_n=\sum_{k=1}^{n} b_k\varphi_k.$$

Тогда (ср. § 10)

$$|v_n-u_n|^2=\left(R_n\left(b^{(n)}-a^{(n)}\right),\ b^{(n)}-a^{(n)}\right). \qquad (14.17)$$

Введем обозначения

$$C_n^*a^{(n)}=\alpha^{(n)}, \quad C_n^*b^{(n)}=\beta^{(n)}. \qquad (14.18)$$

Как мы видели выше (см. формулу (14.14)), $R_n=C_nC_n^*$, и соотношение (14.17) дает:

$$|v_n-u_n|=\|\beta^{(n)}-\alpha^{(n)}\|. \qquad (14.19)$$

В силу соотношений (14.14) уравнения (14.7) и (14.16) дают:

$$C_n(I_n+\Delta_n)\,\alpha^{(n)}=f^{(n)},$$

$$\left[C_n(I_n+\Delta_n)+\Gamma_n\left(C_n^*\right)^{-1}\right]\beta^{(n)}=f^{(n)}+\delta^{(n)}.$$

Отсюда легко находим:

$$\beta^{(n)}-\alpha^{(n)}+(I_n+\Delta_n)^{-1}C_n^{-1}\Gamma_n\left(C_n^*\right)^{-1}\left(\beta^{(n)}-\alpha^{(n)}\right)=$$
$$=(I_n+\Delta_n)^{-1}C_n^{-1}\Gamma_n a^{(n)}+(I_n+\Delta_n)^{-1}C_n^{-1}\delta^{(n)}. \qquad (14.20)$$

---

[1]) Достаточность условий теоремы 14.2 доказана Г. Н. Ясковой и М. Н. Яковлевым [1], необходимость — Г. М. Вайникко [1].

Если координатная система сильно минимальна в H_0, то величины $\|(I_n+\Delta_n)^{-1}\|$ и $\|a^{(n)}\|$ ограничены независимо от n; пусть $\|(I_n+\Delta_n)^{-1}\| \leqslant c = \text{const}$ и $\|a^{(n)}\| \leqslant c' = \text{const}$. Далее, $\|C_n^{-1}\| = \|(C_n^*)^{-1}\| \leqslant \lambda_0^{-\frac{1}{2}}$. Потребуем, чтобы

$$\|\Gamma_n\| < \frac{1}{2}\frac{\lambda_0}{c}.$$

Тогда из равенств (14.19) и (14.20) следует:

$$|v_n - u_n| = \|\beta^{(n)} - \alpha^{(n)}\| < 2\left(cc'\lambda_0^{-\frac{1}{2}}\|\Gamma_n\| + c\lambda_0^{-\frac{1}{2}}\|\delta^{(n)}\|\right),$$

и достаточность условий теоремы 14.2 доказана.

Обратимся к доказательству необходимости. Полагая $\Gamma_n = 0$, найдем из соотношений (14.19) и (14.20):

$$|v_n - u_n| = \|\beta^{(n)} - \alpha^{(n)}\| = \|(I_n + \Delta_n)^{-1} C_n^{-1}\delta^{(n)}\|. \tag{14.21}$$

Обозначим

$$(I_n + \Delta_n)^{-1} C_n^{-1}\delta^{(n)} = \sigma^{(n)},$$

тогда

$$\|C_n^{-1}\delta^{(n)}\| \leqslant \|I_n + \Delta_n\| \, \|\sigma^{(n)}\|.$$

Выше мы видели, что $\Delta_n \to \Delta$ по норме. Отсюда следует существование такой положительной постоянной μ, что $\|I_n + \Delta_n\| \leqslant \mu$. Теперь

$$\|\sigma^{(n)}\| \geqslant \frac{1}{\mu}\|C_n^{-1}\delta^{(n)}\|.$$

Выберем $\delta^{(n)}$ так, чтобы

$$\|C_n^{-1}\delta^{(n)}\| = \|C_n^{-1}\| \, \|\delta^{(n)}\|.$$

Тогда, в силу тождества (14.21),

$$|v_n - u_n| \geqslant \frac{1}{\mu}\|C_n^{-1}\| \, \|\delta^{(n)}\|.$$

Если координатная система не сильно минимальна в пространстве H_0, то

$$\|C_n^{-1}\| = \frac{1}{\sqrt{\lambda_1^{(n)}}} \xrightarrow[n\to\infty]{} \infty,$$

и приближенное решение неустойчиво.

В статье Г. М. Вайникко [1] исследован более общий вопрос о необходимых условиях устойчивости приближенного решения, построенного по методу Галёркина — Петрова.

§ 15. Замечания об использовании не сильно минимальных систем

Ниже, в гл. V, будут указаны способы построения почти ортонормированных или, по крайней мере, сильно минимальных координатных систем для сравнительно широкого круга задач. Однако этот круг далеко не всеобъемлющий, и в ряде случаев придется прибегать к координатным системам, не сильно минимальным. В настоящем параграфе будут приведены некоторые соображения об использовании не сильно минимальных координатных систем; мы ограничимся при этом наиболее простым и важным случаем процесса Ритца.

Итак, пусть координатная система $\{\varphi_n\}$ не сильно минимальна в энергетическом пространстве H_A. Рассмотрим «точную» и «неточную» системы Ритца

$$R_n a^{(n)} = f^{(n)}, \tag{15.1}$$

$$(R_n + \Gamma_n) b^{(n)} = f^{(n)} + \delta^{(n)}. \tag{15.2}$$

Определим наименьшее собственное число $\lambda_1^{(n)}$ матрицы R_n и наибольшее собственное число той же матрицы $\lambda_n^{(n)}$; при достаточно больших n число $\lambda_1^{(n)}$ будет малым. Зададим число β, $0 < \beta < 1$, и потребуем прежде всего, чтобы

$$\| \Gamma_n \| \leqslant \beta \lambda_1^{(n)}. \tag{15.3}$$

Матрица погрешностей возникает из-за того, что энергетические произведения $[\varphi_k, \varphi_j]$, представляющие собой элементы матрицы R_n, вычисляются приближенно; взяв порядок приближения достаточно высоким, мы сможем удовлетворить неравенству (15.3).

Повторяя рассуждения § 9, мы получим формулы, аналогичные формулам (9.13) и (9.13′):

$$\| \eta^{(n)} \| = \| b^{(n)} - a^{(n)} \| \leqslant \frac{\left[\lambda_1^{(n)}\right]^{-\frac{3}{2}} \| \Gamma_n \| \, | u_0 | + \left[\lambda_1^{(n)}\right]^{-1} \| \delta^{(n)} \|}{1 - \beta}, \tag{15.4}$$

$$\| \eta^{(n)} \| = \| b^{(n)} - a^{(n)} \| \leqslant \frac{C \left[\lambda_1^{(n)}\right]^{-\frac{3}{2}} \| \Gamma_n \| + \left[\lambda_1^{(n)}\right]^{-1} \| \delta^{(n)} \|}{1 - \beta}. \tag{15.4′}$$

Здесь, как и в формуле (9.13′), C — какая-нибудь известная верхняя граница для $|u_0|$.

Формулы (15.4) и (15.4′) верны в предположении, что векторы $a^{(n)}$ и $b^{(n)}$ точно удовлетворяют уравнениям (15.1) и (15.2) соответственно; практически это означает, что при решении системы (15.2) следует все промежуточные выкладки производить с достаточно большим числом запасных знаков.

Погрешность «неточного» приближения по Ритцу

$$v_n = \sum_{k=1}^{n} b_k^{(n)} \varphi_k$$

можно оценить по формуле (10.8), которую мы здесь воспроизведем:

$$|v_n - u_n| \leqslant \sqrt{\| \eta^{(n)} \| \left(\| \delta^{(n)} \| + \| \Gamma_n \| \, \| b^{(n)} \| \right)}.$$

Величина $\| b^{(n)} \|$ заранее неизвестна, поэтому мы преобразуем последнюю оценку так:

$$\| b^{(n)} \| \leqslant \| a^{(n)} \| + \| \eta^{(n)} \|.$$

Нетрудно получить формулу

$$\| a^{(n)} \| \leqslant \left[\lambda_1^{(n)}\right]^{-\frac{1}{2}} |u_0| \leqslant C \left[\lambda_1^{(n)}\right]^{-\frac{1}{2}},$$

аналогичную формуле (9.12). Теперь

$$|v_n - u_n| \leqslant \sqrt{\| \eta^{(n)} \| \left\{ \| \delta^{(n)} \| + \| \Gamma_n \| \left(C \left[\lambda_1^{(n)}\right]^{-\frac{1}{2}} + \| \eta^{(n)} \| \right) \right\}}; \qquad (15.5)$$

формулы (15.4) и (15.5) позволяют указать, какая погрешность допустима при вычислении матрицы R_n и вектора $f^{(n)}$.

Из формулы (15.4′) и (15.5) легко усмотреть, каковы должны быть погрешности Γ_n и $\delta^{(n)}$, чтобы погрешность $v_n - u_n$ была достаточно малой. Именно, если допустить, что

$$\| \Gamma_n \| = O\left(\left[\lambda_1^{(n)}\right]^{1+\varkappa} \right), \qquad \| \delta^{(n)} \| = O\left(\left[\lambda_1^{(n)}\right]^{\frac{1}{2}+\varkappa} \right), \qquad (15.6)$$

где $\varkappa$ — положительная постоянная, то, как это вытекает из формул (15.4′) и (15.5),

$$\| \eta^{(n)} \| = O\left(\left[\lambda_1^{(n)}\right]^{\varkappa - \frac{1}{2}} \right) \qquad (15.7)$$

и

$$|v_n - u_n| = O\left(\left[\lambda_1^{(n)}\right]^{\varkappa} \right). \qquad (15.8)$$

Интересно отметить, что при $0 < \varkappa < \frac{1}{2}$ величина $\| \eta^{(n)} \|$ может бесконечно возрастать при $n \to \infty$, тогда как погрешность приближенного решения стремится к нулю.

Из сказанного здесь следует также, что в *случае не сильно минимальной координатной системы можно столбец свободных членов $f^{(n)}$ вычислять с несколько меньшей точностью, чем матрицу Ритца R_n.*

Для примера рассмотрим задачу

$$Au = -\frac{d}{dx}\left(p(x) \frac{du}{dx} \right) + q(x) u = f(x), \qquad u(0) = u(1) = 0, \quad (15.9)$$

где $p(x)$, $p'(x)$, $q(x)$ непрерывны,

$$p_0 \leqslant p(x) \leqslant p_1, \qquad 0 \leqslant q(x) \leqslant q_1$$

и p_0, p_1, q_1 — положительные постоянные. Оператор A — положительно определенный; при этом

$$|u|_A^2 = \int_0^1 (pu'^2 + qu^2)\,dx. \tag{15.10}$$

Заметим, что оператор A — сходный с оператором B, где

$$Bu = -\frac{d^2u}{dx^2}, \quad u(0) = u(1) = 0.$$

Сравним $|u|_A$ и $|u|_B$. Прежде всего,

$$|u|_B^2 = \int_0^1 u'^2\,dx,$$

отсюда $|u|_A^2 \geqslant p_0 |u|_B^2$. Далее,

$$|u|_A^2 \leqslant p_1 \int_0^1 u'^2\,dx + q_1 \int_0^1 u^2\,dx = p_1 |u|_B^2 + q_1 \|u\|^2.$$

Наименьшее собственное число оператора B равно π^2, поэтому, если $u(0) = u(1) = 0$, то $\|u\|^2 \leqslant \frac{1}{\pi^2} |u|_B^2$, и мы получаем искомое соотношение между нормами:

$$p_0 |u|_B^2 \leqslant |u|_A^2 \leqslant (p_1 + \pi^{-2} q_1) |u|_B^2. \tag{15.11}$$

Из формулы (15.11) и ранее отмеченного неравенства $\|u\|^2 \leqslant \pi^{-2} |u|_B^2$ вытекает, что

$$|u|_A^2 \geqslant p_0 \pi^2 \|u\|^2,$$

так что число γ, характеризующее положительную определенность оператора A, может быть взято равным $\sqrt{p_0}\,\pi$.

В качестве координатной возьмем систему функций

$$\varphi_k(x) = x^k (1 - x) \qquad (k = 1, 2, \ldots). \tag{15.12}$$

Докажем, что эта система не сильно минимальна в H_B, для чего в соответствии с формулой (2.9) покажем, что некоторые из сумм

$$\sum_{k=1}^{n} [\sigma_{kj}^{(n)}]^2,$$

где $\sigma_{kj}^{(n)}$ — элементы обратной матрицы Ритца, неограничены. Имеем

$$[\varphi_k, \varphi_j]_B = \int_0^1 [kx^{k-1} - (k+1)x^k][jx^{j-1} - (j+1)x^j]\,dx =$$
$$= \frac{2jk}{(j+k)[(j+k)^2-1]}.$$

Величины $\sigma_{kj}^{(n)}$ удовлетворяют соотношению

$$\sum_{k=1}^{n} \frac{2jk\sigma_{kj}^{(n)}}{(j+k)[(j+k)^2-1]} = 1.$$

По неравенству Коши

$$\sum_{k=1}^{n} [\sigma_{kj}^{(n)}]^2 \sum_{k=1}^{n} \frac{4j^2k^2}{(j+k)^2[(j+k)^2-1]^2} \geqslant 1.$$

При $j=n$ имеем:

$$\sum_{k=1}^{n} \frac{4k^2n^2}{(n+k)^2[(n+k)^2-1]} < \sum_{k=1}^{n} \frac{1}{(n+k)^2}.$$

Но

$$\frac{1}{(n+k)^2} < \int_0^1 \frac{dx}{(n+k-x)^2} = \int_{n+k-1}^{n+k} \frac{dy}{y^2}.$$

Отсюда

$$\sum_{k=1}^{n} \frac{1}{(n+k)^2} < \int_n^{2n} \frac{dy}{y^2} = \frac{1}{2n}$$

и из неравенства (15.11) следует:

$$\sum_{k=1}^{n} [\sigma_{kn}^{(n)}]^2 > 2n.$$

Таким образом, в H_B система (15.12) не сильно минимальна. Из соотношения (15.11) и следствия 4.1 вытекает, что та же система не сильно минимальна и в H_A.

Оценим собственные числа матрицы Ритца $R_n = \| [\varphi_k, \varphi_j]_A \|_{j,\,k=1}^{j,\,k=n}$. Пусть $\mu_k^{(n)}$ означают собственные числа матрицы $r_n = \| [\varphi_k, \varphi_j]_B \|_{j,\,k=1}^{j,\,k=n}$.

Из минимаксимального принципа и из неравенств (15.11) вытекает, что

$$p_0\mu_k^{(n)} \leqslant \lambda_k^{(n)} \leqslant (p_1 + \pi^{-2}q_1)\,\mu_k^{(n)}.$$

В статье автора [17] приведены значения $\mu_1^{(n)}$ и $\mu_n^{(n)}$ для некоторых n; соответствующую табличку мы здесь воспроизводим.

n	$\mu_n^{(n)}$	$\mu_1^{(n)}$	Число обусловленности $\rho(r_n)$
2	0,42769839	0,03896921	10,978
3	0,47572671	0,00270726	175,72
5	0,52695	0,00015	3430
7	0,5552	0,00007	7700

Допустим, что приближенное решение задачи (15.9) строилось в виде:

$$u_7(x)=\sum_{k=1}^{7} a_k^{(7)}\varphi_k(x)=x(1-x)\left(a_1^{(7)}+a_2^{(7)}x+a_3^{(7)}x^2+\ldots+a_7^{(7)}x^6\right).$$

Для определенности последующих вычислений примем, что $p_0=1$, $p_1+\pi^{-2}q_1=10$. Тогда $\lambda_1^{(7)}\geqslant 7\cdot 10^{-5}$. Далее, если u_0 есть решение задачи (15.9), то [1]) $|u_0|_A\leqslant\frac{\|f\|}{\gamma}=\frac{\|f\|}{\pi\sqrt{p_0}}$, и в формуле (15.4′) можно положить $C=\frac{\|f\|}{\pi\sqrt{p_0}}=\frac{\|f\|}{\pi}$. Для той же определенности вычислений допустим, что $\|f\|=\pi$ и $C=1$. Наконец, пусть величины $[\varphi_k, \varphi_j]_A$ и (f, φ_j) вычислены с точностью в k знаков после запятой. Тогда

$$|\gamma_{kj}|<\frac{1}{2}10^{-k},\qquad |\delta_k|<\frac{1}{2}10^{-k}$$

и, следовательно,

$$\|\Gamma_7\|\leqslant 3{,}5\cdot 10^{-k},\qquad \|\delta^{(7)}\|\leqslant\frac{\sqrt{7}}{2}10^{-k}<1{,}5\cdot 10^{-k}.$$

Соотношение (15.3) требует, чтобы $3{,}5\cdot 10^{-k}\leqslant 7\beta\cdot 10^{-5}$; возьмем $\beta=0{,}5$, тогда необходимо $k\geqslant 5$. Теперь по формуле (15.4′)

$$\|\eta^{(7)}\|\leqslant 2\left[7^{-1{,}5}\cdot 3{,}5\cdot 10^{7{,}5-k}+\frac{1{,}5}{7}10^{5-k}\right].$$

Второе слагаемое справа существенно меньше первого [2]) и его можно отбросить без заметного ущерба для точности оценки. Мы примем поэтому, что

$$\|\eta^{(7)}\|<2\cdot 3{,}5\cdot 7^{-1{,}5}\cdot 10^{7{,}5-k}=\sqrt{\frac{10}{7}}10^{7-k}<1{,}2\cdot 10^{7-k}.$$

[1]) См. ПМ, стр. 20—21.

[2]) Это связано с замечанием на стр. 84 о том, что в случае не сильно минимальной координатной системы можно свободные члены в системе Ритца вычислять с меньшей точностью, чем матрицу Ритца.

Обратимся к формуле (15.5). В нашем случае $\left[\lambda_1^{(7)}\right]^{-\frac{1}{2}}$ имеет порядок 10^2; в то же время при наименьшем допустимом значении $k=5$ $\|\eta^{(7)}\|<1{,}2\cdot 10^2$, а при $k>5$ порядок $\|\eta^{(7)}\|$ ниже. Будем считать, что $k>5$, тогда $\|\eta^{(7)}\|<\left[\lambda_1^{(7)}\right]^{-\frac{1}{2}}$, и мы заменим в оценке (15.5) $\|\eta^{(7)}\|$ на $\left[\lambda_1^{(7)}\right]^{-\frac{1}{2}}$. Далее, пренебрежем величиной $\|\delta^{(7)}\|$ по сравнению с существенно большей величиной $\|\Gamma_7\|\left[\lambda_1^{(7)}\right]^{-\frac{1}{2}}$. Мы получим тогда:

$$|v_7-u_7|_A \leqslant \sqrt{2\|\eta^{(7)}\|\|\Gamma_7\|\left[\lambda_1^{(7)}\right]^{-\frac{1}{2}}}<10^{5-k};$$

если, следовательно, нам желательно, чтобы погрешность (в энергетической норме) «неточного» решения v_7 имела порядок, например, 10^{-4}, то энергетические произведения $[\varphi_k, \varphi_j]_A$ следует вычислять с 9 верными десятичными знаками.

Заметим, что этот результат согласуется с формулами (15.6)—(15.8) и может быть получен из них при $\varkappa=0{,}8$.

Заметим еще, что решать систему Ритца итерациями в данном случае нецелесообразно: при взятых нами числовых значениях параметров $\lambda_1^{(n)} \leqslant 10\mu_1^{(n)}$ и $\lambda_n^{(n)} \geqslant \mu_n^{(n)}$, поэтому число обусловленности $\rho(R_n) \geqslant 0{,}1\rho(r_n)$; в частности, $\rho(R_7) \geqslant 770$, и при наивыгоднейшем выборе параметра α (см. § 12) итерации будут сходиться как прогрессия со знаменателем $\frac{669}{771}$.

§ 16. Другая точка зрения на устойчивость

Несколько иная точка зрения на устойчивость вычислительного процесса изложена Б. А. Самокишем в его статье [1]. Этот автор рассматривает вычислительный процесс, который мы назовем «абстрактным процессом Бубнова — Галёркина» [1]) и который состоит в следующем.

Пусть линейный оператор A действует из X в Y, где X и Y — банаховы пространства. Чтобы построить приближенное решение уравнения

$$Ax=y, \qquad x\in X, \qquad y\in Y, \tag{16.1}$$

выбираем систему координатных элементов $\varphi_n \in X$ $(n=1, 2, \ldots)$ и систему линейных функционалов l_n $(n=1, 2, \ldots)$, действующих в пространстве Y. Приближенное решение ищем в виде

$$x_n=\sum_{k=1}^{n} a_k^{(n)}\varphi_k, \tag{16.2}$$

[1]) Б. А. Самокиш называет его «абстрактным методом Галёркина».

а коэффициенты $a_k^{(n)}$ определяем из системы уравнений

$$\sum_{k=1}^{n} \alpha_{jk} a_k^{(n)} = l_j y, \qquad \alpha_{jk} = l_j A\varphi_k \qquad (j = 1, 2, \ldots, n). \tag{16.3}$$

Заметим, что если пространства X и Y — гильбертовы, то «абстрактный процесс Бубнова — Галёркина» совпадает с процессом Петрова — Галёркина.

Б. А. Самокиш вводит две числовые характеристики устойчивости процесса (16.3). Одна из них — число обусловленности матрицы $A_n = \|\alpha_{jk}\|_{j,\,k=1}^{j,\,k=n}$ системы (16.3); в данном случае, когда матрица A_n, вообще говоря, неположительна и даже несимметрична, ее число обусловленности определяется по формуле[1])

$$H(A_n) = \|A_n\| \, \|A_n^{-1}\|. \tag{16.4}$$

Вторая характеристика, обозначаемая через μ_n, строится так. В системе (16.3) заменяем числа $l_j y$ произвольными числами β_j такими, что

$$\sum_{j=1}^{n} |\beta_j|^2 = 1; \tag{16.5}$$

далее, определяем коэффициенты $a_k^{(n)}$ из полученной таким образом системы и элементы x_n из формулы (16.2). Число μ_n определяется формулой

$$\mu_n = \frac{\max \|x_n\|}{\min \|x_n\|}; \tag{16.6}$$

максимум и минимум берутся по всевозможным наборам чисел β_1, $\beta_2, \ldots, \beta_n$, удовлетворяющих соотношению (16.5).

Б. А. Самокиш рассматривает величину $H(A_n)$ как характеристику устойчивости процесса (16.3), иначе говоря, устойчивости коэффициентов $a_k^{(n)}$, а величину μ_n — как характеристику устойчивости приближенного решения (16.2).

Точного определения устойчивости в статье Б. А. Самокиша нет, но насколько можно судить, этот автор во всяком случае считает процесс (16.3) устойчивым, если величина $H(A_n)$ ограничена независимо от n, и практически устойчивым, если эти величины при возрастании n растут не слишком быстро.

Аналогично, если величина μ_n ограничена, то процесс вычисления приближенных решений (16.2) считается устойчивым; если же μ_n растет не слишком быстро, то упомянутый процесс практически устойчив.

Подробнее рассмотрен важнейший случай, когда $X = Y$ и $A = I + T$, где I — тождественный, а T — вполне непрерывный

[1]) См. Д. К. Фаддеев и В. Н. Фаддеева [1].

операторы. В этом случае для величин $H(A_n)$ и μ_n получаются оценки, к описанию которых мы и перейдем.

Через X_n обозначим n-мерное подпространство пространства X, натянутое на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$, через E_n — n-мерное унитарное пространство. Введем в рассмотрение операторы Φ_n и Ψ_n. Первый из них действует из E_n в X_n по формуле

$$\Phi_n a = \sum_{k=1}^{n} a_k \varphi_k, \qquad a = (a_1, a_2, \ldots, a_n) \in E_n, \tag{16.7}$$

а второй действует из X_n в E_n по формуле

$$\Psi_n x = (l_1 x, l_2 x, \ldots, l_n x), \qquad x \in E_n. \tag{16.8}$$

Нетрудно убедиться, что операторы Φ_n и Ψ_n обратимы, если $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы и если функционалы l_j таковы, что равенства

$$l_j x = 0, \qquad x \in X_n \qquad (j = 1, 2, \ldots, n)$$

имеют место только при $x = 0$. Эти условия ниже предполагаются выполненными.

Оператор Ψ_n очевидным образом допускает расширение на все пространство X с сохранением той же формулы (16.8); расширенный таким образом оператор обозначим через $\tilde{\Psi}_n$.

Построим проекционный оператор Π_n, действующий из X в X_n таким образом, что $l_j \Pi_n x = l_j x$ $(j = 1, 2, \ldots, n)$. Этот оператор определяется единственным образом; ему нетрудно дать явное выражение, но на этом мы останавливаться не будем. Наконец, положим $\tilde{A}_n = \Pi_n \tilde{A}$, где $\tilde{A}$ — сужение оператора A на подпространство X_n.

Упомянутые выше оценки имеют вид

$$H(A_n) \leqslant \|A\| \|\tilde{A}_n^{-1}\| \|\tilde{\Psi}_n\| \|\Psi_n^{-1}\| \|\Phi_n\| \|\Phi_n^{-1}\|, \tag{16.9}$$

$$\mu_n \leqslant \|A\| \|\tilde{A}_n^{-1}\| \|\tilde{\Psi}_n\| \|\Psi_n^{-1}\|. \tag{16.10}$$

В случае процесса Ритца X есть энергетическое пространство соответствующей задачи, $A = I$, $l_j x = [x, \varphi_j]$, $A_n = R_n$ и

$$H(A_n) = \rho(R_n) = \frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}.$$

Просто вычисляется в этом случае и величина μ_n. Именно, система (16.3), в которой $l_j y$ заменены на β_j, принимает вид:

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]\, a_k^{(n)} = \beta_j$$

или, короче, $R_n a^{(n)} = \beta$, где $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$.

Далее, норма в рассматриваемом случае есть энергетическая норма, поэтому

$$\max_{\sum|\beta_j|^2=1} \|x_n\| = \max_{\|\beta\|=1} |x_n| = \max \frac{|x_n|}{\|R_n a^{(n)}\|}.$$

Напомним, что

$$|x_n|^2 = \left|\sum_{k=1}^{n} a_k \varphi_k\right|^2 = (R_n a^{(n)},\ a^{(n)}).$$

Это дает нам

$$\max_{\sum|\beta_j|^2=1} \|x_n\|^2 = \max_{a^{(n)}\neq 0} \frac{(R_n a^{(n)},\ a^{(n)})}{\|R_n a^{(n)}\|^2}$$

или, если положить $R_n a^{(n)} = c^{(n)}$,

$$\max_{\sum|\beta_j|^2=1} \|x_n\|^2 = \max_{c^{(n)}\neq 0} \frac{\left(R_n^{-1} c^{(n)},\ c^{(n)}\right)}{\|c^{(n)}\|^2} = \frac{1}{\lambda_1^{(n)}}.$$

Аналогично

$$\min_{\sum|\beta_j|^2=1} \|x_n\|^2 = \min_{c^{(n)}\neq 0} \frac{\left(R_n^{-1} c^{(n)},\ c^{(n)}\right)}{\|c^{(n)}\|^2} = \frac{1}{\lambda_n^{(n)}}.$$

Теперь

$$\mu_n = \frac{\max\|x_n\|}{\min\|x_n\|} = \sqrt{\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}} = \sqrt{\rho(R_n)}. \tag{16.11}$$

Таким образом, требование устойчивости процесса Ритца в смысле Б. А. Самокиша равносильно требованию, чтобы координатная система была почти ортонормирована в энергетической метрике задачи.

В случае общего абстрактного процесса Бубнова — Галёркина формула типа (16.11) не имеет места, и величина μ_n может оказаться ограниченной в то время, как величина $H(A_n)$ неограничена.

В связи с развиваемым Б. А. Самокишем представлением о практической устойчивости или неустойчивости абстрактного процесса Бубнова — Галёркина приобретает большое значение оценка величин $H(A_n)$ и μ_n для конкретных классов координатных элементов $\{\varphi_n\}$ и функционалов l_n. В статье [1] Б. А. Самокиша дан ряд примеров таких оценок; один из них мы приводим ниже, предварительно сделав следующие замечания.

Допустим, что дело идет о процессе Ритца или Бубнова — Галёркина, так что функционалы l_n можно отождествить с элементами φ_n, которые принадлежат гильбертову пространству X. В этом случае $\|\widetilde{\Psi}_n\| = \|\Psi_n\|$. Докажем это. С одной стороны, $\widetilde{\Psi}_n$ есть расширение оператора Ψ_n, поэтому $\|\tilde{\Psi}_n\| \geqslant \|\Psi_n\|$. С другой стороны, если x есть произвольный элемент пространства X и x' — проекция этого элемента на подпространство X_n, натянутое на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$, то $\tilde{\Psi}_n x = \tilde{\Psi}_n x' = \Psi_n x'$, поэтому

$$\|\tilde{\Psi}_n x\| = \|\Psi_n x'\| \leqslant \|\Psi_n\| \, \|x'\| \leqslant \|\Psi_n\| \, \|x\|$$

и, следовательно, $\|\tilde{\Psi}_n\| \leqslant \|\Psi_n\|$. Далее, как легко видеть, в случае процесса Ритца или Бубнова — Галёркина операторы Φ_n и Ψ_n — сопряженные, и потому $\|\Psi_n\| = \|\Phi_n\|$, $\|\Psi_n^{-1}\| = \|\Phi_n^{-1}\|$.

В этом случае $\Phi_n = S_n^{-1}$, где S_n — оператор, введенный в § 13. Из формулы (13.12) легко следует, что

$$\|\Phi_n\| = \|\Psi_n\| = \sqrt{\lambda_n^{(n)}},$$

$$\|\Phi_n^{-1}\| = \|\Psi_n^{-1}\| = \frac{1}{\sqrt{\lambda_1^{(n)}}}.$$

Оценки (16.9) и (16.10) принимают вид

$$H(A_n) \leqslant \|A\| \, \|\tilde{A}_n^{-1}\| \frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}, \qquad \mu_n \leqslant \|A\| \, \|\tilde{A}_n^{-1}\| \sqrt{\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}}.$$

Можно доказать, что в случае процесса Бубнова — Галёркина величина $\|\tilde{A}_n^{-1}\|$ ограничена (для процесса Ритца $\|\tilde{A}_n^{-1}\| = 1$), и мы получаем более простую оценку

$$H(A_n) \leqslant C \frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}, \qquad \mu_n \leqslant C \sqrt{\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}}, \qquad C = \text{const}. \tag{16.12}$$

В случае процесса Ритца, как мы видели, $A = I$, $\tilde{A}_n^{-1} = I$, $\mu_n = \sqrt{\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}}$ и, очевидно, $H(R_n) = \rho(R_n) = \frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}$. Таким образом, дело сводится к оценке отношения $\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}}$.

Рассмотрим теперь в вещественном пространстве $X = L_2(-1, 1)$ систему функций $\varphi_n(t) = t^n$ $(n = 0, 1, \ldots)$. Докажем, что для этой системы $\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}} = O\left((\sqrt{2} + 1)^{2n}\right)$. В данном случае

$$\Phi_n a = \sum_{k=0}^{n-1} a_k t^k = p_{n-1}(t)$$

есть полином степени $n-1$, и нам нужно найти верхнюю и нижнюю границы отношения

$$\frac{\int\limits_{-1}^{1} p_{n-1}^2(t)\,dt}{\sum\limits_{k=0}^{n-1} a_k^2}. \tag{16.13}$$

Очевидно,

$$\int\limits_{-1}^{1} p_{n-1}^2(t)\,dt = \int\limits_{\gamma} p_{n-1}^2(z)\,dz,$$

где γ — верхняя или нижняя полуокружность $|z|=1$, с концами в точках $z=-1$ и $z=+1$. Отсюда

$$\int\limits_{-1}^{1} p_{n-1}^2(t)\,dt = \frac{1}{2}\int\limits_{|z|=1} p_{n-1}^2(z)\,\mathrm{sign}\,\mathrm{Im}\,z\,dz \leqslant$$

$$\leqslant \frac{1}{2}\int\limits_{0}^{2\pi} |p_{n-1}(e^{i\vartheta})|^2\,d\vartheta = \pi\sum_{k=0}^{n-1} a_k^2.$$

Таким образом, отношение (16.13) ограничено числом π сверху, и $\lambda_n^{(n)} \leqslant \pi$.

Отношение (16.13) не меняется при умножении вектора a на постоянную; выберем эту постоянную так, чтобы

$$\int\limits_{-1}^{1} p_{n-1}^2(t)\,dt = 1.$$

В последнем интеграле сделаем подстановку $t=\frac{1}{2}(z+z^{-1})$, $z=e^{i\vartheta}$ и введем обозначение

$$g(z) = \frac{z^{n-1}}{2}\,p_{n-1}\left(\frac{1}{2}(z+z^{-1})\right)\sqrt{1-z^2}.$$

Тогда

$$\int\limits_{|z|=1} |g(z)|^2\,d\vartheta = \int\limits_{0}^{2\pi} p_{n-1}^2(\cos\vartheta)\,|\sin\vartheta|\,d\vartheta = 2\int\limits_{-1}^{1} p_{n-1}^2(t)\,dt = 2.$$

Функция $g(z)$ аналитически продолжима внутрь круга $|z|<1$ и регулярна в нем; пусть при $|z|<1$

$$g(z) = \sum_{k=0}^{\infty} \alpha_k z^k.$$

Тогда

$$|g(z)|^2 \leqslant \sum_{k=0}^{\infty} |\alpha_k|^2 \sum_{k=0}^{\infty} |z_k|^{2k} =$$

$$= \frac{1}{2\pi(1-|z|^2)} \int\limits_{|z|=1} |g(z)|^2 \, d\vartheta = \frac{1}{\pi(1-|z|^2)}.$$

Пусть теперь t — произвольное комплексное число, связанное с z тем же соотношением $t = \frac{1}{2}(z + z^{-1})$, $|z| < 1$. Тогда

$$|p_{n-1}(t)| \leqslant |z|^{-n+1} (1-|z|^2)^{-\frac{1}{4}} |g(z)| \leqslant \frac{1}{\pi} |z|^{-n+1} (1-|z|^2)^{-\frac{5}{4}}.$$

В t-плоскости рассмотрим эллипс с фокусами -1 и $+1$, проходящий через точку $t = i$. На этом эллипсе $|z| = \sqrt{2} - 1$ и, следовательно,

$$|p_{n-1}(t)| \leqslant C_1 (\sqrt{2} + 1)^n, \qquad C_1 = \text{const}.$$

Последняя оценка верна также на окружности $|t| = 1$. Но тогда

$$\sum_{k=0}^{n-1} a_k^2 = \frac{1}{2\pi} \int\limits_{|t|=1} |p_n(t)|^2 \, d\vartheta \leqslant C_2 (\sqrt{2} + 1)^{2n}, \qquad C_2 = \text{const}.$$

Теперь ясно, что отношение (16.13) не меньше, чем величина $C_3 (\sqrt{2} + 1)^{-2n}$, $C_3 = \text{const}$ и, следовательно, $\frac{1}{\lambda_1^{(n)}} \leqslant C_4 (\sqrt{2} + 1)^{2n}$, $C_4 = \text{const}$. Вспоминая, что $\lambda_n^{(n)} \leqslant \pi$, мы найдем, что

$$\frac{\lambda_n^{(n)}}{\lambda_1^{(n)}} \leqslant C (\sqrt{2} + 1)^{2n}, \qquad C = \text{const}.$$

ГЛАВА III

ОБ УСТОЙЧИВОСТИ ПРОЦЕССА БУБНОВА — ГАЛЁРКИНА ДЛЯ НЕСТАЦИОНАРНЫХ ЗАДАЧ

§ 17. Схема процесса Бубнова — Галёркина для нестационарных задач

Мы будем рассматривать в этой главе задачу Коши для нестационарного операторного уравнения

$$\frac{d^2}{dt^2} Au + \frac{d}{dt} Bu + Cu = f(t), \tag{17.1}$$

т. е. задачу интегрирования уравнения (17.1) при начальных условиях

$$u\big|_{t=0} = \varphi, \qquad \frac{du}{dt}\bigg|_{t=0} = \psi. \tag{17.2}$$

Задачи такого типа рассматривал С. Л. Соболев [2] для некоторых случаев, когда A, B, C — дифференциальные операторы; мы будем поэтому называть уравнения вида (17.1) уравнениями С. Л. Соболева.

В последующем будем предполагать, что A, B, C — операторы, действующие в некотором сепарабельном гильбертовом пространстве H; $u = u(t)$ и $f(t)$ — функции от t, значения которых суть элементы пространства H; этому же пространству принадлежат и входящие в начальные условия (17.2) элементы φ и ψ. Если $A \equiv 0$, то второе из условий (17.2) не ставится.

Метод Бубнова — Галёркина применялся к разным случаям задачи (17.1) — (17.2) многими авторами; одно из наиболее полных исследований принадлежит М. И. Вишику [1], который рассмотрел и тот случай, когда операторы A, B, C могут зависеть от времени. Этот общий случай мы оставим в стороне и ограничимся предположением, что упомянутые операторы от времени не зависят.

При известных предположениях, которых мы здесь не будем касаться, задача Коши (17.1) — (17.2) имеет решение, и притом единственное.

Остановимся на процессе построения последовательности приближенных решений этой задачи, аналогичном процессу Бубнова — Галёркина; мы будем называть его процессом Бубнова — Галёркина для нестационарных задач. Более подробно об этом процессе см. статью [1] М. И. Вишика.

Для простоты допустим, что операторы A, B, C — положительно определенные [1]). Далее, примем, что их области определения $D(A)$, $D(B)$, $D(C)$ пересекаются по некоторому множеству D, плотному в данном гильбертовом пространстве H и в каждом из энергетических пространств H_A, H_B, H_C. Тогда эти пространства пересекаются по некоторому множеству $D_0 \supset D$, плотному в каждом из них. Нетрудно видеть, что D_0 станет полным гильбертовым пространством, если в этом множестве ввести скалярное произведение по формуле

$$[u, v]_0 = [u, v]_A + [u, v]_B + [u, v]_C, \tag{17.3}$$

норму в D_0 будем обозначать через $|u|_0$.

Выберем координатную систему

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, \tag{17.4}$$

подчиненную следующим условиям: 1) $\varphi_n \in D_0$ $(n = 1, 2, \ldots)$; 2) элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы при любом n; 3) система (17.4) полна в D_0. Заметим, что из последнего требования вытекает полнота системы (17.4) в любом из пространств H_A, H_B, H_C.

Примем, что элементы φ и ψ, входящие в начальные условия (17.2), принадлежат пространству D_0.

Будем искать приближенное решение задачи (17.1) — (17.2) в виде

$$u_n(t) = \sum_{k=1}^{n} a_k^{(n)}(t)\,\varphi_k, \tag{17.5}$$

входящие в формулу (17.5) неизвестные функции $a_k^{(n)}(t)$ будем определять из системы дифференциальных уравнений

$$\sum_{k=1}^{n} \left\{ [\varphi_k, \varphi_j]_A \ddot{a}_k^{(n)} + [\varphi_k, \varphi_j]_B \dot{a}_k^{(n)} + [\varphi_k, \varphi_j]_C a_k^{(n)} \right\} = (f(t), \varphi_j)$$
$$(j = 1, 2, \ldots, n), \tag{17.6}$$

где точка означает дифференцирование по t, и из начальных условий

$$a_k^{(n)}(0) = \alpha_k^{(n)}, \qquad \dot{a}_k^{(n)}(0) = \beta_k^{(n)} \qquad (k = 1, 2, \ldots, n). \tag{17.7}$$

---

[1]) В последующих параграфах настоящей главы мы будем рассматривать и такие случаи, когда некоторые из операторов A, B, C только неотрицательны.

Выбор постоянных $\alpha_k^{(n)}$ и $\beta_k^{(n)}$ в значительной мере произволен и подчинен только одному требованию:

$$\left|\varphi - \sum_{k=1}^{n} \alpha_k^{(n)}\varphi_k\right|_0 \xrightarrow[n\to\infty]{} 0, \qquad \left|\psi - \sum_{k=1}^{n} \beta_k^{(n)}\varphi_k\right|_0 \xrightarrow[n\to\infty]{} 0. \quad (17.8)$$

Можно, например, потребовать, чтобы эти постоянные были определены из условия, чтобы суммы

$$\sum_{k=1}^{n} \alpha_k^{(n)}\varphi_k, \qquad \sum_{k=1}^{n} \beta_k^{(n)}\varphi_k$$

были соответственно проекциями элементов φ и ψ на подпространство пространства D_0, натянутое на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$. В таком случае $\alpha_k^{(n)}$ и $\beta_k^{(n)}$ находятся из следующих систем Ритца:

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]_0 \alpha_k^{(n)} = [\varphi, \varphi_j]_0 \qquad (j = 1, 2, \ldots, n), \qquad (17.9')$$

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]_0 \beta_k^{(n)} = [\psi, \varphi_j]_0 \qquad (j = 1, 2, \ldots, n). \qquad (17.9'')$$

Ниже мы будем иметь дело с теми случаями, когда одно из пространств H_A, H_B, H_C вкладывается в каждое из двух других; если, например, $H_A \subset H_B$ и $H_A \subset H_C$, то метрики пространств H_A и D_0 эквивалентны, в пространстве D_0 можно вместо метрики (17.3) ввести метрику пространства H_A. Тогда $D_0 \equiv H_A$ и системы (17.9′) и (17.9″) заменяются следующими, более простыми по строению их матриц и свободных членов:

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]_A \alpha_k^{(n)} = [\varphi, \varphi_j]_A \qquad (j = 1, 2, \ldots, n), \qquad (17.10')$$

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]_A \beta_k^{(n)} = [\psi, \varphi_j]_A \qquad (j = 1, 2, \ldots, n). \qquad (17.10'')$$

Поясним несколько строение системы дифференциальных уравнений (17.6). Формально ее можно получить так: допустим, что координатные элементы принадлежат множеству D, иначе говоря, что они принадлежат областям определения всех трех операторов A, B, C сразу. Тогда выражение (17.5) можно подставить в левую часть уравнения (17.1). Потребуем (как обычно в процессе Бубнова — Галёркина), чтобы разность между результатом этой подстановки и функцией $f(t)$ была ортогональна в метрике исходного пространства H

к элементам $\varphi_1, \varphi_2, \ldots, \varphi_n$. Мы получаем тогда следующую систему дифференциальных уравнений:

$$\sum_{k=1}^{n} \left\{ (A\varphi_k, \varphi_j)\, \ddot{a}_k^{(n)} + (B\varphi_k, \varphi_j)\, \dot{a}_k^{(n)} + (C\varphi_k, \varphi_j)\, a_k^{(n)} \right\} = (f, \varphi_j)$$
$$(j = 1, 2, \ldots, n).$$

Но $(A\varphi_k, \varphi_j) = [\varphi_k, \varphi_j]_A$ и т. д., и мы приходим к системе (17.6).

Задача Коши (17.6) — (17.7) имеет единственное решение; подставив его в формулу (17.5), получим элемент, который назовем приближенным решением задачи (17.1) — (17.2).

При некоторых довольно сложно формулируемых условиях [1]) приближенное решение $u_n(t)$ стремится в некотором смысле к точному решению $u(t)$; мы не станем формулировать эти результаты, так как для исследования устойчивости процесса Бубнова — Галёркина они нам не понадобятся.

Обратимся теперь к понятию устойчивости процесса Бубнова — Галёркина для нестационарных задач. Введем следующие обозначения для матриц:

$$R_{A_n} = \left\| [\varphi_k, \varphi_j]_A \right\|_{j,\ k=1}^{j,\ k=n}, \qquad R_{B_n} = \left\| [\varphi_k, \varphi_j]_B \right\|_{j,\ k=1}^{j,\ k=n},$$
$$R_{C_n} = \left\| [\varphi_k, \varphi_j]_C \right\|_{j,\ k=1}^{j,\ k=n}$$

и для векторов:

$$a^{(n)}(t) = \left\{ a_1^{(n)}(t), a_2^{(n)}(t), \ldots, a_n^{(n)}(t) \right\},$$
$$f^{(n)}(t) = \left\{ (f(t), \varphi_1), (f(t), \varphi_2), \ldots, (f(t), \varphi_n) \right\},$$
$$\alpha^{(n)} = \left\{ \alpha_1^{(n)}, \alpha_2^{(n)}, \ldots, \alpha_n^{(n)} \right\},$$
$$\beta^{(n)} = \left\{ \beta_1^{(n)}, \beta_2^{(n)}, \ldots, \beta_n^{(n)} \right\}.$$

Уравнения (17.6) и (17.7) запишутся в виде

$$R_{A_n} \ddot{a}^{(n)}(t) + R_{B_n} \dot{a}^{(n)}(t) + R_{C_n} a^{(n)}(t) = f^{(n)}(t), \tag{17.6'}$$

$$a^{(n)}(0) = \alpha^{(n)}, \qquad \dot{a}^{(n)}(0) = \beta^{(n)}. \tag{17.7'}$$

Допустим, что матрицы R_{A_n}, R_{B_n}, R_{C_n} вычислены с погрешностями Γ_{A_n}, Γ_{B_n}, Γ_{C_n}, а векторы $f^{(n)}(t)$, $\alpha^{(n)}$, $\beta^{(n)}$ — с погрешностями $\delta^{(n)}(t)$, $\delta_0^{(n)}$, $\delta_1^{(n)}$ соответственно; здесь Γ_{A_n}, Γ_{B_n}, Γ_{C_n} — симметричные матрицы порядка n, а $\delta^{(n)}(t)$, $\delta_0^{(n)}$, $\delta_1^{(n)}$ — n-компонентные векторы.

---

[1]) См. М. И. Вишик [1].

Вместо задачи (17.6′)—(17.7′) мы на самом деле будем тогда решать следующую задачу:

$$\left(R_{A_n}+\Gamma_{A_n}\right)\ddot{b}^{(n)}(t)+\left(R_{B_n}+\Gamma_{B_n}\right)\dot{b}^{(n)}(t)+\left(R_{C_n}+\Gamma_{C_n}\right)b^{(n)}(t)=$$
$$=f^{(n)}(t)+\delta^{(n)}(t), \qquad (17.11)$$

$$b^{(n)}(0)=\alpha^{(n)}+\delta_0^{(n)}, \qquad \dot{b}^{(n)}(0)=\beta^{(n)}+\delta_1^{(n)}. \qquad (17.12)$$

Понятие об устойчивости, введенное в § 13, в данном случае оказывается недостаточным, так как оно опирается на допущение об ограниченности оператора погрешности Γ_n; нам же придется здесь допустить и неограниченные операторы Γ_n. Поэтому мы изменим понятие об устойчивости следующим образом.

Как и в § 13, рассмотрим две последовательности банаховых пространств X_n и Y_n и две последовательности уравнений

$$A_n x^{(n)}=y^{(n)} \qquad (17.13)$$

и

$$(A_n+\Gamma_n)\,z^{(n)}=y^{(n)}+\Delta^{(n)}; \qquad (17.14)$$

по-прежнему будем считать, что оператор A_n действует из X_n в Y_n, а оператор A_n^{-1} существует при всех n и определен на всем пространстве Y_n. Будем говорить, что процесс (17.13) устойчив, если существуют положительные числа p, q, r такие, что при $\|A_n^{-1}\Gamma_n\|\leqslant r$ уравнение (17.14) разрешимо и выполняется неравенство

$$\|z^{(n)}-x^{(n)}\|\leqslant p\,\|A_n^{-1}\Gamma_n\|+q\,\|\Delta^{(n)}\|. \qquad (17.15)$$

Теорема 17.1. *Для устойчивости процесса* (17.13) *необходимо, чтобы* $\|A_n^{-1}\|\leqslant C_1$, *и достаточно, чтобы* $\|A_n^{-1}\|\leqslant C_1$, $\|A_n^{-1}y^{(n)}\|=\|x^{(n)}\|\leqslant C_2$, *где* C_1 *и* C_2 *не зависят от* n.

Необходимость условия $\|A_n^{-1}\|\leqslant C_1$ доказывается так же, как в теореме 9.1. Чтобы доказать достаточность наших условий, возьмем какое-либо число r, $0<r<1$, и потребуем, чтобы $\|A_n^{-1}\Gamma_n\|\leqslant r$. Теперь вводим уравнения

$$\left(A_n+\Gamma_n\right)z_1^{(n)}=y^{(n)}, \qquad \left(A_n+\Gamma_n\right)z_2^{(n)}=\Delta^{(n)},$$

так что $z^{(n)}-x^{(n)}=\eta^{(n)}=\left(z_1^{(n)}-x^{(n)}\right)+z_2^{(n)}$. Имеем теперь:

$$z_2^{(n)}=\left(I_n+A_n^{-1}\Gamma_n\right)^{-1}A_n^{-1}\Delta^{(n)},$$

откуда

$$\|z_2^{(n)}\|\leqslant\frac{C_1}{1-r}\,\|\Delta^{(n)}\|.$$

Далее, легко видеть, что

$$z_1^{(n)}-x^{(n)}=-\left(I_n+A_n^{-1}\Gamma_n\right)^{-1}A_n^{-1}\Gamma_n x^{(n)}$$

и, следовательно,

$$\| z_1^{(n)} - x^{(n)} \| \leqslant \frac{1}{1-r} \| A_n^{-1}\Gamma_n \| \, \| x^{(n)} \| \leqslant \frac{C_2}{1-r} \| A_n^{-1}\Gamma_n \|.$$

Окончательно

$$\| z^{(n)} - x^{(n)} \| \leqslant \| z_1^{(n)} - x^{(n)} \| + \| z_2^{(n)} \| \leqslant$$
$$\leqslant \frac{C_2}{1-r} \| A_n^{-1}\Gamma_n \| + \frac{C_1}{1-r} \| \Delta^{(n)} \|,$$

и неравенство (17.15) выполнено при $p = \frac{C_2}{1-r}$, $q = \frac{C_1}{1-r}$.

Понятие устойчивости процесса (17.6′) — (17.7′) установим в соответствии с введенным только что общим понятием.

Через X_n обозначим пространство n-компонентных векторных функций, непрерывных в промежутке $0 \leqslant t \leqslant l$, где l — некоторое фиксированное число, которое может быть и бесконечным. Норма в X_n определяется следующей формулой: если $a^{(n)} = a^{(n)}(t)$ и $a^{(n)} \in X_n$, то

$$\| a^{(n)} \|_{X_n} = \max_{0 \leqslant t \leqslant l} \| a^{(n)}(t) \|_{E_n}. \tag{17.16}$$

За Y_n примем пространство троек n-компонентных векторов вида

$$g^{(n)} = (h^{(n)}(t), \ \rho^{(n)}, \ \pi^{(n)}),$$

где $h^{(n)}(t)$ суть непрерывные функции от t, $0 \leqslant t \leqslant l$, а $\rho^{(n)}$ и $\pi^{(n)}$ — постоянные; норму в Y_n зададим формулой

$$\| g^{(n)} \|_{Y_n} = \max_{0 \leqslant t \leqslant l} \| h^{(n)}(t) \|_{E_n} + \| \rho^{(n)} \|_{E_n} + \| \pi^{(n)} \|_{E_n}. \tag{17.17}$$

При $l = \infty$ символ «max» в формулах (17.16) и (17.17) следует заменить на «sup».

Уравнения (17.6′) и (17.7′) определяют некоторый оператор A_n, действующий из X_n в Y_n и переводящий элемент $a^{(n)} = a^{(n)}(t)$ в элемент $g^{(n)} = \left(f^{(n)}(t), \ \alpha^{(n)}, \ \beta^{(n)}\right)$. Уравнения (17.6′) и (17.7′) можно записать в виде одного уравнения (17.13), в котором $x^{(n)}$ и $y^{(n)}$ заменены на $a^{(n)}$ и $g^{(n)}$.

Введем в рассмотрение еще оператор Γ_n, который действует также из X_n в Y_n и который преобразует любой элемент $a^{(n)} = a^{(n)}(t) \in X_n$ в элемент $\zeta^{(n)} \in Y_n$ вида

$$\zeta^{(n)} = \left(\zeta^{(n)}(t), \ 0, \ 0\right),$$

где

$$\zeta^{(n)}(t) = \Gamma_{A_n} \ddot{a}^{(n)}(t) + \Gamma_{B_n} \dot{a}^{(n)}(t) + \Gamma_{C_n} a^{(n)}(t).$$

Теперь уравнения (17.11) и (17.12) можно также записать в виде одного уравнения (17.14), в котором

$$z^{(n)} = b^{(n)}, \qquad \Delta^{(n)} = \left(\delta^{(n)}(t),\ \delta_0^{(n)},\ \delta_1^{(n)}\right).$$

В соответствии с данным выше общим определением будем называть процесс Бубнова — Галёркина для нестационарных задач (процесс (17.6′) — (17.7′)) устойчивым в промежутке $0 \leqslant t \leqslant l$, если существуют такие независящие от n постоянные p, q, r, что при $\| A_n^{-1}\Gamma_n \| \leqslant r$ имеет место неравенство вида (17.15):

$$\| \eta^{(n)} \| = \| b^{(n)} - a^{(n)} \|_{X_n} \leqslant p \| A_n^{-1}\Gamma_n \|_{X_n \to Y_n} + q \| \Delta^{(n)} \|_{Y_n}. \quad (17.15')$$

Устойчивость процесса Бубнова — Галёркина для нестационарных задач была исследована М. А. Велиевым [1] — [2], результаты которого, несколько усиленные и обобщенные, мы ниже в основном и излагаем. Мы будем при этом исходить из теоремы 17.1, в соответствии с которой для доказательства устойчивости процесса (17.6) — (17.7) нам будет достаточно установить ограниченность норм операторов A_n^{-1} и элементов $x^{(n)} = A_n^{-1} y^{(n)}$.

§ 18. Уравнения параболического типа

Будем рассматривать задачу об интегрировании уравнения

$$\frac{du}{dt} + Cu = f(t) \qquad (18.1)$$

при начальном условии

$$u|_{t=0} = \varphi. \qquad (18.2)$$

Здесь, как и в § 17, $u(t)$ и $f(t)$ суть функции от t, значения которых суть элементы гильбертова пространства H, а C — самосопряженный и положительно определенный в этом пространстве оператор. В обозначениях § 17 имеем $A = 0$, $B = I$; отсюда, как легко видеть, следует, что пространство D_0 состоит из тех же элементов, что и пространство H_C, и что $| u |_0^2 = \| u \|^2 + | u |_C^2$. Примем, что $\varphi \in H_C$ и что функция $f(t)$ непрерывна в промежутке $0 \leqslant t \leqslant l$, где l — некоторое положительное число, которое может быть и бесконечным; если $l = \infty$, то будем требовать, чтобы

$$\sup_{0 \leqslant t < \infty} \| f(t) \| < \infty. \qquad (18.3)$$

В данном случае надо потребовать, чтобы координатные элементы $\varphi_n \in H_C$; будем считать выполненными и остальные требования, которые были наложены на координатную систему в § 17. Обозначим через ρ_n матрицу $\rho_n = \| (\varphi_k, \varphi_j) \|_{j,\,k=1}^{j,\,k=n}$, где круглые скобки обозначают скалярное умножение в пространстве H. Система Бубнова —

Галёркина (система (17.6′) — (17.7′)) для задачи (18.1) — (18.2) имеет вид:

$$\rho_n \dot{a}^{(n)}(t) + R_n a^{(n)}(t) = f^{(n)}(t), \qquad R_n = R_{C_n}, \tag{18.4}$$

$$a^{(n)}(0) = \alpha^{(n)}. \tag{18.5}$$

Вектор $\alpha^{(n)}$ мы определим из условия, чтобы элемент

$$\sum_{k=1}^{n} \alpha_k^{(n)} \varphi_k \tag{18.6}$$

был проекцией элемента φ не в пространство D_0, как в § 17, а в пространство H_C. При этом условие (17.8) будет выполнено, а для определения вектора $\alpha^{(n)}$ придется решать не систему (17.9′), а более простую систему

$$\sum_{k=1}^{n} [\varphi_k, \varphi_j]_C \, \alpha_k^{(n)} = [\varphi, \varphi_j]_C. \tag{18.7}$$

Для процесса (18.4) — (18.5) упрощается определение пространства Y_n. Зададимся некоторым промежутком $0 \leqslant t \leqslant l$, где $l \leqslant \infty$; если $l = \infty$, то будем считать, что $0 \leqslant t < \infty$. За Y_n примем пространство пар векторов вида $g^{(n)} = (h^{(n)}(t), \beta^{(n)})$, где $\beta^{(n)}$ — постоянный вектор, а $h^{(n)}(t)$ непрерывен как функция от t на отрезке $0 \leqslant t \leqslant l$; норму в Y_n зададим формулой

$$\| g^{(n)} \|_{Y_n} = \max_{0 \leqslant t \leqslant l} \| h^{(n)}(t) \|_{E_n} + \| \beta^{(n)} \|_{E_n}. \tag{18.8}$$

Оператор A_n действует из X_n в Y_n и переводит вектор $a^{(n)}(t)$ в пару векторов, один из которых равен $\rho_n \dot{a}^{(n)}(t) + R_n a^{(n)}(t)$, а другой равен $a^{(n)}(0)$.

Теорема 18.1. *Если координатная система почти ортонормирована в H, то процесс* (18.4) — (18.5) *устойчив в промежутке* $0 \leqslant t \leqslant l$, *в котором*

$$\sup_{0 \leqslant t \leqslant l} \| f(t) \|_H < \infty.$$

Будучи в H почти ортонормированной, координатная система также и сильно минимальна в этом пространстве; в силу следствия 4.2, эта система сильно минимальна и в H_C.

Эрмитовы матрицы ρ_n и R_n — положительно определенные, поэтому их можно одновременно привести к диагональному виду: существует такая неособенная матрица Z_n, что [1])

$$Z_n^* \rho_n Z_n = I_n, \qquad Z_n^* R_n Z_n = D_n, \tag{18.9}$$

[1]) См., например, Ф. Р. Гантмахер [1], гл. X, § 6.

где D_n — диагональная матрица:

$$D_n = \left[\nu_1^{(n)},\ \nu_2^{(n)},\ \ldots,\ \nu_n^{(n)}\right], \tag{18.10}$$

и положительные числа $\nu_k^{(n)}$ суть корни уравнения

$$\operatorname{Det}(R_n - \nu\rho_n) = 0. \tag{18.11}$$

Важно отметить, что числа $\nu_k^{(n)}$ положительно ограничены снизу. Действительно, наименьшее из них определяется формулой [1])

$$\nu_1^{(n)} = \min \frac{(R_n x^{(n)},\ x^{(n)})}{(\rho_n x^{(n)},\ x^{(n)})},$$

где минимум берется по всем n-компонентным векторам $x^{(n)} = \left(x_1^{(n)},\ x_2^{(n)},\ \ldots,\ x_n^{(n)}\right)$, отличным от нулевого. Положим

$$u = \sum_{k=1}^{n} x_k^{(n)} \varphi_k,$$

где φ_k — координатные элементы. Тогда

$$(R_n x^{(n)},\ x^{(n)}) = |u|_C^2, \qquad (\rho_n x^{(n)},\ x^{(n)}) = \|u\|^2,$$

$$\nu_1^{(n)} = \min_{x_n \neq 0} \frac{|u|_C^2}{\|u\|^2}.$$

Оператор C — положительно определенный и потому существует такая постоянная $\nu_0 > 0$, что для любого элемента $u \in H_C$ верно неравенство $|u|_C^2 \geqslant \nu_0 \|u\|^2$. Теперь ясно, что

$$\nu_1^{(n)} \geqslant \nu_0. \tag{18.12}$$

Положим $Z_n^{-1} = G_n$, тогда из формул (18.9) следует:

$$\rho_n = G_n^* G_n, \qquad R_n = G_n^* D_n G_n. \tag{18.13}$$

Нормы матриц G_n и G_n^{-1} ограничены. Действительно, пусть $\mu_1^{(n)} \leqslant \mu_2^{(n)} \leqslant \ldots \leqslant \mu_n^{(n)}$ — собственные числа матрицы ρ_n. Координатная система в H почти ортонормирована, поэтому существуют положительные постоянные μ_0 и M_0 такие, что $\mu_0 \leqslant \mu_k^{(n)} \leqslant M_0$. Далее, как легко видеть,

$$\|G_n\| = \sqrt{\|\rho_n\|} = \sqrt{\mu_n^{(n)}}, \qquad \|G_n^{-1}\| = \sqrt{\|\rho_n^{-1}\|} = \frac{1}{\sqrt{\mu_1^{(n)}}};$$

отсюда

$$\|G_n\| = \|G_n^*\| \leqslant \sqrt{M_0}, \qquad \|G_n^{-1}\| = \left\|(G_n^*)^{-1}\right\| \leqslant \frac{1}{\sqrt{\mu_0}}. \tag{18.14}$$

---

[1]) См. Ф. Р. Гантмахер [1], гл. X, § 7.

Полагая

$$G_n a^{(n)}(t) = c^{(n)}(t), \qquad G_n \alpha^{(n)} = \gamma^{(n)}, \qquad G_n^{*-1} f^{(n)}(t) = F^{(n)}(t), \tag{18.15}$$

мы приведем уравнения (18.4) и (18.5) к виду:

$$\left.\begin{aligned} \dot{c}^{(n)}(t) + D_n c^{(n)}(t) &= F^{(n)}(t), \\ c^{(n)}(0) &= \gamma^{(n)}. \end{aligned}\right\} \tag{18.16}$$

Система (18.16) распадается на n независимых уравнений:

$$\left.\begin{aligned} \dot{c}_k^{(n)}(t) + \nu_k^{(n)} c_k^{(n)}(t) &= F_k^{(n)}(t), \\ c_k^{(n)}(0) &= \gamma_k^{(n)}, \end{aligned}\right\} \tag{18.16$'$}$$

решения которых даются формулой

$$c_k^{(n)}(t) = \gamma_k^{(n)} e^{-\nu_k^{(n)} t} + \int_0^t F_k^{(n)}(\tau) \exp\left(-\nu_k^{(n)}(t-\tau)\right) d\tau \tag{18.17}$$

$$(k = 1, 2, \ldots, n).$$

Ближайшая наша задача — оценить сверху норму $\| A_n^{-1} \|$. Будем поэтому считать, что в системе (18.4) $f^{(n)}(t)$ есть произвольный вектор из пространства Y_n, и в этом предположении оценим величину $\| a^{(n)} \|_{Y_n}$.

В матричной записи формула (18.17) имеет вид:

$$c^{(n)}(t) = e^{-D_n t} \gamma^{(n)} + \int_0^t e^{-D_n (t-\tau)} F^{(n)}(\tau)\, d\tau. \tag{18.17$'$}$$

Обе части этого равенства умножим слева на G_n^{-1} и заменим $F^{(n)}(\tau)$ и $\gamma^{(n)}$ по формулам (18.15). Мы получим тогда:

$$a^{(n)}(t) = G_n^{-1} e^{-D_n t} G_n \alpha^{(n)} + \int_0^t G_n^{-1} e^{-D_n (t-\tau)} G_n^{*-1} f^{(n)}(\tau)\, d\tau. \tag{18.18}$$

Заметим, что

$$\| e^{-D_n t} \| = e^{-\mu_1^{(n)} t} \leqslant 1.$$

Теперь, полагая, что $0 \leqslant t \leqslant l$, где $l \leqslant \infty$, имеем:

$$\| a^{(n)}(t) \|_{E_n} \leqslant \sqrt{\frac{M_0}{\mu_0}} \| \alpha^{(n)} \|_{E_n} + \frac{1}{\mu_0} \max_{0 \leqslant t \leqslant l} \| f^{(n)}(t) \|_{E_n} \int_0^t e^{-\nu_0 t}\, dt <$$

$$< \sqrt{\frac{M_0}{\mu_0}} \| \alpha^{(n)} \|_{E_n} + \frac{1}{\mu_0 \nu_0} \max_{0 \leqslant t \leqslant l} \| f^{(n)}(t) \|_{E_n}. \tag{18.19}$$

Обозначим через $g^{(n)}$ пару векторов $\left(f^{(n)}(t),\ \alpha^{(n)}\right)$ и через q_0 большее из чисел $\sqrt{\frac{M_0}{\mu_0}}$ и $\frac{1}{\mu_0 \nu_0}$. Тогда

$$\| a^{(n)} \|_{X_n} = \max_{0 \leqslant t \leqslant l} \| a^{(n)}(t) \|_{E_n} \leqslant q_0 \| g^{(n)} \|_{Y_n}.$$

Последнее неравенство означает, что

$$\| A_n^{-1} \| \leqslant q_0 \tag{18.20}$$

и нормы $\| A_n^{-1} \|$ ограничены независимо от n.

Возвратим теперь символу $f^{(n)}(t)$ его первоначальное значение:

$$f^{(n)}(t) = ((f(t),\ \varphi_1),\quad (f(t),\ \varphi_2),\ \ldots,\ (f(t),\ \varphi_n)).$$

Докажем, что

$$\| f^{(n)}(t) \|_{E_n} \leqslant L, \tag{18.21}$$

где L — постоянная, которая не зависит ни от t, ни от n. Поставим задачу о наилучшем приближении в метрике пространства H элемента $f(t)$ линейной комбинацией вида

$$\sum_{k=1}^{n} \alpha_k^{(n)}(t)\, \varphi_k.$$

Тогда $\alpha_k^{(n)}(t)$ удовлетворяют системе (см. § 5) $\rho_n \alpha^{(n)}(t) = f^{(n)}(t)$, где $\alpha^{(n)}(t)$ — вектор $\left(\alpha_1^{(n)}(t),\ \alpha_2^{(n)}(t),\ \ldots,\ \alpha_n^{(n)}(t)\right)$. По формуле (5.6)

$$\| \alpha^{(n)}(t) \|_{E_n} \leqslant \frac{1}{\sqrt{\mu_0}} \| f(t) \|_H \leqslant \frac{1}{\sqrt{\mu_0}} \max_{0 \leqslant t \leqslant l} \| f(t) \|_H.$$

Теперь

$$\| f^{(n)}(t) \|_{E_n} \leqslant \| \rho_n \| \| \alpha^{(n)}(t) \| =$$

$$= \mu_n^{(n)} \| \alpha^{(n)}(t) \| \leqslant \frac{M_0}{\sqrt{\mu_0}} \max \| f(t) \|_H, \tag{18.22}$$

и неравенство (18.21) доказано.

Вектор $\alpha^{(n)}$ решает задачу о наилучшем приближении элемента φ в метрике пространства H_C линейной комбинацией координатных элементов $\varphi_1,\ \varphi_2,\ \ldots,\ \varphi_n$. По той же формуле (5.6)

$$\| \alpha^{(n)} \|_{E_n} \leqslant \frac{1}{\sqrt{\lambda_0}} | \varphi |_C, \tag{18.23}$$

где λ_0 — положительная нижняя граница собственных чисел матрицы R_n.

Воспользуемся введенным выше обозначением $g^{(n)} = \left(f^{(n)}(t),\ \alpha^{(n)}\right)$. Неравенства (18.21) и (18.23) показывают, что

$$\| g^{(n)} \|_{Y_n} \leqslant L_1 = \text{const}.$$

Но тогда

$$\| a^{(n)} \|_{X_n} = \| A_n^{-1} g^{(n)} \|_{X_n} \leqslant \| A_n^{-1} \| \| g^{(n)} \|_{Y_n} \leqslant q_0 L_1. \tag{18.24}$$

Из неравенств (18.20) и (18.24) и теоремы 17.1 вытекает устойчивость процесса (18.4) — (18.5), выражаемая формулой (17.15′).

Неравенство, аналогичное неравенству (17.15′), легко получить и для погрешности приближенного решения, построенного процессом Бубнова — Галёркина. Действительно, пусть «точное» и «неточное» приближения суть соответственно

$$u_n(t) = \sum_{k=1}^{n} a_k^{(n)}(t)\,\varphi_k,$$

$$v_n(t) = \sum_{k=1}^{n} b_k^{(n)}(t)\,\varphi_k.$$

Тогда

$$\| v_n(t) - u_n(t) \|_H^2 = \left\| \sum_{k=1}^{n} \left[b_k^{(n)}(t) - a_k^{(n)}(t) \right] \varphi_k \right\|_H^2 =$$

$$= \left\| \sum_{k=1}^{n} \eta_k^{(n)}(t)\,\varphi_k \right\|_H^2 = \left(\rho_n \eta^{(n)}(t),\ \eta^{(n)}(t)\right)_{E_n} \leqslant$$

$$\leqslant \mu_n^{(n)} \| \eta^{(n)}(t) \|_{E_n}^2 \leqslant M_0 \| \eta^{(n)}(t) \|_{E_n}^2,$$

и, по неравенству (17.15′),

$$\max_{0 \leqslant t \leqslant l} \| v_n(t) - u_n(t) \|_H \leqslant$$

$$\leqslant \sqrt{M_0} \left\{ p \| A_n^{-1} \Gamma_n \|_{X_n \to Y_n} + q \| \Delta^{(n)} \|_{Y_n} \right\}. \tag{18.25}$$

Оценим величину $\| A_n^{-1} \Gamma_n \|_{X_n \to Y_n}$. Обозначим через Γ_{B_n} и Γ_{C_n} погрешности матриц ρ_n и R_n соответственно.

Следуя определению, данному в § 17, положим для $x^{(n)} = x^{(n)}(t)$

$$\Gamma_n x^{(n)} = (\zeta^{(n)}(t),\ 0),$$

где

$$\zeta^{(n)}(t) = \Gamma_{B_n} \dot{x}^{(n)}(t) + \Gamma_{C_n} x^{(n)}(t).$$

Если обозначить $A_n^{-1} \Gamma_n x^{(n)} = \xi^{(n)}(t)$, то $\xi^{(n)}(t)$ есть решение задачи Коши

$$\rho_n \dot{\xi}^{(n)}(t) + R_n \xi^{(n)}(t) = \Gamma_{B_n} \dot{x}^{(n)}(t) + \Gamma_{C_n} x^{(n)}(t),$$

$$\xi^{(n)}(0) = 0.$$

По формуле (18.18)

$$\xi^{(n)}(t)=\int_0^t G_n^{-1}e^{-D_n(t-\tau)}G_n^{*^{-1}}\left[\Gamma_{B_n}\dot{x}^{(n)}(\tau)+\Gamma_{C_n}x^{(n)}(\tau)\right]d\tau.$$

Первый из интегралов, на который распадается правая часть последней формулы, возьмем по частям:

$$\xi^{(n)}(t)=\rho_n^{-1}\Gamma_{B_n}x^{(n)}(t)-\int_0^t G_n^{-1}D_ne^{-D_n(t-\tau)}G_n^{*^{-1}}\Gamma_{B_n}x^{(n)}(\tau)\,d\tau+$$
$$+\int_0^t G_n^{-1}e^{-D_n(t-\tau)}G_n^{*^{-1}}\Gamma_{C_n}x^{(n)}(\tau)\,d\tau-G_n^{-1}e^{-D_nt}G_n^{*^{-1}}\Gamma_{B_n}x_n(0).$$

Отсюда

$$\|\xi^{(n)}(t)\|_{E_n}\leqslant\frac{1}{\mu_0}\|\Gamma_{B_n}\|\,\|x^{(n)}(t)\|_{E_n}+$$
$$+\frac{1}{\mu_0}\|\Gamma_{B_n}\|\max_{0\leqslant t\leqslant l}\|x^{(n)}(t)\|_{E_n}\int_0^t\|D_ne^{-D_n\tau}\|\,d\tau+$$
$$+\frac{1}{\mu_0}\|\Gamma_{C_n}\|\max_{0\leqslant t\leqslant l}\|x^{(n)}(t)\|_{E_n}\int_0^t\|e^{-D_n\tau}\|\,d\tau+$$
$$+\frac{1}{\mu_0}\|\Gamma_{B_n}\|\max_{0\leqslant t\leqslant l}\|x_n(t)\|_{E_n}. \qquad (18.26)$$

Очевидно,

$$\int_0^t\|e^{-D_n\tau}\|\,d\tau=\int_0^t e^{-\mu_1^{(n)}\tau}\,d\tau\leqslant\frac{1}{\mu_0}.$$

Что касается интеграла

$$\int_0^t\|D_ne^{-D_n\tau}\|\,d\tau,$$

то ему легко дать следующую оценку:

$$\int_0^t\|D_ne^{-D_n\tau}\|\,d\tau\leqslant\|D_n\|\int_0^t\|e^{-D_n\tau}\|\,d\tau=\frac{\|D_n\|}{\mu_0}=\frac{\nu_n^{(n)}}{\mu_0}. \qquad (18.27)$$

Этот же интеграл можно оценить и иначе.

Собственные числа матрицы $D_ne^{-D_n\tau}$ суть $\nu_k^{(n)}e^{-\nu_k^{(n)}\tau}$ ($k=1, 2, \ldots, n$) и норма упомянутой матрицы равна наибольшему из этих чисел.

Обозначим через M_k $(k=1, 2, \ldots, n)$ множество тех значений τ, для которых

$$\left\| D_n e^{-D_n \tau} \right\| = \nu_k^{(n)} e^{-\nu_k^{(n)} \tau}.$$

Тогда

$$\int_0^t \left\| D_n e^{-D_n \tau} \right\| d\tau \leqslant \int_0^\infty \left\| D_n e^{-D_n \tau} \right\| d\tau =$$

$$= \sum_{k=1}^n \int_{M_k} \nu_k^{(n)} e^{-\nu_k^{(n)} \tau} d\tau < \sum_{k=1}^n \int_0^\infty \nu_k^{(n)} e^{-\nu_k^{(n)} \tau} d\tau = n. \tag{18.28}$$

Обозначим

$$\sigma(n) = \min\left(\frac{\nu_n^{(n)}}{\mu_0}, n \right). \tag{18.29}$$

Тогда из (18.27) и (18.28) следует:

$$\int_0^t \left\| D_n e^{-D_n \tau} \right\| d\tau \leqslant \sigma(n), \tag{18.30}$$

и из неравенства (18.26) теперь можно легко получить оценку

$$\left\| A_n^{-1} \Gamma_n \right\|_{X_n \to Y_n} \leqslant [p' + p'' \sigma(n)] \left\| \Gamma_{B_n} \right\|_{E_n} + p''' \left\| \Gamma_{C_n} \right\|_{E_n}, \tag{18.31}$$

где p', p'', p''' — постоянные. Из неравенства (17.15′), характеризующего устойчивость процесса (18.4) — (18.5), вытекает теперь очевидным образом следующее неравенство:

$$\left\| \eta^{(n)} \right\|_{X_n} \leqslant [p_1 + p_2 \sigma(n)] \left\| \Gamma_{B_n} \right\|_{E_n} + p_3 \left\| \Gamma_{C_n} \right\|_{E_n} +$$
$$+ p_4 \max_{0 \leqslant t \leqslant l} \left\| \delta^{(n)}(t) \right\|_{E_n} + p_5 \left\| \delta_0^{(n)} \right\|_{E_n}, \tag{18.32}$$

где p_i $(1 \leqslant i \leqslant 5)$ — постоянные.

Аналогично из формулы (18.25) вытекает неравенство для погрешности приближенного решения

$$\max_{0 \leqslant t \leqslant l} \left\| v_n(t) - u_n(t) \right\|_H \leqslant$$
$$\leqslant \sqrt{M_0} \Big\{ [p_1 + p_2 \sigma(n)] \left\| \Gamma_{B_n} \right\|_{E_n} + p_3 \left\| \Gamma_{C_n} \right\|_{E_n} +$$
$$+ p_4 \max_{0 \leqslant t \leqslant l} \left\| \delta^{(n)}(t) \right\|_{E_n} + p_5 \left\| \delta_0^{(n)} \right\|_{E_n} \Big\}. \tag{18.33}$$

Оценки (18.32) и (18.33) верны, если

$$[p' + p'' \sigma(n)] \left\| \Gamma_{B_n} \right\|_{E_n} + p''' \left\| \Gamma_{C_n} \right\|_{E_n} \leqslant r < 1. \tag{18.34}$$

Замечание. Рассмотрим случай, когда координатная система ортонормирована в H. Тогда матрица $\rho_n = I_n$ вычисляется точно, и $\Gamma_{B_n} = 0$. В этом случае численные коэффициенты в формулах (18.32) и (18.33) не зависят от n. Именно этот случай и рассмотрен в работах М. А. Велиева [1], [2]. Аналогичное замечание относится и к оценкам § 21.

§ 19. Более общее уравнение первого порядка

Рассмотрим теперь задачу об интегрировании уравнения

$$\frac{d}{dt} Bu + Cu = f(t) \tag{19.1}$$

при начальном условии

$$u\big|_{t=0} = \varphi. \tag{19.2}$$

Будем считать, что операторы B и C — положительно определенные в данном гильбертовом пространстве H и что из двух энергетических пространств H_B и H_C одно вкладывается в другое.

Случай $H_C \subset H_B$ мало отличается от случая § 18, и здесь верно утверждение, аналогичное теореме 18.1: процесс Бубнова — Галёркина устойчив (в смысле определения § 17), если координатная система почти ортонормирована в H_B; устойчивость имеет место в промежутке $0 \leqslant t \leqslant l$ таком, что

$$\sup_{0 \leqslant t \leqslant l} \| f(t) \|_H < \infty. \tag{19.3}$$

Самый процесс Бубнова — Галёркина определяется уравнениями

$$\left.\begin{aligned} \rho_n \dot{a}^{(n)}(t) + R_n a^{(n)}(t) = f^{(n)}(t), \\ a^{(n)}(0) = \alpha^{(n)}, \end{aligned}\right\} \tag{19.4}$$

где на этот раз $\rho_n = R_{B_n}$, $R_n = R_{C_n}$, а $f^{(n)}(t)$ и $\alpha^{(n)}$ определены так же, как в § 18. Для погрешности $\eta^{(n)} = b^{(n)} - a^{(n)}$ вектора коэффициентов верны формулы (17.15′) и (18.32); соответствующие формулы для погрешности приближенного решения получаются из формул (18.25) и (18.33) заменой левой части на $\max\limits_{0 \leqslant t \leqslant l} | v_n(t) - u_n(t) |_B$. Мы не будем останавливаться на доказательстве нашего утверждения — для этого было бы достаточно повторить рассуждения § 18.

Отметим случай, когда не только $H_C \subset H_B$, то и $H_B \subset H_C$, так что энергетические пространства H_B и H_C состоят из одних и тех же элементов, и нормы в них эквивалентны.

В этом случае числа $\nu_k^{(n)}$ (формула (18.10)) ограничены сверху. Действительно, если $x_k^{(n)}$ — собственный вектор уравнения $R_n x - \nu_k^{(n)} \rho_n x = 0$, то

$$\nu_k^{(n)} = \frac{\left(R_n x_k^{(n)}, x_k^{(n)}\right)}{\left(\rho_n x_k^{(n)}, x_k^{(n)}\right)} = \frac{|u|_C^2}{|u|_B^2}, \qquad u = \sum_{j=1}^{n} x_{kj}^{(n)} \varphi_j, \tag{19.5}$$

где $x_{kj}^{(n)}$ — составляющие вектора $x_k^{(n)}$, а φ_j — координатные элементы. Нормы в H_B и H_C эквивалентны, поэтому существуют такие положительные c_1 и c_2, что $c_1 |u|_B \leqslant |u|_C \leqslant c_2 |u|_B$, и $\nu_k^{(n)} \leqslant c_2^2$.

При достаточно больших n формула (18.29) дает $\sigma(n) \leqslant \frac{c_2^2}{\mu_0} = \text{const}$, и вместо формул (18.32) и (18.33) мы получаем следующие формулы (обозначения постоянных изменены):

$$\begin{aligned} \| \eta^{(n)} \|_{X_n} \leqslant p_1 \| \Gamma_{B_n} \|_{E_n} + p_2 \| \Gamma_{C_n} \|_{E_n} + \\ + p_3 \max_{0 \leqslant t \leqslant l} \| \delta^{(n)}(t) \|_{E_n} + p_4 \| \delta_0^{(n)} \|_{E_n}, \end{aligned} \tag{19.6}$$

$$\begin{aligned} | v_n(t) - u_n(t) |_B \leqslant \sqrt{M_0} \Big\{ p_1 \| \Gamma_{B_n} \|_{E_n} + p_2 \| \Gamma_{C_n} \|_{E_n} + \\ + p_3 \max_{0 \leqslant t \leqslant l} \| \delta^{(n)}(t) \|_{E_n} + p_4 \| \delta_0^{(n)} \|_{E_n} \Big\}, \end{aligned} \tag{19.7}$$

в которых коэффициенты при нормах уже не зависят от n. Как это следует из формулы (18.34), оценки (19.6) и (19.7) верны, если

$$\left(p' + p'' \frac{c_2^2}{\mu_0}\right) \| \Gamma_{B_n} \|_{E_n} + p''' \| \Gamma_{C_n} \|_{E_n} \leqslant r < 1. \tag{19.8}$$

Докажем, что в рассматриваемом случае оценки вида (19.6) и (19.7), только с другими значениями постоянных p_i, справедливы и для погрешностей производных $\dot{\eta}^{(n)}(t)$ и $\dot{v}^{(n)}(t) - \dot{u}^{(n)}(t)$, если только нормы $\| \Gamma_{B_n} \|_{E_n}$ и $\| \Gamma_{C_n} \|_{E_n}$ достаточно малы. Координатная система, в силу теоремы 4.3, почти ортонормирована и в пространстве H_C; отсюда следует, что собственные числа $\lambda_k^{(n)}$ матриц R_n ограничены положительными числами снизу и сверху. Пусть $\lambda_0 \leqslant \lambda_k^{(n)} \leqslant \Lambda_0$. «Неточный» вектор коэффициентов Ритца $b^{(n)}(t)$ удовлетворяет уравнению

$$\left(\rho_n + \Gamma_{B_n}\right) \dot{b}^{(n)}(t) + \left(R_n + \Gamma_{C_n}\right) b^{(n)}(t) = f^{(n)}(t) + \delta^{(n)}(t).$$

Вычитая отсюда уравнение (19.4), получим:

$$\begin{aligned} \dot{\eta}^{(n)}(t) = \left(\rho_n + \Gamma_{B_n}\right)^{-1} \Big\{ \delta^{(n)}(t) - \left(R_n + \Gamma_{C_n}\right) \eta^{(n)}(t) - \\ - \Gamma_{B_n} \dot{a}^{(n)}(t) - \Gamma_{C_n} a^{(n)}(t) \Big\}. \end{aligned} \tag{19.9}$$

Величины $\|\Gamma_{B_n}\|_{E_n}$ и $\|\Gamma_{C_n}\|_{E_n}$ удовлетворяют неравенству (19.8). Дополнительно потребуем, чтобы $\|\Gamma_{B_n}\| \leqslant r\mu_0$, где μ_0 — положительная нижняя граница собственных чисел $\mu_k^{(n)}$ матрицы ρ_n. По формулам (18.21) и (18.24) величины $\|f^{(n)}(t)\|_{E_n}$ и $\|a^{(n)}(t)\|_{E_n}$ ограничены независимо от n и t. Далее, $\|\rho_n^{-1}\| \leqslant \frac{1}{\mu_0}$ и $\|R_n\|_{E_n} \leqslant \Lambda_0$, и из уравнения (19.4) следует, что величина $\|\dot{a}^{(n)}(t)\|_{E_n}$ также ограничена независимо от n и t. Наконец,

$$\left\|\left(\rho_n + \Gamma_{B_n}\right)^{-1}\right\| \leqslant \|\rho_n^{-1}\| \left\|\left(I_n + \rho_n^{-1}\Gamma_{B_n}\right)^{-1}\right\| \leqslant \frac{M_0}{1-r}.$$

Теперь из соотношений (19.7) и (19.9) вытекает неравенство вида

$$\|\dot{\eta}^{(n)}\|_{X_n} \leqslant q_1 \|\Gamma_{B_n}\|_{E_n} + q_2 \|\Gamma_{C_n}\|_{E_n} + \\ + q_3 \max_{0 \leqslant t \leqslant l} \|\delta^{(n)}(t)\|_{E_n} + q_4 \left\|\delta_0^{(n)}\right\|_{E_n}, \tag{19.10}$$

где q_i — постоянные; отсюда и из соотношения

$$|\dot{v}_n(t) - \dot{u}_n(t)|_B^2 = (\rho_n \dot{\eta}^{(n)}(t),\ \dot{\eta}^{(n)}(t))_{E_n}$$

следует неравенство:

$$|\dot{v}_n(t) - \dot{u}_n(t)|_B \leqslant \sqrt{M_0} \left\{ q_1 \|\Gamma_{B_n}\|_{E_n} + q_2 \|\Gamma_{C_n}\|_{E_n} + \right. \\ \left. + q_3 \max_{0 \leqslant t \leqslant l} \|\delta^{(n)}(t)\|_{E_n} + q_4 \left\|\delta_0^{(n)}\right\|_{E_n} \right\}. \tag{19.11}$$

Обратимся к случаю, когда $H_B \subset H_C$, но $H_C \not\subset H_B$. В этом случае мы дополнительно предположим, что $\varphi = 0$. Система (19.4) Бубнова — Галёркина принимает вид:

$$\rho_n \dot{a}^{(n)}(t) + R_n a^{(n)}(t) = f^{(n)}(t), \qquad a^{(n)}(0) = 0. \tag{19.12}$$

Примем на этот раз, что пространство Y_n совпадает с X_n. Уравнения (19.12) определяют оператор $A_n \in (X_n \to Y_n)$, действующий по формуле $\left(a^{(n)} = a^{(n)}(t)\right)$:

$$A_n a^{(n)} = \rho_n \dot{a}^{(n)}(t) + R_n a^{(n)}(t).$$

Этот оператор определен на векторах $a^{(n)} \in X_n$, непрерывно дифференцируемых по t и равных нулю при $t = 0$.

Теорема 19.1. *Пусть $H_B \subset H_C$. Процесс* (19.12) *устойчив в промежутке $0 \leqslant t \leqslant l$, $l < \infty$, если координатная система почти ортонормирована в H_C и если выполнено неравенство* (19.3).

Сохраняя обозначения § 18, мы приведем уравнения (19.12) к виду

$$\dot{c}_k^{(n)}(t)+\nu_k^{(n)}c_k^{(n)}(t)=F_k^{(n)}(t), \qquad c_k^{(n)}(0)=0.$$

Отсюда легко следует, что

$$a^{(n)}(t)=\int_0^t G_n^{-1}e^{-D_n(t-\tau)}G_n^{*^{-1}}f^{(n)}(\tau)\,d\tau. \tag{19.13}$$

В наших условиях нормы матриц G_n^{-1} $\left(\text{и } G_n^{*^{-1}}\right)$ ограничены независимо от n. Действительно, в данном случае координатная система сильно минимальна в H_B, поэтому $\mu_k^{(n)}\geqslant\mu_0=\text{const}>0$, где $\mu_k^{(n)}$ — собственные числа матрицы ρ_n; теперь, как и в § 18 (нормы взяты в метрике E_n),

$$\left\|G_n^{-1}\right\|=\left\|G_n^{*^{-1}}\right\|=\sqrt{\left\|\rho_n^{-1}\right\|}=\frac{1}{\sqrt{\mu_1^{(n)}}}\leqslant\frac{1}{\sqrt{\mu_0}}.$$

Числа $\nu_k^{(n)}$ положительны, поэтому $\left\|e^{-D_n(t-\tau)}\right\|\leqslant 1$ и

$$\left\|a^{(n)}(t)\right\|_{E_n}\leqslant\frac{1}{\mu_0}\int_0^t\left\|f^{(n)}(\tau)\right\|_{E_n}d\tau\leqslant$$

$$\leqslant\frac{l}{\mu_0}\max_{0\leqslant t\leqslant l}\left\|f^{(n)}(t)\right\|_{E_n}=\frac{l}{\mu_0}\left\|f^{(n)}\right\|_{X_n}.$$

Беря максимум левой части, получим:

$$\left\|a^{(n)}\right\|_{X_n}\leqslant\frac{l}{\mu_0}\left\|f^{(n)}\right\|_{X_n}. \tag{19.14}$$

Это неравенство означает, что нормы операторов A_n^{-1} ограничены в совокупности:

$$\left\|A_n^{-1}\right\|_{X_n}\leqslant\frac{1}{\mu_0}. \tag{19.15}$$

Докажем теперь, что и величины $\left\|f^{(n)}\right\|_{X_n}$ ограничены в совокупности. Имеем:

$$f^{(n)}(t)=((f(t),\ \varphi_1),\quad (f(t),\ \varphi_2),\ \ldots,\ (f(t),\ \varphi_n)).$$

Далее,

$$(f(t),\ \varphi_k)_H=\left(CC^{-1}f(t),\ \varphi_k\right)_H=\left[C^{-1}f(t),\ \varphi_k\right]_C.$$

Как и в § 18, поставим задачу о наилучшем приближении в метрике H_C элемента $h(t)=C^{-1}f(t)$ линейной комбинацией вида

$$\sum_{k=1}^{n}\alpha_k^{(n)}(t)\,\varphi_k.$$

Вектор $\alpha^{(n)}(t)=\left(\alpha_1^{(n)}(t),\ \alpha_2^{(n)}(t),\ \ldots,\ \alpha_n^{(n)}(t)\right)$ удовлетворяет уравнению $R_n\alpha^{(n)}(t)=h^{(n)}(t)$, где

$$h^{(n)}(t)=([h(t),\ \varphi_1]_C,\ [h(t),\ \varphi_2]_C,\ \ldots,\ [h(t),\ \varphi_n]_C).$$

Но

$$[h(t),\ \varphi_k]_C=(f(t),\ \varphi_k),$$

поэтому $h^{(n)}(t)=f^{(n)}(t)$, и $\alpha^{(n)}(t)$ удовлетворяет уравнению $R_n\alpha^{(n)}(t)=f^{(n)}(t)$. По формуле (5.6)

$$\|\alpha^{(n)}(t)\|_{E_n}\leqslant\frac{1}{\sqrt{\lambda_0}}|h(t)|_C=\frac{1}{\sqrt{\lambda_0}}\|f(t)\|_H\leqslant\frac{1}{\sqrt{\lambda_0}}\max_{0\leqslant t\leqslant l}\|f(t)\|_H$$

и, следовательно,

$$\|f^{(n)}(t)\|_{E_n}=\|R_n^{-1}\alpha^{(n)}(t)\|_{E_n}\leqslant\frac{\Lambda_0}{\sqrt{\lambda_0}}\max_{0\leqslant t\leqslant l}\|f(t)\|.$$

Взяв максимум левой части, найдем:

$$\|f^{(n)}\|_{X_n}\leqslant\frac{\Lambda_0}{\sqrt{\lambda_0}}\max_{0\leqslant t\leqslant l}\|f(t)\|=\text{const},$$

что и требовалось доказать. Из неравенства (19.14) следует теперь, что нормы $\|a^{(n)}\|_{X_n}$ ограничены в совокупности, и устойчивость процесса (19.12) вытекает из теоремы 17.1.

Докажем, что для погрешности $\eta^{(n)}(t)$ вектора коэффициентов Ритца верна оценка вида (19.6), в которой, разумеется, отсутствует слагаемое $p_4\|\delta_0^{(n)}\|_{E_n}$. Для этого оценим норму $\|A_n^{-1}\Gamma_n\|$. В данном случае оператор Γ_n действует из X_n в X_n; если положить $\Gamma_n x^{(n)}=\zeta^{(n)}$, то

$$\zeta^{(n)}(t)=\Gamma_{B_n}\dot{x}^{(n)}(t)+\Gamma_{C_n}x^{(n)}(t).$$

Обозначим $A_n^{-1}\Gamma_n x^{(n)}=\xi^{(n)}$. Повторяя рассуждения § 18, мы опять придем к оценкам (18.26) и (18 27). Но в данном случае числа $\nu_k^{(n)}$ ограничены, как это видно из формулы (19.5), и из только что упомянутых оценок вытекает, что

$$\|A_n^{-1}\Gamma_n\|\leqslant p'\|\Gamma_{B_n}\|_{E_n}+p''\|\Gamma_{C_n}\|_{E_n},\qquad p',\ p''=\text{const}.$$

На основании последнего неравенства легко получается оценка вида

$$\|\eta^{(n)}\|_{X_n}\leqslant p_1\|\Gamma_{B_n}\|_{E_n}+p_2\|\Gamma_{C_n}\|_{E_n}+p_3\max_{0\leqslant t\leqslant l}\|\delta^{(n)}(t)\|_{E_n}. \qquad (19.16)$$

Теперь из соотношения

$$|v_n(t)-u_n(t)|_C^2=(R_n\eta^{(n)}(t),\ \eta^{(n)}(t))_{E_n}$$

вытекает также неравенство

$$|v_n(t)-u_n(t)|_C \leqslant \sqrt{\Lambda_0}\,\{p_1\|\Gamma_{B_n}\|_{E_n}+p_2\|\Gamma_{C_n}\|_{E_n}+ \\ +p_3\max_{0\leqslant t\leqslant l}\|\delta^{(n)}(t)\|_{E_n}\}, \qquad (19.17)$$

где Λ_0 — верхняя граница собственных чисел матрицы R_n.

§ 20. Уравнения С. Л. Соболева

Обратимся к задаче § 17 и выясним достаточные условия устойчивости процесса (17.6′) — (17.7′). Ограничимся простейшим случаем, когда метрики пространств H_A, H_B, H_C эквивалентны и, следовательно, эти пространства состоят из одних и тех же элементов. Для определенности примем, что векторы $\alpha^{(n)}$ и $\beta^{(n)}$, входящие в начальные условия (17.7′), определены из уравнений (17.10′) и (17.10″).

Теорема 20.1. *Если метрики H_A, H_B, H_C эквивалентны и координатная система почти ортонормирована в одной из этих метрик, то процесс* (17.6′) — (17.7′) *устойчив в любом конечном промежутке изменения t.*

Обозначим $\varkappa_k^{(n)}$, $\mu_k^{(n)}$, $\lambda_k^{(n)}$ собственные числа матриц R_{A_n}, R_{B_n}, R_{C_n} соответственно. Координатная система почти ортонормирована в любой из метрик H_A, H_B, H_C, поэтому упомянутые собственные числа положительно ограничены сверху и снизу:

$$\varkappa_0\leqslant\varkappa_k^{(n)}\leqslant K_0,\quad \mu_0\leqslant\mu_k^{(n)}\leqslant M_0,\quad \lambda_0\leqslant\lambda_k^{(n)}\leqslant\Lambda_0, \qquad (20.1)$$

где $\varkappa_0$, K_0, μ_0, M_0, λ_0, Λ_0 — положительные постоянные.

Займемся оценкой нормы $\|A_n^{-1}\|$, где A_n — оператор, введенный в § 17. Аналогично тому, как это было сделано в § 18, положим:

$$R_{A_n}=G_n^*G_n,\qquad R_{B_n}=G_n^*D_nG_n,\qquad D_n=[\nu_1^{(n)},\ \nu_2^{(n)},\ \ldots,\ \nu_n^{(n)}];$$

по-прежнему доказывается, что нормы матриц G_n и G_n^{-1} ограничены, именно,

$$\|G_n\|=\|G_n^*\|\leqslant\sqrt{M_0},\qquad \|G_n^{-1}\|=\|G_n^{*-1}\|\leqslant\frac{1}{\sqrt{\mu_0}}.$$

Легко также видеть (ср. § 19), что числа $\nu_k^{(n)}$ положительно ограничены сверху и снизу. Пусть

$$0<\nu_0\leqslant\nu_k^{(n)}\leqslant N_0,\qquad \nu_0,\ N_0=\text{const}. \qquad (20.2)$$

Положим еще

$$G_na^{(n)}(t)=c^{(n)}(t),\quad G_n\alpha^{(n)}=\gamma^{(n)},\quad G_n\beta^{(n)}=\tilde\gamma^{(n)},\quad G_n^{*-1}f^{(n)}(t)=F^{(n)}(t). \qquad (20.3)$$

Уравнения (17.6′) и (17.7′) приводятся к виду

$$\left.\begin{aligned} &\ddot{c}^{(n)}(t)+D_n\dot{c}^{(n)}(t)+P_nc^{(n)}(t)=F^{(n)}(t),\\ &c^{(n)}(0)=\gamma^{(n)},\qquad \dot{c}^{(n)}(0)=\tilde{\gamma}^{(n)},\end{aligned}\right\} \tag{20.4}$$

где

$$P_n=G_n^{*^{-1}}R_{C_n}G_n^{-1}. \tag{20.5}$$

Нетрудно убедиться, что матрица P_n — положительно определенная и что ее собственные числа, которые мы обозначим через $\vartheta_k^{(n)}$, положительно ограничены сверху и снизу; пусть

$$0<\vartheta_0\leqslant\vartheta_k^{(n)}\leqslant\theta_0,\qquad \vartheta_0=\mathrm{const},\qquad \theta_0=\mathrm{const}. \tag{20.6}$$

Обе части первого уравнения (20.4) проинтегрируем по t в пределах $(0,\ l)$; приняв во внимание остальные уравнения (20.4), получим:

$$\dot{c}^{(n)}(t)+D_nc^{(n)}(t)=\int_0^t F^{(n)}(\tau)\,d\tau+\tilde{\gamma}^{(n)}+D_n\gamma^{(n)}-\int_0^t P_nc^{(n)}(\tau)\,d\tau.$$

Отсюда следует соотношение, аналогичное соотношению (18.17′):

$$c^{(n)}(t)=e^{-D_nt}\gamma^{(n)}+\int_0^t e^{-D_n(t-\tau)}\left[\tilde{\gamma}^{(n)}+D_n\gamma^{(n)}+\int_0^\tau F^{(n)}(\tau_1)\,d\tau_1-\right.$$
$$\left.-\int_0^\tau P_nc^{(n)}(\tau_1)\,d\tau_1\right]d\tau.$$

Это равенство легко приводится к (векторному) интегральному уравнению типа Вольтерра

$$c^{(n)}(t)+\int_0^t K_n(t-\tau)\,c^{(n)}(\tau)\,d\tau=\Phi^{(n)}(t), \tag{20.7}$$

где

$$K_n(t)=D_n^{-1}\left(I_n-e^{-D_nt}\right)P_n, \tag{20.8}$$

$$\Phi^{(n)}(t)=\gamma^{(n)}+D_n^{-1}\left(I_n-e^{-D_nt}\right)\tilde{\gamma}^{(n)}+\int_0^t D_n^{-1}\left(I_n-e^{-D_n(t-\tau)}\right)F^{(n)}(\tau)\,d\tau. \tag{20.9}$$

Уравнение (20.7) решается, как обычно, последовательными приближениями, и нетрудно дать оценку этому решению. Прежде всего,

$$\|K_n(t)\|\leqslant\|D_n^{-1}\|\,\|I_n-e^{-D_nt}\|\,\|P_n\|=$$
$$=\frac{1}{\nu_1^{(n)}}\left(1-e^{-\nu_1^{(n)}t}\right)\vartheta_n^{(n)}\leqslant q,\qquad q=\frac{\theta_0}{\nu_0}.$$

Далее,

$$\|\Phi^{(n)}(t)\|_{E_n} \leqslant \|\gamma^{(n)}\| + \frac{1}{\nu_1^{(n)}}\left(1 - e^{-\nu_n^{(n)} t}\right)\|\tilde{\gamma}^{(n)}\| +$$
$$+ \frac{l}{\nu_1^{(n)}} \max_{0 \leqslant t \leqslant l} \|F^{(n)}(t)\| \leqslant \|\gamma^{(n)}\| + \frac{1}{\nu_0}\|\tilde{\gamma}^{(n)}\| + \frac{l}{\nu_0} \max_{0 \leqslant t \leqslant l} \|F^{(n)}(t)\|_{E_n}.$$

Обозначая величину справа через σ_n, имеем:

$$\|\Phi^{(n)}(t)\| \leqslant \sigma_n.$$

Решение уравнения (20.7) имеет вид:

$$c^{(n)}(t) = \sum_{m=0}^{\infty} (-1)^m c^{(n,\, m)}(t),$$

где

$$c^{(n,\, 0)}(t) = \Phi^{(n)}(t), \qquad c^{(n,\, m)}(t) = \int_0^t K_n(t - \tau)\, c^{(n,\, m-1)}(\tau)\, d\tau.$$

Отсюда легко следуют оценки:

$$\|c^{(n,\, m)}(t)\|_{E_n} \leqslant \frac{\sigma_n q^m t^m}{m!}$$

и

$$\|c^{(n)}(t)\|_{E_n} \leqslant \sigma_n e^{qt} \leqslant \sigma_n e^{ql}.$$

Теперь

$$\|a^{(n)}(t)\|_{E_n} = \|G_n^{-1} c^{(n)}(t)\|_{E_n} \leqslant \sigma_n \frac{e^{ql}}{\sqrt{\mu_0}}.$$

Беря максимум левой части, получим:

$$\|a^{(n)}\|_{X_n} = \max_{0 \leqslant t \leqslant l} \|a^{(n)}(t)\|_{E_n} \leqslant \sigma_n \frac{e^{ql}}{\sqrt{\mu_0}}. \tag{20.10}$$

Далее, по формулам (20.3)

$$\sigma_n = \|G_n \alpha^{(n)}\| + \frac{1}{\nu_0}\|G_n \beta^{(n)}\| + \frac{l}{\nu_0} \max_{0 \leqslant t \leqslant l} \left\|G_n^{*^{-1}} f^{(n)}(t)\right\|_{E_n} \leqslant$$
$$\leqslant \sqrt{M_0}\,\|\alpha^{(n)}\| + \frac{\sqrt{M_0}}{\nu_0}\|\beta^{(n)}\| + \frac{l}{\nu_0 \sqrt{\mu_0}} \max_{0 \leqslant t \leqslant l} \|f^{(n)}(t)\|_{E_n}.$$

Пусть

$$q_1 = \max\left\{\sqrt{M_0},\ \frac{\sqrt{M_0}}{\nu_0},\ \frac{l}{\nu \sqrt{\mu_0}}\right\}.$$

Обозначим еще через $g^{(n)}$ тройку векторов $\left(f^{(n)}(t),\ \alpha^{(n)},\ \beta^{(n)}\right)$ Элемент $g^{(n)}$ принадлежит пространству Y_n и

$$\sigma_n \leqslant q_1 \|g^{(n)}\|_{Y_n}.$$

Теперь из оценки (20.10) следует, что

$$\| a^{(n)} \|_{X_n} \leqslant q_2 \| g^{(n)} \|_{Y_n}, \qquad q_2 = \frac{q_1 e^{ql}}{\sqrt{\mu_0}} = \mathrm{const}.$$

Последнее неравенство означает, что

$$\| A_n^{-1} \| \leqslant q_2,$$

где A_n — оператор, введенный в § 17 и переводящий вектор $a^{(n)} \in X_n$ в тройку векторов $g^{(n)} = \left(f^{(n)}(t),\ \alpha^{(n)},\ \beta^{(n)}\right) \in Y_n$.

Повторив с необходимыми изменениями соответствующие рассуждения § 18, найдем, что нормы $\| a^{(n)} \|_{X_n}$ ограничены в совокупности. Теперь теорема 20.1 вытекает из теоремы 17.1.

Для вектора погрешности $\eta^{(n)}(t) = b^{(n)}(t) - a^{(n)}(t)$ нетрудно теперь получить оценку вида

$$\| \eta^{(n)} \|_{X_n} \leqslant p_1 \| \Gamma_{A_n} \| + p_2 \| \Gamma_{B_n} \| + p_3 \| \Gamma_{C_n} \| + p_4 \max_{0 \leqslant t \leqslant l} \| \delta^{(n)}(t) \| + \\ + p_5 \| \delta_0^{(n)} \| + p_6 \| \delta_1^{(n)} \|, \qquad (20.11)$$

где p_i — постоянные, и нормы справа взяты в метрике E_n. Аналогичные неравенства можно установить для величин $\| \dot{\eta}^{(n)} \|_{X_n}$, $| v_n(t) - u_n(t) |$ и $| \dot{v}_n(t) - \dot{u}_n(t) |$, где символ $|\ \ |$ означает норму в любом из пространств H_A, H_B, H_C.

§ 21. Уравнения гиперболического типа

Условие эквивалентности метрик H_A, H_B, H_C, принятое в § 20, не необходимо для того, чтобы процесс Бубнова — Галёркина для уравнения типа С. Л. Соболева оказался устойчивым. Поясним это на примере гиперболического уравнения

$$\frac{d^2 u}{dt^2} + B \frac{du}{dt} + Cu = f(t), \qquad (21.1)$$

которое мы будем интегрировать, как всегда, при начальных условиях

$$u|_{t=0} = \varphi, \qquad u_t|_{t=0} = \psi. \qquad (21.2)$$

Примем, что B — неотрицательная постоянная, а C — положительно определенный оператор.

Система Бубнова — Галёркина имеет вид (мы записываем ее в векторной форме):

$$\rho_n \ddot{a}^{(n)}(t) + B \rho_n \dot{a}^{(n)}(t) + R_n a^{(n)}(t) = f^{(n)}(t), \qquad (21.3)$$

$$a^{(n)}(0) = \alpha^{(n)}, \qquad \dot{a}^{(n)}(0) = \beta^{(n)}. \qquad (21.4)$$

Здесь $\rho_n = \|(\varphi_k, \varphi_j)\|_{j,\,k=1}^{j,\,k=n}$, $R_n = \|[\varphi_k, \varphi_j]_C\|_{j,\,k=1}^{j,\,k=n}$, а $f^{(n)}(t)$ имеет то же значение, что и в предшествующих параграфах; векторы $\alpha^{(n)}$ и $\beta^{(n)}$ будем определять из уравнений

$$R_n\alpha^{(n)} = \varphi^{(n)}, \qquad R_n\beta^{(n)} = \psi^{(n)}, \tag{21.5}$$

где векторы $\varphi^{(n)}$ и $\psi^{(n)}$ определены формулами

$$\left.\begin{aligned} \varphi^{(n)} &= ([\varphi, \varphi_1]_C, [\varphi, \varphi_2]_C, \ldots, [\varphi, \varphi_n]_C), \\ \psi^{(n)} &= ([\psi, \varphi_1]_C, [\psi, \varphi_2]_C, \ldots, [\psi, \varphi_n]_C). \end{aligned}\right\} \tag{21.6}$$

Справедлива следующая теорема.

Теорема 21.1. *Пусть координатная система почти ортонормирована в данном гильбертовом пространстве H и пусть в промежутке $0 \leqslant t \leqslant l$, $l < \infty$,*

$$\max_{0 \leqslant t \leqslant l} \|f(t)\| < \infty. \tag{21.7}$$

Тогда процесс (21.3) — (21.4) *устойчив* (*в смысле определения* § 17) *в промежутке $0 \leqslant t \leqslant l$. Если $B > 0$, то процесс* (21.3) — (21.4) *устойчив в промежутке $0 \leqslant t < \infty$.*

Для упрощения выкладок проведем доказательство при одном дополнительном предположении, о котором скажем несколько ниже.

Подстановкой

$$a^{(n)}(t) = e^{-\frac{B}{2}t}\tilde{a}^{(n)}(t) \tag{21.8}$$

система (21.3) — (21.4) сводится к следующей:

$$\left.\begin{aligned} &\rho_n\ddot{\tilde{a}}^{(n)}(t) + \left(R_n - \frac{B^2}{4}\rho_n\right)\tilde{a}^{(n)}(t) = e^{\frac{B}{2}t}f^{(n)}(t), \\ &\tilde{a}^{(n)}(0) = \alpha^{(n)}, \qquad \dot{\tilde{a}}^{(n)}(0) = \frac{B}{2}\alpha^{(n)} + \beta^{(n)}. \end{aligned}\right\} \tag{21.9}$$

Как и в предшествующих параграфах, положим:

$$\rho_n = G_n^*G_n, \qquad R_n = G_n^*D_nG_n, \qquad D_n = \left[\nu_1^{(n)}, \nu_2^{(n)}, \ldots, \nu_n^{(n)}\right],$$

при этом $\nu_k^{(n)} \geqslant \nu_0 = \mathrm{const} > 0$. *Мы примем в последующем, что* $\nu_0 > \frac{B^2}{4}$.

Полагая, далее,

$$G_n\tilde{a}^{(n)}(t) = c^{(n)}(t), \quad G_n\alpha^{(n)} = \gamma^{(n)}, \quad G_n\left(\frac{B}{2}\alpha^{(n)} + \beta^{(n)}\right) = \tilde{\gamma}^{(n)},$$

$$e^{\frac{B}{2}t}G_n^{*-1}f^{(n)}(t) = F^{(n)}(t),$$

мы приведем задачу (21.9) к виду:

$$\ddot{c}^{(n)}(t)+\left(D_n-\frac{B^2}{4}I_n\right)c^{(n)}(t)=F^{(n)}(t),$$

$$c^{(n)}(0)=\gamma^{(n)}, \qquad \dot{c}^{(n)}(0)=\tilde{\gamma}^{(n)}.$$

Решение этой последней задачи имеет вид:

$$c^{(n)}(t)=\cos\left(\left(D_n-\frac{B^2}{4}I_n\right)^{\frac{1}{2}}t\right)\gamma^{(n)}+$$
$$+\left(D_n-\frac{B^2}{4}I_n\right)^{-\frac{1}{2}}\sin\left(\left(D_n-\frac{B^2}{4}I_n\right)^{\frac{1}{2}}t\right)\tilde{\gamma}^{(n)}+$$
$$+\left(D_n-\frac{B^2}{4}I_n\right)^{-\frac{1}{2}}\int_0^t\sin\left(D_n-\frac{B^2}{4}I_n\right)^{\frac{1}{2}}(t-\tau)F^{(n)}(\tau)\,d\tau. \qquad (21.10)$$

Перейдем к оценкам.

В силу условия $\frac{B^2}{4}<\nu_0$ матрица $D_n-\frac{B^2}{4}I_n$ — положительно определенная, поэтому

$$\left\|\cos\left(D_n-\frac{B^2}{4}I_n\right)^{\frac{1}{2}}t\right\|\leqslant 1,$$

$$\left\|\left(D_n-\frac{B^2}{4}I_n\right)^{-\frac{1}{2}}\sin\left(D_n-\frac{B^2}{4}I_n\right)^{\frac{1}{2}}t\right\|\leqslant t,$$

теперь из формулы (21.10) следует:

$$\|c^{(n)}(t)\|_{E_n}\leqslant\|\gamma^{(n)}\|+t\|\tilde{\gamma}^{(n)}\|+\int_0^t(t-\tau)\|F^{(n)}(\tau)\|\,d\tau\leqslant$$
$$\leqslant\|\gamma^{(n)}\|+t\|\tilde{\gamma}^{(n)}\|+\frac{t^2}{2}\max_{0\leqslant t\leqslant l}\|F^{(n)}(t)\|_{E_n}. \qquad (21.11)$$

Пусть $l<\infty$. Тогда

$$\|c^{(n)}(t)\|_{E_n}\leqslant\|\gamma^{(n)}\|+l\|\tilde{\gamma}^{(n)}\|+\frac{l^2}{2}\max_{0\leqslant t\leqslant l}\|F^{(n)}(t)\|_{E_n},$$

и так как $\|G_n\|$ и $\|G_n^{-1}\|$ ограничены, то

$$\|c^{(n)}(t)\|_{E_n}\leqslant q\|g^{(n)}\|_{Y_n}, \qquad (21.12)$$

где q — некоторая постоянная, $g^{(n)}$ — тройка векторов $\left(f^{(n)}(t),\alpha^{(n)},\beta^{(n)}\right)$, а Y_n — пространство, введенное в § 17. Далее,

$$\|a^{(n)}(t)\|_{E_n}=\left\|e^{-\frac{B}{2}t}G_nc^{(n)}(t)\right\|\leqslant\|G_n\|\,\|c^{(n)}(t)\|_{E_n}\leqslant q_1\|g^{(n)}\|_{Y_n},$$
$$q_1=\mathrm{const}$$

и, следовательно,

$$\| a^{(n)} \|_{X_n} = \max_{0 \leqslant t \leqslant l} \| a^{(n)}(t) \|_{E_n} \leqslant q_1 \| g^{(n)} \|_{Y_n}, \tag{21.13}$$

Таким образом, первое из условий теоремы 17.1 выполнено. Те же рассуждения, что и в предшествующих параграфах настоящей главы, показывают, что величина $\| g^{(n)} \|_{Y_n}$ ограничена, а тогда, в силу неравенства (21.13), ограничена и величина $\| a^{(n)} \|_{X_n}$, второе условие теоремы 17.1 также выполнено, и наша теорема доказана для случая $l < \infty$.

Пусть теперь $l = \infty$. В этом случае полагаем $B > 0$. В силу формулы (21.11)

$$\| a^{(n)}(t) \|_{E_n} \leqslant \| G_n \| e^{-\frac{1}{2}Bt} \| c^{(n)}(t) \| \leqslant$$
$$\leqslant \| G_n \| \left\{ \| \gamma^{(n)} \| + \frac{2}{Be} \| \tilde{\gamma}^{(n)} \| + \frac{8}{B^2 e^2} \max_{0 \leqslant t \leqslant l} \| F^{(n)}(t) \|_{E_n} \right\},$$

отсюда легко следует, что $a^{(n)}$ удовлетворяет неравенству вида (21.13); дальнейшие рассуждения протекают, как и выше.

ГЛАВА IV

О НЕВЯЗКЕ ПРИБЛИЖЕННОГО РЕШЕНИЯ

Вопрос, трактуемый в настоящей главе, связан с применением процесса Ритца к функционалу энергетического метода. Пусть нам дано уравнение

$$Au = f, \tag{IV. 1}$$

где оператор A — положительно определенный в некотором гильбертовом пространстве H, и пусть процессом Ритца построено приближение

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k \tag{IV. 2}$$

к точному решению u_0 уравнения (IV. 1). Вопрос о том, в какой мере u_n удовлетворяет уравнению (IV. 1), в общем случае лишен смысла: координатные элементы, а с ними и элемент u_n, могут не входить в $D(A)$, а тогда невозможно подставить u_n в уравнение (IV. 1). Но если даже $\varphi_k \in D(A)$, то, вообще говоря, $Au_n \not\to f$; более того, если при любом выборе координатной системы из $D(A)$ будет $Au_n \underset{n\to\infty}{\longrightarrow} f$, то оператор A — ограниченный [1]).

Однако при специальном выборе координатной системы может оказаться, что невязка приближенного решения, $Au_n - f$, стремится к нулю при $n \to \infty$ и тогда, когда оператор A неограничен. Для того случая, когда оператор A имеет точечный спектр, такой выбор указан в статье автора [11] [2]). Теорема, данная автором в статье [11], нуждается в уточнении, которое по существу содержится в статье П. Е. Соболевского [1].

В настоящей главе дано подробное изложение результатов указанных статей; глава начинается с доказательства одной теоремы Н. И. Польского [1]; точнее говоря, предложение, доказываемое в ближайшем параграфе, является нужным для дальнейшего частным случаем только что упомянутой теоремы.

---

[1]) См. ПМ, стр. 41.

[2]) Некоторые обобщения даны в статье О. К. Богаряна [1].

§ 22. Теорема Н. И. Польского

Пусть u — произвольный элемент некоторого гильбертова пространства $\mathfrak{H}$; скалярное произведение и норму в $\mathfrak{H}$ будем обозначать обычными символами $(\quad,\quad)$ и $\|\quad\|$. Пусть в этом пространстве даны две полные последовательности элементов $\{\varphi_k\}$ и $\{\psi_k\}$; мы примем еще, что при любом n как элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$, так и элементы $\psi_1, \psi_2, \ldots, \psi_n$ линейно независимы. Положим

$$u_n = \sum_{k=1}^{n} a_k \varphi_k, \tag{22.1}$$

и пусть коэффициенты[1]) a_k определяются из условия, что $u_n - u$ ортогонально к $\psi_1, \psi_2, \ldots, \psi_n$:

$$(u_n, \psi_j) = \sum_{k=1}^{n} a_k (\varphi_k, \psi_j) = (u, \psi_j) \qquad (j = 1, 2, \ldots, n). \tag{22.2}$$

Поставим вопрос: какие условия достаточно наложить на последовательности $\{\varphi_k\}$ и $\{\psi_k\}$, чтобы система (22.2) была разрешима при любом n и чтобы $u_n \to u$?

Обозначим через L_n и M_n подпространства, натянутые на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ и $\psi_1, \psi_2, \ldots, \psi_n$ соответственно; через P_n обозначим оператор проектирования на подпространство M_n. Подчиним наши последовательности следующему условию: существует такая постоянная C, что для любого n и для любого элемента $v \in L_n$ справедливо неравенство

$$\|v\| \leqslant C \|P_n v\|. \tag{22.3}$$

Теорема 22.1. *Если выполнено условие* (22.3), *то система* (22.2) *единственным образом разрешима при любом* n, *и* $\|u - u_n\| \underset{n\to\infty}{\longrightarrow} 0$.

Система (22.2) равносильна уравнению

$$P_n u_n = P_n u, \qquad u_n \in L_n. \tag{22.4}$$

Соответствующая однородная система имеет вид:

$$P_n u_n = 0, \qquad u_n \in L_n.$$

Пусть u_n — какое-либо ее решение. По неравенству (22.3)

$$\|u_n\| \leqslant C \|P_n u_n\| = 0.$$

Отсюда $u_n = 0$, и так как элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы, то соответствующие коэффициенты $a_k = 0$ $(k = 1, 2, \ldots, n)$;

---

[1]) Здесь нам нет необходимости подчеркивать зависимость коэффициентов от n, и потому мы пишем a_k вместо более точного символа $a_k^{(n)}$.

отсюда в свою очередь следует, что система (22.2) имеет единственное решение при любом n.

Переходим к доказательству второй части теоремы. Система $\{\psi_n\}$ полна, поэтому $P_n u \to u$; из равенства (22.4) следует теперь, что $P_n u_n \to u$. Пусть Π_n — оператор проектирования на подпространство L_n. Очевидно, $u_n - \Pi_n u \in L_n$; по соотношениям (22.3) и (22.4)

$$\| u_n - \Pi_n u \| \leqslant C \| P_n u_n - P_n \Pi_n u \| = C \| P_n (u - \Pi_n u) \|,$$

и так как $\|P_n\| = 1$, то

$$\| u_n - \Pi_n u \| \leqslant C \| u - \Pi_n u \|.$$

Правая часть этого неравенства стремится к нулю при $n \to \infty$ в силу полноты системы $\{\varphi_n\}$, поэтому $u_n - \Pi_n u \to 0$. Но тогда

$$u - u_n = [(u - \Pi_n u) - (u_n - \Pi_n u)] \underset{n \to \infty}{\to} 0,$$

и теорема доказана.

§ 23. Теорема о невязке

Как известно, оператор, самосопряженный в некотором гильбертовом пространстве H, называется оператором с точечным спектром [1]), если система его собственных элементов полна в H. Если этот оператор еще и положительно определенный, то система его собственных элементов полна и в соответствующем энергетическом пространстве. Действительно, пусть A — положительно определенный оператор с точечным спектром; пусть λ_k и ω_k — его собственные числа и нормированные собственные элементы, так что $A\omega_k = \lambda\omega_k$ и $(Au, u) \geqslant \gamma^2 \| u \|^2$. Тогда $\lambda_k \geqslant \gamma^2 > 0$ и

$$A^{\frac{1}{2}} u = \sum_{k=1}^{\infty} \sqrt{\lambda_k} (u, \omega_k) \omega_k, \qquad u \in D\left(A^{\frac{1}{2}}\right),$$

причем сходится ряд

$$\sum_{k=1}^{\infty} \lambda_k | (u, \varphi_k) |^2 = \left\| A^{\frac{1}{2}} u \right\|^2.$$

Одновременно, так как система $\{\omega_k\}$ ортонормирована и полна в H, то

$$u = \sum_{k=1}^{\infty} (u, \omega_k) \omega_k, \qquad \sum_{k=1}^{\infty} | (u, \omega_k) |^2 = \| u \|^2.$$

---

[1]) См., например, В. И. Смирнов [3], стр. 463 и 578.

Но в случае положительно определенного оператора множества $D\left(A^{\frac{1}{2}}\right)$ и H_A состоят из одних и тех же элементов [1]), причем, если $u \in D\left(A^{\frac{1}{2}}\right)$, то $|u|_A = \left\| A^{\frac{1}{2}} u \right\|$. Положим

$$\zeta_n = \sum_{k=1}^{n} (u, \omega_k)\, \omega_k.$$

Очевидно, $\zeta_n \in D\left(A^{\frac{1}{2}}\right)$ и

$$A^{\frac{1}{2}} \zeta_n = \sum_{k=1}^{n} \sqrt{\lambda_k}\, (u, \omega_k)\, \omega_k.$$

Теперь

$$|u - \zeta_n|_A^2 = \left\| A^{\frac{1}{2}} u - A^{\frac{1}{2}} \zeta_n \right\|^2 = \sum_{k=n+1}^{\infty} \lambda_k \, |(u, \omega_k)|^2 \underset{n \to \infty}{\longrightarrow} 0.$$

Таким образом, любой элемент $u \in H_A$ аппроксимируется в H_A линейными комбинациями элементов ω_k; это и означает, что система $\{\omega_k\}$ полна в H_A.

Теорема 23.1. *Пусть A и B — сходные положительно определенные операторы и пусть B — оператор с точечным спектром. Допустим еще, что операторы A и B удовлетворяют неравенству*

$$|(Au, Bu)| \geqslant a \, \|Au\|^2, \tag{23.1}$$

где a — положительная постоянная. Если систему $\{\varphi_k\}$ собственных элементов оператора B принять за координатную для уравнения

$$Au = f \tag{23.2}$$

и если [2])

$$u_n = \sum_{k=1}^{n} a_k \varphi_k \tag{23.3}$$

есть n-е приближение по Ритцу к решению уравнения (23.2), *то невязка $Au_n - f$ стремится к нулю при $n \to \infty$.*

В силу неравенства (3.1) можно заменить соотношение (23.1) ему равносильным

$$|(Au, Bu)| \geqslant a' \, \|Bu\|^2, \qquad a' = \text{const}, \tag{23.1$'$}$$

которым мы обычно и будем пользоваться; систему $\{\varphi_n\}$ будем считать ортонормированной в H.

---

[1]) См. ПМ, стр. 24, замечание 3.

[2]) Как и выше, мы не подчеркиваем здесь зависимости коэффициентов Ритца от n, и обозначаем их через a_k вместо $a_k^{(n)}$.

Сходные операторы A и B являются также и полусходными (§ 3); пространства H_A и H_B состоят из одних и тех же элементов и метрики в этих пространствах эквивалентны. Поэтому система $\{\varphi_k\}$, полная в H_B, полна и в H_A и может быть использована как координатная для уравнения (23.2). Более того, очевидно, что $\varphi_k \in D(B)$, и так как A и B — сходные операторы, то $\varphi_k \in D(A)$ $(k=1, 2, \ldots)$. В таком случае можно считать, что коэффициенты a_k в выражении (23.3) удовлетворяют следующей системе уравнений:

$$\sum_{k=1}^{n} (A\varphi_k, \varphi_j)\, a_k = (f, \varphi_j) \qquad (j=1, 2, \ldots, n). \tag{23.4}$$

Пусть u — решение уравнения (23.1). Положим

$$Bu = w, \qquad Bu_n = w_n. \tag{23.5}$$

Уравнению (23.2) можно придать вид

$$w = BA^{-1}f; \tag{23.6}$$

при этом, в силу соотношения (23.3),

$$w_n = \sum_{k=1}^{n} c_k\varphi_k, \qquad c_k = \lambda_k a_k. \tag{23.7}$$

Преобразуем систему (23.4). Положительно определенные операторы A и B симметричны. Кроме того, в силу равенства (23.1), $\varphi_k = \lambda_k B^{-1}\varphi_k$. Отсюда

$$(A\varphi_k, \varphi_j) = (\varphi_k, A\varphi_j) = \lambda_k \left(B^{-1}\varphi_k, A\varphi_j\right) = \lambda_k \left(\varphi_k, B^{-1}A\varphi_j\right).$$

Выражение (f, φ_j) преобразуем так:

$$(f, \varphi_j) = \left(AB^{-1}BA^{-1}f, \varphi_j\right) = \left(BA^{-1}f, B^{-1}A\varphi_j\right) = \left(w, B^{-1}A\varphi_j\right).$$

Теперь система (23.4) принимает такую форму:

$$\sum_{k=1}^{n} c_k\left(\varphi_k, B^{-1}A\varphi_j\right) = \left(w, B^{-1}A\varphi_j\right) \qquad (j=1, 2, \ldots, n). \tag{23.8}$$

Система (23.8) с точностью до обозначений совпадает с системой (22.2); следует только положить $\psi_k = B^{-1}A\varphi_k$. Докажем, что в нашем случае имеет место неравенство типа (22.3).

В рассматриваемом случае L_n и M_n суть подпространства пространства H, базисы которых суть $\varphi_1, \varphi_2, \ldots, \varphi_n$ и $B^{-1}A\varphi_1, B^{-1}A\varphi_2, \ldots, B^{-1}A\varphi_n$ соответственно. Проектор P_n можно построить так: найдем такие постоянные μ_{jk}, чтобы

$$\left\| \varphi_j - \sum_{k=1}^{n} \mu_{jk} B^{-1}A\varphi_k \right\|^2 = \min \qquad (j=1, 2, \ldots, n), \tag{23.9}$$

тогда для любого элемента $v \in L_n$ имеем:

$$v = \sum_{k=1}^{n} \gamma_k \varphi_k, \qquad P_n v = \sum_{j,\,k=1}^{n} \gamma_j \mu_{jk} B^{-1} A \varphi_k; \tag{23.10}$$

неравенство (22.3) принимает вид:

$$\left\| \sum_{k=1}^{n} \gamma_k \varphi_k \right\|^2 \leqslant c^2 \left\| \sum_{j,\,k=1}^{n} \gamma_j \mu_{jk} B^{-1} A \varphi_k \right\|^2$$

или, что то же,

$$\sum_{k=1}^{n} |\gamma_k|^2 \leqslant c^2 \sum_{j,\,k=1}^{n} \gamma_j \bar{\gamma}_k \sum_{r,\,s=1}^{n} \mu_{jr} \bar{\mu}_{ks} (B^{-1} A \varphi_r,\ B^{-1} A \varphi_s), \tag{23.11}$$

и достаточно доказать, что наименьшее собственное число эрмитовой формы

$$\Gamma = \sum_{j,\,k=1}^{n} \gamma_j \bar{\gamma}_k \sum_{r,\,s=1}^{n} \mu_{jr} \bar{\mu}_{ks} (B^{-1} A \varphi_r,\ B^{-1} A \varphi_s) \tag{23.12}$$

ограничено снизу положительным числом, не зависящим от n.

Положим

$$\sum_{j=1}^{n} \gamma_j \mu_{jr} = \delta_r. \tag{23.13}$$

Тогда

$$\Gamma = \sum_{r,\,s=1}^{n} \delta_r \delta_s (B^{-1} A \varphi_r,\ B^{-1} A \varphi_s) = \left\| B^{-1} A \sum_{r=1}^{n} \delta_r \varphi_r \right\|^2.$$

Обозначим временно

$$\sum_{r=1}^{n} \delta_r \varphi_r = \zeta, \qquad B^{-1} A \zeta = \eta.$$

Тогда $\zeta = A^{-1} B \eta$ и $\|\zeta\| \leqslant \|A^{-1}B\| \, \|\eta\|$; что оператор $A^{-1}B$ ограничен, было доказано в § 3. Теперь $\|\eta\| \geqslant \|A^{-1}B\|^{-1} \|\zeta\|$ или

$$\Gamma \geqslant \|A^{-1}B\|^{-2} \left\| \sum_{r=1}^{n} \delta_r \varphi_r \right\|^2 = \|A^{-1}B\|^{-2} \sum_{r=1}^{n} |\delta_r|^2. \tag{23.14}$$

Обозначим через M_n матрицу преобразования (23.13):

$$M_n = \| \mu_{jr} \|_{j,\,r=1}^{j,\,r=n}.$$

Докажем, что обратная матрица M_n^{-1} существует и $\|M_n^{-1}\| \leqslant C_1$, где C_1 — постоянная, которая не зависит от n.

Задача о минимуме величины (23.9) приводит к следующей системе уравнений для неизвестных μ_{jk}:

$$\sum_{k=1}^{n} \left(B^{-1}A\varphi_k,\ B^{-1}A\varphi_m\right)\mu_{jk} = \left(\varphi_j,\ B^{-1}A\varphi_m\right)$$
$$(m,\ j = 1,\ 2,\ \ldots,\ n). \tag{23.15}$$

Обозначим через Φ_n и Ψ_n матрицы, элементы которых соответственно равны $\left(B^{-1}A\varphi_k,\ B^{-1}A\varphi_m\right)$ и $\left(\varphi_k,\ B^{-1}A\varphi_m\right)$. Тогда система (23.15) записывается совсем просто:

$$M_n\Phi_n = \Psi_n. \tag{23.16}$$

Докажем, что матрица Ψ_n^{-1} существует.

Имеем:

$$\left(\varphi_k,\ B^{-1}A\varphi_m\right) = \left(B^{-1}\varphi_k,\ A\varphi_m\right) = \frac{1}{\lambda_k}\left(\varphi_k,\ A\varphi_m\right) = \frac{1}{\lambda_k}\left(A\varphi_k,\ \varphi_m\right).$$

Обозначим через Λ_n диагональную матрицу

$$\Lambda_n = (\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n)$$

и через R_n матрицу Ритца для оператора A и координатной системы $\{\varphi_k\}$. Тогда, очевидно, $\Psi_n = \Lambda_n^{-1}R_n$ и, следовательно, существует матрица $\Psi_n^{-1} = R_n^{-1}\Lambda_n$. Теперь из (23.16) следует: $M_n^{-1} = \Phi_n\Psi_n^{-1}$ и $\left\| M_n^{-1} \right\| \leqslant \left\| \Phi_n \right\| \left\| \Psi_n^{-1} \right\|$. Легко видеть, что величина $\left\|\Phi_n\right\|$ ограничена. Действительно, если $t = (t_1,\ t_2,\ \ldots,\ t_n)$ — произвольный вектор, то

$$\begin{aligned}\|\Phi_n t\|^2 &= \sum_{k,\ m=1}^{n} \left(B^{-1}A\varphi_k,\ B^{-1}A\varphi_m\right) t_k\bar{t}_m = \\ &= \left\| B^{-1}A\left(\sum_{k=1}^{n} t_k\varphi_k\right)\right\|^2 \leqslant \left\|B^{-1}A\right\|^2 \left\|\sum_{k=1}^{n} t_k\varphi_k\right\|^2 = \\ &= \left\|B^{-1}A\right\|^2 \sum_{k=1}^{n} |t_k|^2 = \left\|B^{-1}A\right\|^2 \|t\|^2,\end{aligned} \tag{23.17}$$

отсюда $\left\|\Phi_n\right\| \leqslant \left\|B^{-1}A\right\|$. Остается показать, что величина $\left\|\Psi_n^{-1}\right\|$ ограничена. Пусть, по-прежнему, t — произвольный вектор. Нам будет достаточно доказать, что имеет место неравенство

$$\|\Psi_n t\| \geqslant c_0\|t\|, \tag{23.18}$$

где c_0 не зависит от n и t. Обозначая для краткости $B^{-1}A\varphi_m = \psi_m$, имеем:

$$\|\Psi_n t\|^2 = \sum_{k=1}^{n}\left|\sum_{m=1}^{n} (\varphi_k,\ \psi_m)\, t_m\right|^2,$$

что легко преобразуется к виду

$$\left.\begin{aligned} \|\Psi_n t\|^2 &= \sum_{k=1}^{n} |(\psi, \varphi_k)|^2, \\ \psi = \sum_{m=1}^{n} \bar{t}_m \psi_m &= B^{-1}A \sum_{m=1}^{n} \bar{t}_m \varphi_m. \end{aligned}\right\} \tag{23.19}$$

Обратимся к неравенству (23.1′). Полагая в нем $Bu = v$, мы приведем его к виду

$$|(AB^{-1}v, v)| \geqslant a' \|v\|^2;$$

здесь v — любой элемент пространства H. Но

$$|(AB^{-1}v, v)| = |(v, B^{-1}Av)| = |(B^{-1}Av, v)|,$$

и мы приходим к неравенству

$$|(B^{-1}Av, v)| \geqslant a' \|v\|^2.$$

Полагая в нем

$$v = \sum_{m=1}^{n} \bar{t}_m \varphi_m,$$

находим

$$|(\psi, v)| \geqslant c_0 \|v\|^2 = a' \|t\|^2.$$

С другой стороны,

$$|(\psi, v)|^2 \leqslant \sum_{m=1}^{n} |t_m| |(\psi_m, \varphi_m)| \leqslant \|t\|^2 \sum_{m=1}^{n} |(\psi, \varphi_m)|^2.$$

Отсюда, по формуле (23.19),

$$\|\Psi_n t\| = \sum_{m=1}^{n} |(\psi, \varphi_m)|^2 \geqslant \|t\|^{-2} |(\psi, v)|^2 \geqslant a'^2 \|t\|^2,$$

и неравенство (23.8) доказано, причем можно положить $c_0 = a'$.

Таким образом, норма $\|\Psi_n^{-1}\|$ ограничена независимо от n; из полученных выше соотношений $\|M_n^{-1}\| \leqslant \|\Phi_n\| \|\Psi_n^{-1}\|$ и $\|\Phi_n\| \leqslant \leqslant \|B^{-1}A\|$ вытекает, что тем же свойством обладает и матрица M_n^{-1}, так что

$$\|M_n^{-1}\| \leqslant C_1 = \frac{\|B^{-1}A\|}{c}.$$

Систему уравнений (23.13) можно записать в виде $M_n^* \gamma = \delta$, где γ и δ – векторы с составляющими $\gamma_1, \gamma_2, \ldots, \gamma_n$ и $\delta_1, \delta_2, \ldots, \delta_n$ соответственно. Отсюда $\|\delta\| \geqslant C_1^{-1} \|\gamma\|$, и формула (23.14) дает

$$\Gamma \geqslant \|A^{-1}B\|^{-2} C_1^{-2} \|\gamma\|^2;$$

из последнего неравенства вытекает, что собственные числа формы Г ограничены снизу положительным числом, которое не зависит от n, и для интересующего нас случая неравенство (22.3) установлено.

Из теоремы Н. И. Польского следует теперь, что $w_n \underset{n\to\infty}{\longrightarrow} w$ или $Bu_n \underset{n\to\infty}{\longrightarrow} Bu$ в метрике пространства H. Воздействуя на обе части последнего соотношения ограниченным оператором AB^{-1}, получим $Au_n \underset{n\to\infty}{\longrightarrow} Au = f$, что и требовалось доказать.

Замечание 1. Легко указать важный класс случаев, когда оператор B, сходный с оператором A, имеет точечный спектр. Это всегда будет иметь место, когда A — оператор с дискретным спектром. Действительно, как хорошо известно, для этого необходимо и достаточно, чтобы всякое множество, ограниченное в метрике пространства H_A, было компактно в метрике исходного пространства H [1]). Пусть теперь A — оператор с дискретным спектром и B — сходный с A оператор. Пусть множество M ограничено в H_B. Операторы A и B также и полусходные (см. § 3), и по неравенству (3.3) множество M ограничено и в H_A; так как спектр A дискретен, то M компактно в H. Таким образом, множество, ограниченное в H_B, необходимо компактно в H, и спектр оператора B дискретен.

Замечание 2. Условие (23.1) для операторов A и B не необходимо: достаточно, чтобы оно выполнялось для операторов A и $B+kI$, где I — тождественный оператор, а k — подходящая положительная постоянная. Так как собственные элементы операторов B и $B+kI$ совпадают, то и в этом случае применение упомянутых элементов в качестве координатных дает невязку, стремящуюся к нулю.

Введем следующее определение. Пусть A и B — сходные операторы. Назовем оператор B *родственным* оператору A, если при подходящем выборе положительной постоянной k операторы A и $B+kI$ удовлетворяют неравенству (23.1).

---

[1]) Самосопряженный оператор имеет дискретный спектр, если этот спектр — точечный и собственные числа сгущаются только на бесконечности. Достаточность условия, сформулированного в тексте, см., например, ПМ, стр. 57 (теорема 2). Необходимость можно доказать так. Если спектр оператора A дискретен, то оператор A^{-1} вполне непрерывен; таков же и оператор $A^{-\frac{1}{2}}$; если некоторое множество $M_1 \subset H$ ограничено в H, то множество $A^{-\frac{1}{2}}M_1$ компактно в H. Пусть теперь $M \subset H_A$ и $|u|_A \leqslant C = \text{const}$, $u \in M$. Положим $A^{\frac{1}{2}}u = v$. Тогда $\|v\| = \left\|A^{\frac{1}{2}}u\right\| = |u|_A \leqslant C$, и множество $A^{\frac{1}{2}}M$ ограничено в H. Тогда $M = A^{-\frac{1}{2}}A^{\frac{1}{2}}M$ компактно в H, что и требовалось доказать.

Очевидно, в теореме 23.1 можно условие (23.1) заменить условием, чтобы оператор B был родственным оператору A.

В ближайших параграфах настоящей главы будут рассмотрены некоторые классы операторов, к которым применима теорема 23.1.

§ 24. Операторы, различающиеся младшими членами

Пусть A и B — сходные операторы и $C = A - B$. Допустим, что справедливо неравенство

$$\| Cu \| \leqslant \alpha \, | u |_A, \qquad \alpha = \text{const}, \qquad u \in D(A). \tag{24.1}$$

Покажем, что тогда оператор B — родственный оператору A. Действительно, полагая $B' = B + kI$, имеем:

$$(Au,\ B'u) = \| Au \|^2 + k \, | u |_A^2 - (Au,\ Cu).$$

Отсюда

$$(Au,\ B'u) \,| \geqslant \| Au \|^2 + k \, | u |_A^2 - \| Au \| \, \| Cu \| \geqslant$$
$$\geqslant \| Au \|^2 + k \, | u |_A^2 - \alpha \, | u |_A \, \| Au \|.$$

При любом $\varepsilon > 0$ верно неравенство

$$| u |_A \, \| Au \| \leqslant \frac{1}{2\varepsilon} | u |_A^2 + \frac{\varepsilon}{2} \| Au \|^2.$$

Выберем k и ε так, чтобы удовлетворялись неравенства

$$\frac{\alpha\varepsilon}{2} < 1, \qquad k \geqslant \frac{\alpha}{2\varepsilon}.$$

Тогда

$$|(Au,\ B'u)| \geqslant \left(1 - \frac{\alpha\varepsilon}{2}\right) \| Au \|^2 + \left(k - \frac{\alpha}{2\varepsilon}\right) | u |_A^2 \geqslant \left(1 - \frac{\alpha\varepsilon}{2}\right) \| Au \|^2. \tag{24.2}$$

Приведем несколько примеров.

1. Пусть $\tilde{A}$ — невырождающийся эллиптический оператор второго порядка

$$\tilde{A}u = -\sum_{j,\ k=1}^{m} \frac{\partial}{\partial x_j} \left(A_{jk}(x) \frac{\partial u}{\partial x_k} \right) + C(x)\, u \tag{24.3}$$

при краевом условии

$$\left[\sum_{j,\ k=1}^{m} A_{jk} \frac{\partial u}{\partial x_k} \cos(\nu,\ x_j) + \sigma u \right]_S = 0. \tag{24.4}$$

Здесь S — граница конечной области Ω; принимаем, что S — $(m-1)$-мерная дважды непрерывно дифференцируемая поверхность. Далее,

ν — внешняя нормаль к S, σ — ограниченная неотрицательная измеримая функция, заданная на S. Допустим еще, что в замкнутой области $\overline{\Omega} = \Omega \cup S$ коэффициенты A_{jk} непрерывно дифференцируемы дважды, а коэффициент C — один раз. Наконец, примем, что справедливо неравенство

$$\sum_{j,\,k=1}^{m} A_{jk} t_j \overline{t_k} \geqslant \mu \sum_{k=1}^{m} |t_k|^2 \qquad \mu = \text{const} > 0$$

и что

$$C(x) > -\mu_1,$$

где μ_1 — наименьшее собственное число оператора

$$-\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}(x) \frac{\partial u}{\partial x_k}\right)$$

при краевом условии (24.4).

При перечисленных условиях оператор $\widetilde{A}$ — положительно определенный в пространстве $H = L_2(\Omega)$ (ВМ, § 24); будем под A понимать его расширение по Фридрихсу[1]). Опираясь на результаты Ж. Жиро[2]) и О. А. Ладыженской[3]), можно доказать, что область $D(A)$ определения оператора A состоит из функций класса $W_2^{(2)}(\Omega)$, удовлетворяющих, в некотором обобщенном смысле, условию (24.4). Область $D(A)$ не зависит от коэффициента C, поэтому, если мы обозначим через B оператор, который является расширением по Фридрихсу положительно определенного оператора[4])

$$-\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk} \frac{\partial u}{\partial x_k}\right) + u \tag{24.5}$$

при краевом условии (24.4), то операторы A и B сходные. Далее, разность $Cu = Au - Bu$ удовлетворяет неравенству (24.1), так как

$$\|Cu\| \leqslant \max C \, \|u\| \leqslant \frac{\max C}{\gamma} |u|_A,$$

где γ — постоянная, характеризующая положительную определенность оператора $\widetilde{A}$. Если мы будем решать энергетическим методом уравнение

$$\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{j,k} \frac{\partial u}{\partial x_k}\right) + Cu = f(x) \tag{24.6}$$

---

[1]) См. К. Фридрихс [1], а также ПМ, § 5.
[2]) См. статью этого автора [1], а также К. Миранда [1], § 22.
[3]) См. книгу этого автора [1], гл. II, § 3.
[4]) Если $\sigma \not\equiv 0$, то можно слагаемое u в выражении (24.5) не писать.

при краевом условии (24.4), а в качестве координатных выберем собственные функции оператора

$$-\sum_{j,k=1}^{m}\frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right)$$

при том же краевом условии, то невязка уравнения (24.6) будет стремиться к нулю.

Соображения настоящего параграфа можно, разумеется, применить и к более сложным задачам, например, к уравнениям более высокого порядка. Ограничимся одним примером.

2. В пространстве $H=L_2(\Omega)$, где Ω — срединная область пластинки, рассмотрим оператор задачи об изгибе пластинки, подверженной усилиям, действующим в ее срединной плоскости (ВМ, § 18):

$$\Delta^2 w-\frac{h}{D}\left[T_x\frac{\partial^2 w}{\partial x^2}+2T_{xy}\frac{\partial^2 w}{\partial x\,\partial y}+T_y\frac{\partial^2 w}{\partial y^2}\right] \tag{24.7}$$

при краевых условиях

$$w|_S=0,\qquad \left.\frac{\partial w}{\partial \nu}\right|_S=0, \tag{24.8}$$

где S — контур пластинки, ν — внешняя нормаль к S. Через A обозначим расширение этого оператора по Фридрихсу, через B — расширение по Фридрихсу оператора $\Delta^2 w$ при тех же краевых условиях (24.8). Будем считать, что напряжения T_x, T_y, T_z либо растягивающие, либо достаточно малые по абсолютной величине, так что оператор A — положительно определенный. Если контур S и функции T_x, T_{xy}, T_y достаточно гладкие, то, как это вытекает из результатов О. В. Гусевой [1] и А. И. Кошелева [1], области определения операторов A и B совпадают с множеством $\overset{0}{W}{}_2^{(4)}(\Omega)$ функций, которые имеют четвертые обобщенные производные, квадратично суммируемые в Ω, и которые удовлетворяют условиям (24.8). Таким образом, операторы A и B — сходные. Нетрудно видеть, что разность $C=A-B$ удовлетворяет неравенству (24.1). Действительно, по неравенству Буняковского,

$$\|Cw\|^2=\frac{h}{D}\left\|T_x\frac{\partial^2 w}{\partial x^2}+2T_{xy}\frac{\partial^2 w}{\partial x\,\partial y}+T_y\frac{\partial^2 w}{\partial y^2}\right\|^2\leqslant$$

$$\leqslant\frac{h}{D}\int_\Omega\!\!\int\left(T_x^2+2T_{xy}^2+T_y^2\right)dx\,dy\,\times$$

$$\times\int_\Omega\!\!\int\left\{\left|\frac{\partial^2 w}{\partial x^2}\right|^2+2\left|\frac{\partial^2 w}{\partial x\,\partial y}\right|^2+\left|\frac{\partial^2 w}{\partial y^2}\right|^2\right\}dx\,dy.$$

Первый интеграл справа есть постоянная, а второй равен $|w|_B^2$ (см. ВМ, § 27, формула (18)). Таким образом, имеет место неравенство вида

$$\|Cw\| \leqslant \alpha |w|_B, \qquad \alpha = \text{const}; \tag{24.9}$$

так как операторы A и B сходные, то неравенства (24.9) и (24.1) равносильны.

Из доказанного следует, что оператор B родствен оператору A.

Результаты настоящего параграфа могут быть некоторым образом усилены. Чтобы выяснить это обстоятельство, заметим прежде всего следующее. Допустим, что оператор A (а, следовательно, и сходный с ним оператор B) имеет дискретный спектр. Тогда из неравенства (24.1) следует, что оператор CA^{-1} (а с ним и операторы $A^{-1}C$, CB^{-1}, $B^{-1}C$) вполне непрерывен в данном пространстве H. Действительно, из соотношений

$$\left\| CA^{-\frac{1}{2}}u \right\|^2 \leqslant \alpha^2 \left| A^{-\frac{1}{2}}u \right|_A^2 = \alpha^2 \left(AA^{-\frac{1}{2}}u,\ A^{-\frac{1}{2}}u \right) = \alpha^2 \| u \|^2$$

вытекает, что оператор $CA^{-\frac{1}{2}}$ ограничен в H; так как $A^{-\frac{1}{2}}$ в H вполне непрерывен, то таким же будет и оператор $CA^{-1} = CA^{-\frac{1}{2}}A^{-\frac{1}{2}}$. Оказывается, что неравенство (24.1) не необходимо: невязка будет стремиться к нулю, если только произведение CA^{-1} вполне непрерывно в H. Ниже мы даем достаточно подробную формулировку и доказательство этого утверждения.

Рассмотрим уравнение

$$Au = f, \tag{24.10}$$

где A — линейный оператор, действующий в сепарабельном гильбертовом пространстве H. Пусть линейный оператор B таков, что $D(A) = D(B)$. Для приближенного решения уравнения (24.10) иногда используется следующий процесс, называемый «методом моментов» [1]): выбирается координатная последовательность $\varphi_k \in D(A)$ $(k = 1, 2, \ldots)$ и строится приближенное решение в обычном виде

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k; \tag{24.11}$$

коэффициенты $a_k^{(n)}$ определяются из линейной алгебраической системы

$$(Au_n,\ B\varphi_j) = (f,\ B\varphi_j) \qquad (j = 1,\ 2,\ \ldots,\ n). \tag{24.12}$$

---

[1]) См. М. Ф. Кравчук [1].

Теорема 24.1[1]. *Пусть*: 1) *уравнение* (24.10) *разрешимо единственным образом*; 2) *оператор* B^{-1} *ограничен и определен на всем пространстве* H; 3) *оператор* $T = CB^{-1}$, $C = A - B$, *вполне непрерывен в* H; 4) *при любом* n *элементы* $\varphi_1, \varphi_2, \ldots, \varphi_n$ *линейно независимы*; 5) *система* $\{B\varphi_k\}$ *полна в* H. *Тогда для достаточно больших* n *система* (24.12) *разрешима единственным образом и имеют место соотношения*

$$u_n \to u_0, \tag{24.13}$$

$$Au_n \to f, \tag{24.14}$$

где u_0 — *решение уравнения* (24.10).

Доказательство. Положим

$$B\varphi_n = \psi_n, \qquad Bu_n = v_n = \sum_{k=1}^{n} a_k^{(n)} \psi_k, \qquad Bu_0 = v_0.$$

Элемент v_0 удовлетворяет уравнению

$$v_0 + Tv_0 = f, \tag{24.15}$$

в котором оператор T вполне непрерывен; в таком случае к уравнению (24.15) можно применить процесс Бубнова — Галёркина (ВМ, стр. 397, замечание), причем, в силу условий 2), 4) и 5) теоремы, систему $\{\psi_k\}$ можно принять за координатную, а это приведет нас к системе уравнений (24.12). Отсюда сразу вытекает, что эта система единственным образом разрешима, если только n достаточно велико. Далее, имеем $v_n \to v_0$ или $Bu_n \to Bu_0$; так как оператор B^{-1} ограничен, то $u_n \to u_0$. Наконец,

$$Cu_n = Tv_n \to Tv_0 = Cu_0$$

и, следовательно,

$$Au_n = Bu_n + Cu_n \to Bu_0 + Cu_0 = Au_0 = f.$$

Допустим теперь дополнительно, что операторы A и B — самосопряженные и положительно определенные, спектр оператора B точечный и φ_k ($k = 1, 2, \ldots$) суть собственные элементы оператора B. Пусть μ_k — собственное число этого оператора, соответствующее собственному элементу φ_k. Тогда $B\varphi_k = \mu_k \varphi_k$; сократив обе части уравнения (24.12) на μ_j, приведем его к виду

$$(Au_n, \varphi_j) = (f, \varphi_j) \qquad (j = 1, 2, \ldots, n). \tag{24.16}$$

Система (24.16) есть система Ритца для уравнения (24.10); как следствие из теоремы 24.1 мы получаем следующую теорему.

---

[1]) См. статью автора [16].

Теорема 24.2. *Пусть A и B — сходные самосопряженные положительно определенные операторы, действующие в гильбертовом пространстве H, спектр оператора B точечный и оператор $(A-B)B^{-1}$ вполне непрерывен в H. Если φ_k $(k=1, 2, \ldots)$ — полная система собственных элементов оператора B и*

$$u_n = \sum_{k=1}^{n} a_k \varphi_k$$

есть n-е приближенное решение по Ритцу уравнения (24.10), *то при $n \to \infty$ невязка $Au_n - f \to 0$.*

§ 25. Невырождающийся обыкновенный дифференциальный оператор второго порядка

В этом и в ближайших параграфах настоящей главы будут рассмотрены некоторые классы дифференциальных операторов, к которым можно подобрать сравнительно простые родственные операторы.

Рассмотрим задачу

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)\,u = f(x), \tag{25.1}$$

$$u(0) = u(1) = 0 \tag{25.2}$$

при простейших предположениях относительно $p(x)$ и $q(x)$: мы примем, что $p(x)$, $p'(x)$ и $q(x)$ непрерывны при $0 \leqslant x \leqslant 1$ и что $p(x) \geqslant p_0 = \text{const} > 0$, а

$$q(x) > -\mu_1,$$

где μ_1 есть наименьшее собственное число оператора

$$-\frac{d}{dx}\left(p\frac{du}{dx}\right)$$

при краевых условиях (25.2). Левая часть уравнения (25.1) и граничные условия (25.2) порождают оператор, который мы обозначим через $\tilde{A}$ и который действует в пространстве $H = L_2(0, 1)$; этот оператор действует по формуле

$$\tilde{A}u = -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)\,u,$$

а за область его определения примем множество функций, дважды непрерывно дифференцируемых на отрезке $0 \leqslant x \leqslant 1$ и равных нулю на его концах. Оператор $\tilde{A}$ — положительно определенный[1]) и потому

[1]) См. ВМ, § 20.

допускает самосопряженное расширение по Фридрихсу[1]); это расширение обозначим через A. Напомним[1]), что обобщенное решение задачи (25.1) — (25.2), к которому приводит энергетический метод, на самом деле есть решение уравнения $Au = f$.

Выясним, из каких функций состоит область определения $D(A)$ оператора A. Как известно[1]), она состоит из всевозможных обобщенных решений задачи (25.1) — (25.2), соответствующих всевозможным функциям $f \in L_2(0, 1)$. Такие решения во всяком случае принадлежат энергетическому пространству H_A нашей задачи.

Исходя из формулы[2])

$$|u|_A^2 = \int_0^1 (p(x)\,|u'(x)|^2 + q(x)\,|u(x)|^2)\,dx, \tag{25.3}$$

легко доказать, что пространство H_A состоит из функций, которые на отрезке $0 \leqslant x \leqslant 1$ абсолютно непрерывны, имеют суммируемую с квадратом первую производную и удовлетворяют условиям (25.2); обратно, каждая функция, обладающая перечисленными свойствами, входит в H_A.

Обобщенное решение задачи (25.1) — (25.2) реализует в H_A минимум функционала

$$F(u) = |u|_A^2 - 2\,\mathrm{Re}\,(u, f) =$$
$$= \int_0^1 \{p(x)\,|u'(x)|^2 + q(x)\,|u(x)|^2 - 2\,\mathrm{Re}\,(f(x)\,u(x))\}\,dx,$$

поэтому $F(u_0 + \eta) \geqslant F(u_0)$, где η — произвольный элемент пространства H_A. Отсюда легко следует, что

$$\mathrm{Re}\,\{[u_0, \eta]_A - (f, \eta)\} + |\eta|_A^2 \geqslant 0$$

и, следовательно, $\mathrm{Re}\,\{[u_0, \eta] - (f, \eta)\} = 0$. Заменив η на $i\eta$, $i = \sqrt{-1}$, найдем, что $\mathrm{Im}\,\{[u_0, \eta] - (f, \eta)\} = 0$, и окончательно

$$[u_0, \eta]_A - (f, \eta) = 0, \qquad \eta \in H_A. \tag{25.4}$$

Равенство (25.4) означает, что в точке минимума вариация функционала F равна нулю. В рассматриваемом нами примере равенство (25.4) принимает вид

$$\int_0^1 \{p(x)\,u_0'(x)\,\overline{\eta'(x)} + [q(x)\,u_0(x) - f(x)]\,\overline{\eta(x)}\}\,dx = 0. \tag{25.5}$$

---

[1]) См. К. Фридрихс [1], а также ПМ, § 5.
[2]) См. ВМ, § 20.

Второй член преобразуем так:

$$\int_0^1 [q(x)u_0(x)-f(x)]\,\overline{\eta(x)}\,dx =$$
$$= \int_0^1 \overline{\eta(x)}\,\frac{d}{dx}\left(\int_0^x [q(t)u_0(t)-f(t)]\,dt\right)dx =$$
$$= \int_0^1 \overline{\eta'(x)} \int_0^x [q(t)u_0(t)-f(t)]\,dt\,dx;$$

внеинтегральный член исчезает, так как $\eta \in H_A$ и потому $\eta(0) = \eta(1) = 0$. Теперь равенство (25.5) принимает вид

$$\int_0^1 \overline{\eta'(x)}\left\{p(x)u_0'(x) - \int_0^x [q(t)u_0(t)-f(t)]\,dt\right\}dx = 0. \qquad (25.5')$$

Производная $\eta'(x)$ удовлетворяет тождеству

$$\int_0^1 \eta'(x)\,dx = \eta(1)-\eta(0) = 0$$

и принадлежит, следовательно, подпространству $\tilde{L}_2(0, 1)$ пространства $L_2(0, 1)$, ортогональному к единице. Очевидно и обратное: любую функцию из $\tilde{L}_2(0, 1)$ можно рассматривать как производную некоторой функции, являющейся элементом пространства H_A. В самом деле, если $\zeta \in \tilde{L}_2(0, 1)$, то

$$\int_0^1 \zeta(x)\,dx = 0,$$

функция

$$\eta(x) = \int_0^x \zeta(t)\,dt$$

абсолютно непрерывна на отрезке $[0, 1]$, имеет почти всюду производную $\eta'(x) = \zeta(x)$, суммируемую с квадратом на том же отрезке, и $\eta(0) = \eta(1) = 0$. Теперь тождество (25.5′) означает, что выражение в фигурных скобках ортогонально к $\tilde{L}_2(0, 1)$. Но тогда это выражение есть постоянное:

$$p(x)u_0'(x) - \int_0^x [q(t)u_0(t)-f(t)]\,dt = \text{const} \qquad (25.6)$$

Из равенства (25.6) видно, что если $u_0(x) \in D(A)$, то первая производная $u_0'(x)$ абсолютно непрерывна, вторая производная $u_0''(x)$ суммируема с квадратом и почти всюду удовлетворяется дифференциальное уравнение

$$-\frac{d}{dx}\left(p(x)\frac{du_0}{dx}\right)+q(x)u_0=f(x).$$

Обратно, если некоторая функция $u_0(x)$ обладает только что перечисленными свойствами и, разумеется, удовлетворяет условиям (25.2), то $u_0 \in D(A)$.

Действительно, функция $f(x) \in L_2(0, 1)$, где

$$f(x)=-\frac{d}{dx}(pu_0')+qu_0 \tag{25.7}$$

и функция $u_0(x)$ является обобщенным решением задачи (25.1)—(25.2), в которой $f(x)$ определяется равенством (25.7).

Таким образом, в нашем примере множество $D(A)$ совпадает с множеством функций, обладающих следующими свойствами: сами функции, так же как и их первые производные, абсолютно непрерывны на отрезке $0 \leqslant x \leqslant 1$, эти функции обращаются в нуль на его концах, а их вторые производные на том же отрезке суммируемы с квадратом.

Из сказанного следует, в частности, что область $D(A)$ не зависит от конкретного выбора функций $p(x)$ и $q(x)$. Поэтому, если обозначить через B самосопряженное расширение по Фридрихсу оператора задачи (25.1)—(25.2) при

$$p(x)\equiv 1, \qquad q(x)\equiv 0,$$

то $D(A)=D(B)$, т. е. операторы A и B сходные. Собственные функции оператора B суть нетривиальные решения уравнения $u''+\lambda u=0$ при начальных условиях (25.2): они хорошо известны и имеют вид

$$\varphi_n(x)=c_n\sin n\pi x \qquad (n=1,\ 2,\ \ldots); \tag{25.8}$$

коэффициенты c_n могут быть выбраны из того или иного условия нормировки.

Докажем теперь, что построенный выше оператор B родственен оператору A.

Имеем

$$(Au,\ (B+kI)u)=(Au,\ Bu)+k\,|u|_A^2.$$

Далее,

$$(Au,\ Bu)=\int_0^1 \overline{u''(x)}\,\{(p(x)\,u'(x))'-q(x)\,u(x)\}\,dx=$$

$$=\int_0^1 \{p(x)\,|u''(x)|^2+[p'(x)\,u'(x)-q(x)\,u(x)]\,\overline{u''(x)}\}\,dx$$

и, следовательно,

$$|(Au,\ (B+kI)\,u)| \geqslant p_0 \int_0^1 |u''(x)|^2\,dx + k\,|u|_A^2 -$$
$$- \left| \int_0^1 \overline{u''(x)}\,[p'(x)\,u'(x) - q(x)\,u(x)]\,dx \right| =$$
$$= p_0 \|Bu\|^2 + k\,|u|_A^2 - \left| \int_0^1 \overline{u''(x)}\,[p'(x)\,u'(x) - q(x)\,u(x)]\,dx \right|. \quad (25.9)$$

При любом положительном ε справедливо неравенство

$$\left| \int_0^1 \overline{u''(x)}\,[p'(x)\,u'(x) - q(x)\,u(x)]\,dx \right| \leqslant$$
$$\leqslant \frac{\varepsilon}{2} \int_0^1 |u''(x)|^2\,dx + \frac{1}{2\varepsilon} \int_0^1 |p'(x)\,u'(x) - q(x)\,u(x)|^2\,dx. \quad (25.10)$$

По предположению, функции $p'(x)$ и $q(x)$ непрерывны и потому ограничены, и второй интеграл справа не превосходит величины

$$c \int_0^1 \{|u'(x)|^2 + |u(x)|^2\}\,dx, \qquad c = \text{const}.$$

Далее, из тождества

$$u(x) = u(0) + \int_0^x u'(t)\,dt = \int_0^x u'(t)\,dt$$

легко вытекает, что

$$\int_0^1 |u(x)|^2\,dx \leqslant \int_0^1 |u'(x)|^2\,dx,$$

и правая часть формулы (25.10) не превосходит величины

$$\frac{\varepsilon}{2} \int_0^1 |u''(x)|^2\,dx + \frac{c}{\varepsilon} \int_0^1 |u'(x)|^2\,dx = \frac{\varepsilon}{2} \|Bu\|^2 + \frac{c}{\varepsilon} |u|_B^2.$$

Теперь из неравенства (25.9) следует

$$|(Au,\ (B+kI)\,u)| \geqslant \left(p_0 - \frac{\varepsilon}{2}\right) \|Bu\|^2 + k\,|u|_A^2 - \frac{c}{\varepsilon}\,|u|_B^2.$$

Операторы A и B сходные, поэтому

$$\|Bu\|^2 \geqslant c' \|Au\|^2, \qquad |u|_B^2 \leqslant c''\,|u|_A^2, \qquad c',\ c'' = \text{const}.$$

Выберем $\varepsilon < 2p_0$. Тогда

$$|(Au,\ (B+kI)\,u)| \geqslant c'\left(p_0 - \frac{\varepsilon}{2}\right)\|Au\|^2 + \left(k - \frac{cc''}{\varepsilon}\right)|u|_A^2.$$

Теперь достаточно взять $k \geqslant \frac{cc''}{\varepsilon}$, и неравенство (23.1) установлено для операторов A и $B+kI$.

Как это вытекает из теоремы 23.1, если функции (25.8) выбраны как координатные для задачи (25.1) — (25.2) и

$$u_n(x) = \sum_{k=1}^{n} a_k\varphi_k(x) = \sum_{k=1}^{n} a_k \sin k\pi x$$

есть приближенное решение этой задачи по Ритцу, то в метрике пространства $H = L_2(0,\ 1)$ невязка $\|Au_n - f\| \to 0$ или

$$\left\| p\left(u_n'' - u_0''\right) + p'\left(u_n' - u_0'\right) - q\left(u_n - u_0\right)\right\| \xrightarrow[n\to\infty]{} 0.$$

Но $u_n \to u_0$ в метрике H_A; из этого следует, что

$$\|u_n' - u_0'\|_{L_2} \xrightarrow[n\to\infty]{} 0, \qquad \|u_n - u_0\|_{L_2} \xrightarrow[n\to\infty]{} 0,$$

и так как $p(x)$ положительно ограничено снизу, то

$$\|u_n'' - u_0''\|_{L_2} \xrightarrow[n\to\infty]{} 0.$$

Таким образом, если в задаче (25.1) — (25.2) воспользоваться координатными функциями (25.8), то вторые производные приближенных решений сходятся (в метрике $L_2(0,\ 1)$) ко второй производной точного решения. Отсюда, между прочим, следует, что сходимость первых производных — равномерная.

§ 26. Вырождающийся обыкновенный дифференциальный оператор второго порядка

Рассмотрим дифференциальный оператор второго порядка

$$Au = -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)\,u, \tag{26.1}$$

где функции $p(x)$ и $q(x)$ подчинены тем же ограничениям, что и в § 25, за одним исключением: мы примем, что $p(x) > 0$ при $x > 0$ и $p(0) = 0$; более определенно, мы допустим, что

$$p(x) = x^\alpha \tilde{p}(x), \qquad \alpha = \text{const}, \qquad 0 < \alpha < 2, \tag{26.2}$$

где функция $\tilde{p}(x)$ на отрезке $0 \leqslant x \leqslant 1$ непрерывно дифференцируема и строго положительна, так что $\tilde{p}(x) \geqslant p_0$, где p_0 — положительная постоянная.

Приведем некоторые факты, относящиеся к вырождающимся операторам указанного выше вида.

За область $D(A)$ определения оператора A, действующего по формуле (26.1) в пространстве $H = L_2(0, 1)$, примем множество функций $u(x)$, удовлетворяющих следующим условиям:

а) Функция $p(x)u'(x)$ абсолютно непрерывна на отрезке $0 \leqslant x \leqslant 1$.

б) $Au \in L_2(0, 1)$.

в) $u(1) = 0$.

г) Если $1 \leqslant \alpha < 2$, то $p(x)u'(x) = o\left(x^{\frac{1}{2}}\right)$.

Из условий в) и г) вытекает, что при x, близких к нулю,

$$u(x) = -\int_x^1 u'(t)\,dt = \begin{cases} o\left(x^{\frac{3}{2}-\alpha}\right), & \frac{3}{2} < \alpha < 2, \\ o\left(\ln\frac{1}{x}\right), & \alpha = \frac{3}{2}; \end{cases} \tag{26.3}$$

при $0 < \alpha < \frac{3}{2}$ функция $u(x)$ непрерывна в точке $x = 0$. Очевидно, так определенная область $D(A)$ плотна в $L_2(0, 1)$.

д) Если $0 < \alpha < 1$, то $u(0) = 0$.

При таком определении оказывается, что оператор A — самосопряженный и положительно определенный и что его спектр дискретен [1]). Докажем это. Составим скалярное произведение

$$(Au, u) = -\int_0^1 \overline{u(x)}(p(x)u'(x))'\,dx + \int_0^1 q(x)|u(x)|^2\,dx.$$

Первый интеграл возьмем по частям:

$$-\int_0^1 \overline{u(x)}(p(x)u'(x))'\,dx = -\lim_{\delta\to 0}\int_\delta^1 \overline{u(x)}(p(x)u'(x))'\,dx =$$
$$= \lim_{\delta\to 0}\left\{u(\delta)\,p(\delta)\,u'(\delta) + \int_\delta^1 p(x)|u'(x)|^2\,dx\right\}.$$

Докажем, что внеинтегральный член имеет предел, равный нулю. Это очевидно, если $0 < \alpha < 1$, так как тогда $u(\delta) \to 0$, а $p(\delta)u'(\delta)$ имеет конечный предел. Если $1 \leqslant \alpha < \frac{3}{2}$, то $u(\delta)$ имеет конечный предел, а $p(\delta)u'(\delta) \to 0$. Наконец, если $\frac{3}{2} \leqslant \alpha < 2$, то $p(\delta)u'(\delta) =$

---

[1]) Эти утверждения можно получить как следствия общей теории, развитой в книгах А. И. Ахиезера и И. М. Глазмана [1] и М. А. Наймарка [1].

$=o\left(\delta^{\frac{1}{2}}\right)$, а $u(\delta)$ во всяком случае стремится к бесконечности медленнее, чем $\delta^{-\frac{1}{2}}$. Из доказанного следует, что интеграл справа также имеет предел при $\delta\to 0$ и, следовательно,

$$-\int_0^1 \overline{u(x)}\,(p(x)\,u'(x))'\,dx=\int_0^1 p(x)\,|u'(x)|^2\,dx.$$

Теперь

$$(Au,\ u)=\int_0^1 \{p(x)\,|u'(x)|^2+q(x)\,|u(x)|^2\}\,dx, \tag{26.4}$$

выражение $(Au,\ u)$ вещественно (оно даже неотрицательно) и оператор A симметричен.

Докажем теперь, что A — самосопряженный оператор. Для определенности рассмотрим случай, когда $1\leqslant\alpha<2$. Пусть A^* — оператор, сопряженный с A, $u\in D(A)$ и $v\in D(A^*)$. Тогда

$$(Au,\ v)=(u,\ v^*),\qquad v^*\in L_2(0,\ 1).$$

Более подробно

$$-\int_0^1 (p(x)\,u'(x))'\,\overline{v(x)}\,dx+\int_0^1 q(x)\,u(x)\,\overline{v(x)}\,dx=\int_0^1 u(x)\,\overline{v^*(x)}\,dx.$$

Второй интеграл перенесем направо. Полагая

$$v^*(x)-q(x)\,v(x)=v_1(x),\qquad v_1(x)\in L_2(0,\ 1),$$

имеем

$$-\int_0^1 (p(x)\,u'(x))'\,\overline{v(x)}\,dx=\int_0^1 u(x)\,\overline{v_1(x)}\,dx. \tag{26.5}$$

Положим

$$v_2(x)=-\int_0^x v_1(t)\,dt.$$

Функция $v_2(x)$ абсолютно непрерывна на отрезке $0\leqslant x\leqslant 1$ и $v_2(0)=0$. Более того, по неравенству Буняковского,

$$|v_2(x)|\leqslant\sqrt{x}\int_0^x |v_1(t)|^2\,dt$$

и, следовательно, $v_2(x) = o(\sqrt{x})$. Интегрируя по частям, находим

$$\int_0^1 u(x)\overline{v_1(x)}\,dx = -u(x)\overline{v_2(x)}\Big|_0^1 + \int_0^1 u'(x)\overline{v_2(x)}\,dx = \\ = \int_0^1 u'(x)\overline{v_2(x)}\,dx.$$

Положим еще

$$v_3(x) = -\int_x^1 \frac{v_2(t)}{p(t)}\,dt. \tag{26.5'}$$

Произведение $p(x)v_3'(x) = v_2(x)$ абсолютно непрерывно на отрезке $0 \leqslant x \leqslant 1$. Далее, $v_3(1) = 0$. Вычислим в заключение функцию

$$-\frac{d}{dx}\left(p\frac{dv_3}{dx}\right) + qv_3 = v_1 + qv_3;$$

очевидно, что эта функция принадлежит пространству $L_2(0, 1)$.

Из всего сказанного вытекает, между прочим, что $v_3 \in D(A)$. Интегрируя еще раз по частям, получаем

$$\int_0^1 u'(x)\overline{v_2(x)}\,dx = p(x)u'(x)\overline{v_3(x)}\Big|_0^1 - \int_0^1 \overline{v_3(x)}(p(x)u'(x))'\,dx.$$

На верхнем пределе внеинтегральный член исчезает, так как $v_3(1) = 0$; на нижнем пределе он также исчезает: по формуле (26.5') $v_3(x) = o\left(x^{-\frac{1}{2}}\right)$, тогда как $p(x)u'(x) = o\left(x^{\frac{1}{2}}\right)$.

Окончательно

$$\int_0^1 u'(x)\overline{v_2(x)}\,dx = -\int_0^1 \overline{v_3(x)}(p(x)u'(x))'\,dx.$$

Формула (26.5) дает теперь тождество

$$\int_0^1 (p(x)u'(x))'(\overline{v(x)} - \overline{v_3(x)})\,dx = 0, \tag{26.6}$$

верное, если $u \in D(A)$. Докажем, что множество значений оператора $(p(x)u'(x))'$, определенного на $D(A)$, совпадает с $L_2(0, 1)$. Для этого возьмем произвольную функцию $f(x) \in L_2(0, 1)$ и покажем, что существует решение $u \in D(A)$ уравнения

$$(p(x)u'(x))' = f(x).$$

Действительно, таким решением является функция

$$u(x) = -\int_x^1 \frac{f_1(t)}{p(t)}\,dt, \qquad f_1(t) = \int_0^t f(t_1)\,dt_1.$$

Теперь из тождества (26.6) следует, что разность $v - v_3$ ортогональна ко всему пространству $L_2(0, 1)$. Отсюда $v_3(x) \equiv v(x)$; следовательно, $v(x) \in D(A)$ и оператор A самосопряжен.

Нетрудно доказать, что оператор A положительно определен. Действительно, по формуле (26.4)

$$(Au, u) \geqslant p_0 \int_0^1 x^\alpha |u'(x)|^2\,dx. \tag{26.7}$$

Далее,

$$u(x) = -\int_x^1 u'(t)\,dt = -\int_x^1 \frac{1}{t^{\frac{\alpha}{2}}}\, t^{\frac{\alpha}{2}} u'(t)\,dt$$

и, по неравенству Буняковского (для определенности берем $\alpha > 1$),

$$|u(x)|^2 \leqslant \frac{1}{\alpha - 1}\left(\frac{1}{x^{\alpha-1}} - 1\right)\int_0^1 t^\alpha |u'(t)|^2\,dt.$$

Интегрируя это по x, легко найдем

$$(Au, u) \geqslant \gamma^2 \|u\|^2, \qquad \gamma^2 = \frac{p_0}{2 - \alpha}.$$

Докажем теперь, что оператор A имеет дискретный спектр. Для случая $0 < \alpha < 1$ это вытекает из результатов ВМ, § 35, и нам остается рассмотреть случай $1 \leqslant \alpha < 2$.

Из формулы (26.4) легко усмотреть, что в данном случае энергетическое пространство H_A образовано из функций, обладающих следующими свойствами: 1) эти функции абсолютно непрерывны на отрезке $\delta \leqslant x \leqslant 1$ при любом δ, $0 < \delta < 1$; 2) они обращаются в нуль при $x = 1$; 3) они сообщают конечное значение интегралу (26.7).

Пусть $M \subset H_A$ — множество, ограниченное по энергии, так что $|u|_A \leqslant C$, $u \in M$. Докажем, что это множество компактно в $H = L_2(0, 1)$. Прежде всего, в силу неравенства (26.7)

$$\int_0^1 x^\alpha |u'(x)|^2\,dx \leqslant \frac{C^2}{p_0},$$

так что множество M_1 функций $v(x)=x^{\frac{\alpha}{2}}u'(x)$ ограничено в $L_2(0,1)$. Далее, если $u \in M$, то

$$u(x)=-\int_x^1 u'(t)\,dt=-\int_x^1 t^{-\frac{\alpha}{2}}v(t)\,dt,$$

или, если положить

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

то

$$u(x)=\int_0^1 K(x,t)\,v(t)\,dt. \tag{26.8}$$

Докажем, что ядро $K(x,t)$ фредгольмово, т. е. что

$$\int_0^1\int_0^1 K^2(x,t)\,dx\,dt<\infty.$$

Действительно, пусть $1<\alpha<2$. Тогда

$$\int_0^1\int_0^1 K^2(x,t)\,dx\,dt=\int_0^1 \frac{1}{\alpha-1}\left(\frac{1}{x^{\alpha-1}}-1\right)dx=\frac{1}{2-\alpha}.$$

Если $\alpha=1$, то

$$\int_0^1\int_0^1 K^2(x,t)\,dx\,dt=\int_0^1 \ln\frac{1}{x}\,dx=1.$$

Из доказанного следует, что интегральный оператор (26.8) вполне непрерывен в $L_2(0,1)$ и, следовательно, переводит ограниченное множество M_1 в компактное множество M. А тогда (ПМ, стр. 57) оператор A имеет дискретный спектр.

Приведенные выше свойства оператора A (включая и структуру его области определения) не зависят от конкретного вида функций $\tilde{p}(x)$ и $q(x)$. Поэтому, положив $\tilde{p}(x)\equiv 1$, $q(x)\equiv 0$ и сохранив неизменной область определения оператора, мы получим новый оператор

$$Bu=-\frac{d}{dx}\left(x^{\alpha}\frac{du}{dx}\right), \tag{26.9}$$

положительно определенный и сходный с A; спектр нового оператора также дискретен. Докажем, что он родствен оператору A.

Имеем:

$$(Au,\ (B+kI)\,u)=\int_0^1 (x^\alpha \bar{u}')'\,[(\tilde{p}\,(x)\,x^\alpha u')' - qu]\,dx + k\,|\,u\,|_A^2 =$$

$$=\int_0^1 \tilde{p}\,(x)\,|(x^\alpha u')'\,|^2\,dx + k\,|\,u\,|_A^2 + \int_0^1 (x^\alpha u')'\,[\tilde{p}'\,(x)\,x^\alpha u' - q\,(x)\,u]\,dx.$$

Отсюда

$$|(Au,\ (B+kI)\,u)| \geqslant$$

$$\geqslant p_0\,\|\,Bu\,\|^2 + k\,|\,u\,|_A - \left|\int_0^1 (x^\alpha u')'\,[\tilde{p}'\,(x)\,x^\alpha u' - q\,(x)\,u]\,dx\right|.$$

Если ε — любое положительное число, то

$$\left|\int_0^1 (x^\alpha u')'\,[\tilde{p}'\,(x)\,x^\alpha u' - q\,(x)\,u]\,dx\right| \leqslant$$

$$\leqslant \frac{\varepsilon}{2}\,\|Bu\,\|^2 + \frac{1}{2\varepsilon}\int_0^1 [\tilde{p}'\,(x)\,x^\alpha\,|\,u'\,| + q\,(x)\,|\,u\,|]^2\,dx.$$

Обозначим

$$\lambda = \max_{0\leqslant x\leqslant 1} x^\alpha\,(\tilde{p}'\,(x))^2, \qquad \mu = \max_{0\leqslant x\leqslant 1} q^2(x).$$

Тогда

$$\int_0^1 [\tilde{p}'\,(x)\,x^\alpha\,|\,u'\,| + q\,(x)\,|\,u\,|]^2\,dx \leqslant$$

$$\leqslant 2\int_0^1 [(\tilde{p}'\,(x))^2\,x^{2\alpha}\,|\,u'\,|^2 + q^2\,(x)\,|\,u\,|^2]\,dx \leqslant$$

$$\leqslant 2\int_0^1 [\lambda x^\alpha\,|\,u'\,|^2 + \mu\,|\,u\,|^2]\,dx \leqslant 2\lambda\,|\,u\,|_B^2 + 2\mu\,\|\,u\,\|^2.$$

Оператор B — положительно определенный и сходный с оператором A, поэтому имеют место неравенства

$$\|\,u\,\| \leqslant c_1\,|\,u\,|_B \leqslant c_2\,|\,u\,|_A, \qquad c_1,\ c_2 = \text{const};$$

отсюда

$$\int_0^1 [\tilde{p}'\,(x)\,x^\alpha\,|\,u'\,| + q\,(x)\,|\,u\,|]^2\,dx \leqslant c_3\,|\,u\,|_A^2, \qquad c_3 = \text{const},$$

и

$$|(Au,\ (B+kI)\,u)| \geqslant \left(p_0-\frac{\varepsilon}{2}\right)\|Bu\|^2+\left(k-\frac{c_3}{2\varepsilon}\right)|\,u\,|_A^2.$$

Но $\|Au\|^2\leqslant c_4\|Bu\|^2$, $c_4=\mathrm{const}$, поэтому, если взять $\varepsilon<2p_0$ и $k\geqslant\frac{c_3}{2\varepsilon}$, то получим

$$|(Au,\ (B+kI)\,u)| \geqslant \frac{1}{c_4}\left(p_0-\frac{\varepsilon}{2}\right)\|Au\|^2.$$

Таким образом, решая задачу

$$-\frac{d}{dx}\left(\tilde{p}(x)\,x^{\alpha}\frac{du}{dx}\right)+q(x)\,u=f(x) \tag{26.10}$$

при краевых условиях $u(1)=0$ (и если $0<\alpha<1$, то и $u(0)=0$) процессом Ритца и взяв в качестве координатных собственные функции оператора B, мы добьемся сходимости невязки к нулю.

Собственные функции оператора B являются решениями дифференциального уравнения

$$\frac{d}{dx}\left(x^{\alpha}\frac{du}{dx}\right)+\lambda u=0,$$

которое легко приводится к уравнению Бесселя[1]). Отсюда легко найти выражение этих собственных функций:

$$\varphi_n(x)=c_n x^{\frac{1-\alpha}{2}} J_\nu\left(\gamma_{\nu,\,n}x^{\frac{2-\alpha}{2}}\right). \tag{26.11}$$

Здесь $\nu=\left|\frac{\alpha-1}{2-\alpha}\right|$, $\gamma_{\nu,\,n}$ — n-й положительный корень функции Бесселя $J_\nu(z)$; множитель c_n может быть выбран из того или иного условия нормировки.

§ 27. Обыкновенный дифференциальный оператор более высокого порядка

Родственные операторы удается построить и для некоторых дифференциальных операторов порядка выше второго. Так, в пространстве $H=L_2(0,\ 1)$ рассмотрим оператор

$$Au=\frac{d^2}{dx^2}\left(p_2(x)\frac{d^2u}{dx^2}\right)-\frac{d}{dx}\left(p_1(x)\frac{du}{dx}\right)+p_0(x)\,u, \tag{27.1}$$

заданный на функциях, третьи производные которых абсолютно непрерывны на отрезке $0\leqslant x\leqslant 1$, а четвертые производные на том же промежутке квадратично суммируемы; потребуем еще, чтобы функ-

[1]) См., например, В. И. Смирнов [1], п. 49.

ции, на которых мы определяем оператор (27.1), удовлетворяли краевым условиям

$$u(0)=u(1)=0, \qquad u''(0)=u''(1)=0. \tag{27.2}$$

Допустим, что коэффициенты $p_k(x)$ непрерывны и k раз непрерывно дифференцируемы и что $p_0(x)$ и $p_1(x)$ неотрицательны, а $p_2(x)\geqslant c=\mathrm{const}>0$.

Повторяя рассуждения предшествующего параграфа, легко доказать, что оператор A — самосопряженный. Докажем, что он положительно определенный. Интегрируя по частям и учитывая условия (27.2), получим

$$(Au,\ u)=\int_0^1 \{p_2(x)\,|u''(x)|^2+p_1(x)\,|u'(x)|^2+p_0(x)\,|u(x)|^2\}\,dx\geqslant$$

$$\geqslant c\int_0^1 |u''(x)|^2\,dx. \tag{27.3}$$

Напишем неравенство Пуанкаре [1]) для функции $u'(x)$:

$$\int_0^1 |u'(x)|^2\,dx\leqslant\frac{1}{2}\int_0^1 |u''(x)|^2\,dx+\left|\int_0^1 u'(x)\,dx\right|^2.$$

Но в силу краевых условий (27.2)

$$\int_0^1 u'(x)\,dx=u(1)-u(0)=0.$$

---

[1]) Для случая одной переменной, меняющейся на отрезке $0\leqslant x\leqslant 1$, неравенство Пуанкаре имеет вид

$$\int_0^1 |u(x)|^2\,dx\leqslant\frac{1}{2}\int_0^1 |u'(x)|^2\,dx+\left|\int_0^1 u(x)\,dx\right|^2$$

и доказывается очень просто. Имеем

$$u(x)-u(y)=\int_y^x u'(t)\,dt.$$

Отсюда

$$u(x)|^2+|u(y)|^2-u(x)\,\overline{u(y)}-\overline{u(x)}\,u(y)=|u(x)-u(y)|^2\leqslant$$

$$\leqslant|x-y|\left|\int_0^1 |u'(t)|^2\,dt\right|\leqslant\int_0^1 |u'(t)|^2\,dt.$$

Интегрируя это неравенство по x и y в пределах от нуля до единицы, мы и получим неравенство Пуанкаре. О неравенстве Пуанкаре см. Р. Курант и Д. Гильберт [1], гл. VII, и С. Л. Соболев [1].

Теперь

$$\int_0^1 |u''(x)|^2\,dx \geqslant 2\int_0^1 |u'(x)|^2\,dx.$$

Из тех же краевых условий вытекает (см. ВМ, § 20)

$$\int_0^1 |u'(x)|^2\,dx \geqslant \int_0^1 |u(x)|^2\,dx = \|u\|^2.$$

Из всего сказанного следует, что

$$(Au,\ u) \geqslant 2c\,\|u\|^2,$$

и оператор A оказывается положительно определенным.

Оператор B, действующий по формуле

$$Bu = \frac{d^4u}{dx^4} \tag{27.4}$$

и имеющий общую с A область определения, является сходным с оператором A. Докажем, что оператор (27.4) также родственен оператору A.

Имеем

$$(Au,\ (B+kI)\,u) = (Au,\ Bu) + k\,|u|_A^2 =$$
$$= \int_0^1 \overline{u^{(4)}}\,[(p_0u'')'' - (p_1u')' + p_2u]\,dx + k\,|u|_A^2.$$

Отсюда

$$|(Au,\ (B+kI)\,u)| \geqslant c\int_0^1 |u^{(4)}(x)|^2\,dx + k\,|u|_A^2 - 2\left|\int_0^1 p_2'u'''\overline{u^{(4)}}\,dx\right| -$$
$$-\left|\int_0^1 [p_0''u'' - (p_1u')' + p_2u]\,\overline{u^{(4)}}\,dx\right| =$$
$$= c\,\|Bu\|^2 + k\,|u|_A - 2\left|\int_0^1 p_0'u'''\overline{u^{(4)}}\,dx\right| -$$
$$-\left|\int_0^1 [p_2''u'' - (p_1u')' + p_2u]\,\overline{u^{(4)}}\,dx\right|. \tag{27.5}$$

Просто оценивается последний интеграл: он не превосходит величины

$$\frac{\varepsilon}{2}\int_0^1 |u^{(4)}|^2\,dx + \frac{1}{2\varepsilon}\int_0^1 |p_0''u'' - (p_1'u')' + p_2u|^2\,dx.$$

Первый интеграл равен $\|Bu\|^2$, а второй, как нетрудно видеть, оценивается сверху величиной $c_1|u|_A^2$, $c_1 = \text{const}$. Таким образом, последний интеграл в (27.5) имеет оценку

$$\frac{\varepsilon}{2}\|Bu\|^2 + \frac{c_1}{2\varepsilon}|u|_A^2. \qquad (27.6)$$

Далее, производная p_2' ограничена: $|p_2'(x)| \leqslant c_2 = \text{const}$, поэтому

$$\left|\int_0^1 p_2' u''' \overline{u^{(4)}}\, dx\right| \leqslant \frac{c_2\varepsilon}{2}\int_0^1 |u^{(4)}|^2\, dx + \frac{c_2}{2\varepsilon}\int_0^1 |u'''|^2\, dx =$$

$$= \frac{c_2\varepsilon}{2}\|Bu\|^2 + \frac{c_2}{2\varepsilon}\int_0^1 |u'''(x)|^2\, dx. \qquad (27.7)$$

Интегрирование по частям дает

$$\int_0^1 |u'''(x)|^2 dx = \int_0^1 u'''(x)\overline{u'''(x)}\, dx =$$

$$= u''(x)\overline{u'''(x)}\Big|_0^1 - \int_0^1 u''(x)\overline{u^{(4)}(x)}\, dx = -\int_0^1 u''(x)\overline{u^{(4)}(x)}\, dx,$$

внеинтегральные члены исчезают в силу условий (27.2). Далее,

$$\int_0^1 |u'''(x)|^2 dx \leqslant \frac{\varepsilon'}{2}\int_0^1 |u^{(4)}(x)|^2 dx + \frac{1}{2\varepsilon'}\int_0^1 |u''(x)|^2 dx$$

или, если воспользоваться неравенством (27.3),

$$\int_0^1 |u'''(x)|^2 dx \leqslant \frac{\varepsilon'}{2}\|Bu\|^2 + \frac{1}{2c\varepsilon'}|u|_A^2. \qquad (27.8)$$

Из неравенств (27.5) — (27.8) следует:

$$|(Au,\ (B+kI)u)| \geqslant \left(c - \frac{\varepsilon}{2} - \frac{c_2\varepsilon}{2} - \frac{c_2\varepsilon'}{4\varepsilon}\right)\|Bu\|^2 +$$

$$+ \left(k - \frac{c_1}{2\varepsilon} - \frac{c_2}{4c\varepsilon'\varepsilon}\right)|u|_A^2.$$

Выберем ε и ε' столь малыми, а k столь большим, чтобы

$$c - \frac{\varepsilon}{2}(1+c_2) - \frac{c_2\varepsilon'}{4\varepsilon} \geqslant \frac{c}{2}, \qquad k - \frac{c_1}{2\varepsilon} - \frac{c_2}{4c\varepsilon'\varepsilon} \geqslant 0.$$

Тогда

$$|(Au,\ (B+kI)u)| \geqslant \frac{c}{2}\|Bu\|^2. \qquad (27.9)$$

Так как операторы A и B сходные, то последнее неравенство равносильно условию (23.1) для операторов A и $B+kI$.

Собственные функции оператора B нетрудно найти: они удовлетворяют уравнению

$$u^{(4)}-\lambda u=0$$

и краевым условиям (27.2). Эти функции имеют вид

$$\varphi_n(x)=c_n\sin n\pi x \qquad (n=1,\ 2,\ \ldots); \tag{27.10}$$

как всегда, коэффициенты c_n можно определить из тех или иных условий нормировки.

Рассмотрим теперь задачу об интегрировании уравнения

$$Au=\frac{d^2}{dx^2}\left(p_2(x)\frac{d^2u}{dx^2}\right)-\frac{d}{dx}\left(p_1(x)\frac{du}{dx}\right)+p_0(x)\,u=f(x), \tag{27.11}$$
$$f(x)\in L_2(0,\ 1),$$

при краевых условиях (27.2). Если функции (27.10) выбраны как координатные и $u_n(x)$ есть n-е приближенное решение по Ритцу, а $u_0(x)$ — точное решение задачи, то $\|Au_n-f\|\xrightarrow[n\to\infty]{}0$, или $\|A(u_n-u_0)\|\xrightarrow[n\to\infty]{}0$. Операторы A и B — сходные и по неравенству (3.1)

$$\|B(u_n-u_0)\|\leqslant\frac{1}{c_1}\|A(u_n-u_0)\|\xrightarrow[n\to\infty]{}0.$$

Последнее неравенство означает, что

$$\left\|\frac{d^4u_n}{dx^4}-\frac{d^4u_0}{dx^4}\right\|_{L_2}\xrightarrow[n\to\infty]{}0;$$

отсюда, в частности, следует, что разности

$$\frac{d^ku_n}{dx^k}-\frac{d^ku_0}{dx^k} \qquad (k=0,\ 1,\ 2,\ 3)$$

равномерно стремятся к нулю.

При произвольном выборе координатной системы можно только утверждать, что $|u_n-u_0|_A\xrightarrow[n\to\infty]{}0$, или, по неравенству (3.3), что $|u_n-u_0|_B\xrightarrow[n\to\infty]{}0$, т. е. что

$$\left\|\frac{d^2u_n}{dx^2}-\frac{d^2u_0}{dx^2}\right\|_{L_2}\xrightarrow[n\to\infty]{}0.$$

§ 28. Эллиптический оператор второго порядка

Рассмотрим эллиптический оператор второго порядка

$$Au=-\sum_{j,\,k=1}^{m}\frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right)+Cu \tag{28.1}$$

в конечной области Ω пространства координат $x_1, x_2, \ldots, x_m$. Границу S области Ω будем считать дважды непрерывно дифференцируемой. Коэффициент C предполагаем непрерывным, а коэффициенты A_{jk} дважды непрерывно дифференцируемыми в замкнутой области $\bar{\Omega} = \Omega \cup S$. Будем считать также, что оператор A — эллиптический невырождающийся, т. е. что выполняется неравенство

$$\sum_{j,\,k=1}^{m} A_{jk} t_j \bar{t}_k \geqslant \mu \sum_{k=1}^{m} |t_k|^2, \qquad \mu = \mathrm{const} > 0.$$

Оператор A зададим на множестве $\overset{0}{W}{}_2^{(2)}(\Omega)$ функций, имеющих вторые обобщенные производные, квадратично суммируемые в Ω, и удовлетворяющих краевому условию

$$u|_S = 0. \tag{28.2}$$

В работах ряда авторов[1]) для функций множества $\overset{0}{W}{}_2^{(2)}(\Omega)$ установлено неравенство

$$\|Au\|^2 = \int_{\Omega} |Au|^2\,dx \geqslant c \int_{\Omega} \sum_{j,\,k=1}^{m} \left| \frac{\partial^2 u}{\partial x_j\,\partial x_k} \right| dx, \quad c = \mathrm{const} > 0. \tag{28.3}$$

Опираясь на это неравенство, а также на свойства решений задачи Дирихле для эллиптических уравнений с гладкой правой частью[2]), можно доказать, что оператор A — самосопряженный. Наметим доказательство.

Оператор A — положительно определенный в пространстве $H = L_2(\Omega)$ (ВМ, § 24) и потому допускающий самосопряженное расширение по Фридрихсу. Пусть $\bar{A}$ — это расширение. Возьмем произвольную функцию $f(x) \in L_2(\Omega)$ и пусть $u_0(x)$ есть обобщенное решение уравнения $Au_0 = f(x)$, или, что то же, решение уравнения $\bar{A}u_0 = f(x)$; по самому определению оператора A, функция u_0 удовлетворяет краевому условию (28.2). Аппроксимируем $f(x)$ в метрике $L_2(\Omega)$ последовательностью достаточно гладких функций $f_n(x)$.

Уравнение $Au_n = f_n(x)$ имеет решение $u_n(x)$, вторые производные которого во всяком случае непрерывны в $\bar{\Omega}$ (К. Миранда [1], стр. 150, теорема 36.1). Из неравенства (28.3) находим:

$$\int_{\Omega} \left| \frac{\partial^2 u_n}{\partial x_j\,\partial x_k} - \frac{\partial^2 u_p}{\partial x_j\,\partial x_k} \right|^2 dx \leqslant \frac{1}{c}\, \|f_n - f_p\| \xrightarrow[n,\,p\to\infty]{} 0$$
$$(j,\ k = 1,\ 2,\ \ldots,\ m).$$

---

[1]) См. К. Миранда [1].

[2]) См. С. Л. Соболев [1], а также ПМ, стр. 76.

Положительно определенный оператор A имеет обратный, ограниченный в $L_2(\Omega)$ (ПМ, стр. 21), поэтому $\|u_n - u_0\| \to 0$. Отсюда и из замкнутости оператора обобщенного дифференцирования[1]) вытекает, что функция $u_0(x)$ имеет квадратично суммируемые в Ω обобщенные вторые производные. Но в таком случае $D(\bar{A}) \subset \overset{0}{W}{}_2^{(2)}(\Omega) = D(A)$. Но $\bar{A}$ есть расширение A, поэтому $D(\bar{A}) = D(A)$, $\bar{A} = A$ и оператор A — самосопряженный.

На том же множестве $\overset{0}{W}{}_2^{(2)}(\Omega)$ рассмотрим два невырождающихся эллиптических оператора общего вида:

$$A^{(p)}u = -\sum_{j,\,k=1}^{m} A_{jk}^{(p)} \frac{\partial^2 u}{\partial x_j\,\partial x_k} + \sum_{k=1}^{m} A_k^{(p)} \frac{\partial u}{\partial x_k} + A^{(p)}u \qquad (p = 1,\ 2).$$

В работах П. Е. Соболевского [1], [2], использовавшего прием О. А. Ладыженской ([1], гл. II, § 3), установлено неравенство

$$\int_\Omega A^{(1)}u \cdot A^{(2)}u\,dx \geqslant M \int_\Omega \sum_{j,\,k=1}^{m} \left(\frac{\partial^2 u}{\partial x_j\,\partial x_k}\right)^2 dx -$$
$$- N \int_\Omega \left[\sum_{k=1}^{m} \left(\frac{\partial u}{\partial x_k}\right)^2 + u^2\right] dx, \qquad M,\ N = \text{const} > 0. \qquad (28.4)$$

Неравенство (28.4) получено П. Е. Соболевским при следующих предположениях: функция $u(x)$ вещественна; коэффициенты $A^{(p)}$ непрерывны в $\bar{\Omega}$, в той же области $A_k^{(p)}$ непрерывно дифференцируемы один раз, $A_{jk}^{(p)}$ — два раза; наконец, область Ω допускает дважды непрерывно дифференцируемое отображение на шар. От последнего ограничения освободилась О. А. Ладыженская [2].

Мы не воспользуемся непосредственно неравенством (28.4), но используем соображения, которые к нему приводят; их мы вкратце излагаем ниже.

Существенным для построений П. Е. Соболевского и О. А. Ладыженской является получаемое интегрированием по частям тождество (мы записываем его для случая вещественных коэффициентов операторов $A^{(p)}$ и комплексных функций):

$$\int_\Omega A^{(1)}u \cdot \overline{A^{(2)}u}\,dx = \int_\Omega \sum_{j,\,k,\,r,\,s=1}^{m} A_{jk}^{(1)} A_{rs}^{(2)} \frac{\partial^2 u}{\partial x_s\,\partial x_k} \frac{\overline{\partial^2 u}}{\partial x_r\,\partial x_j}\,dx + J, \qquad (28.5)$$

---

[1]) См. С. Л. Соболев [1], а также ПМ, стр. 76.

где J есть сумма некоторых объемных и поверхностных интегралов; вторым важным моментом является то, что, как оказывается, величина J допускает оценку

$$|J| \leqslant \varepsilon \int_{\Omega} \sum_{j,\,k=1}^{m} \left| \frac{\partial^2 u}{\partial x_j \partial x_k} \right|^2 dx + \frac{c}{\varepsilon} \int_{\Omega} \left\{ \sum_{k=1}^{m} \left| \frac{\partial u}{\partial x_k} \right|^2 + |u|^2 \right\} dx, \quad (28.6)$$

где ε — любая, сколь угодно малая постоянная, а c не зависит ни от u, ни от ε.

Дальнейшие рассуждения упомянутых авторов основаны на следующем замечании: матрица a коэффициентов $A_{rs}^{(2)}$ положительно определенная, поэтому существует (также положительно определенная) матрица $\gamma = \|\gamma_{jk}\|$ такая, что $\gamma^2 = a$. Оценим снизу сумму под знаком интеграла (28.4).

Операторы $A^{(p)}$ — эллиптические невырождающиеся, поэтому существуют такие постоянные $\rho_p > 0$, что

$$\sum_{j,\,k=1}^{m} A_{jk}^{(p)} \bar{t}_j t_k \geqslant \rho_p \sum |t_k|^2 \qquad (p = 1,\ 2). \qquad (28.7)$$

Имеем, далее,

$$A_{rs}^{(2)} = \sum_{q=1}^{m} \gamma_{rq} \gamma_{qs} = \sum_{q=1}^{m} \gamma_{rq} \bar{\gamma}_{sq}.$$

Отсюда

$$\sum_{j,\,k,\,r,\,s=1}^{m} A_{jk}^{(1)} A_{rs}^{(2)} \frac{\partial^2 u}{\partial x_s \partial x_k} \frac{\partial^2 \bar{u}}{\partial x_r \partial x_s} =$$

$$= \sum_{j,\,k,\,q=1}^{m} A_{jk}^{(1)} \sum_{r=1}^{m} \gamma_{rq} \frac{\partial^2 u}{\partial x_r \partial x_j} \sum_{s=1}^{m} \bar{\gamma}_{sq} \frac{\partial^2 \bar{u}}{\partial x_s \partial x_k} = \sum_{q=1}^{m} \sum_{j,\,k=1}^{m} A_{jk}^{(1)} \sigma_{qj} \bar{\sigma}_{qk},$$

где положено

$$\sigma_{qk} = \sum_{s=1}^{m} \gamma_{sq} \frac{\partial^2 u}{\partial x_s \partial x_k}.$$

По неравенству (28.7)

$$\sum_{j,\,k=1}^{m} A_{jk}^{(1)} \sigma_{qj} \bar{\sigma}_{qk} \geqslant \rho_1 \sum_{k=1}^{m} |\sigma_{qk}|^2.$$

Теперь

$$\sum_{j,\,k,\,r,\,s=1}^{m} A_{jk}^{(1)} A_{rs}^{(2)} \frac{\partial^2 u}{\partial x_s\,\partial x_k}\,\overline{\frac{\partial^2 u}{\partial x_r\,\partial x_j}} \geqslant \rho_1 \sum_{k,\,q=1}^{m} |\sigma_{qk}|^2 =$$

$$= \rho_1 \sum_{k,\,q=1}^{m} \sigma_{qk}\bar{\sigma}_{qk} = \rho_1 \sum_{q,\,k,\,r,\,s=1}^{m} \gamma_{qr}\gamma_{qs} \frac{\partial^2 u}{\partial x_s\,\partial x_k}\,\overline{\frac{\partial^2 u}{\partial x_r\,\partial x_k}} =$$

$$= \rho_1 \sum_{k,\,r,\,s=1}^{m} A_{rs}^{(2)} \frac{\partial^2 u}{\partial x_s\,\partial x_k}\,\overline{\frac{\partial^2 u}{\partial x_r\,\partial x_k}} \geqslant \rho_1\rho_2 \sum_{r,\,k=1}^{m} \left|\frac{\partial^2 u}{\partial x_r\,\partial x_k}\right|^2. \tag{28.8}$$

Положим теперь $A^{(1)} = A$, $A^{(2)} = B + kI$, где k — положительная постоянная, оператор A определен формулой (28.1), а

$$B = -\Delta = -\sum_{k=1}^{m} \frac{\partial^2 u}{\partial x_k^2} \tag{28.9}$$

и оба оператора определены на множестве $\overset{0}{W}{}_2^{(2)}(\Omega)$. Заметим, что A и B — самосопряженные положительно определенные и сходные операторы.

В силу формул (28.5), (28.6) и (28.8) имеем

$$(Au,\ (B + kI)\,u)| \geqslant (\rho_1\rho_2 - \varepsilon) \int_\Omega \sum_{j,\,k=1}^{m} \left|\frac{\partial^2 u}{\partial x_j\,\partial x_k}\right|^2 +$$

$$+ k\,|u|_A^2 - \frac{c}{\varepsilon} \int_\Omega \left\{ \sum_{k=1}^{m} \left|\frac{\partial u}{\partial x_k}\right|^2 + |u|^2 \right\} dx.$$

Последний интеграл просто оценивается через $|u|_A^2$: существует такая постоянная c_1, что

$$\int_\Omega \left\{ \sum_{k=1}^{m} \left|\frac{\partial u}{\partial x_k}\right|^2 + u^2 \right\} dx \leqslant c_1\,|u|_A^2.$$

Возьмем теперь $\varepsilon = \frac{1}{2}\rho_1\rho_2$, $k = \frac{cc_1}{\varepsilon}$. Тогда

$$|(Au,\ (B + kI)\,u)| \geqslant \frac{\rho_1\rho_2}{2} \int_\Omega \sum_{j,\,k=1}^{m} \left|\frac{\partial^2 u}{\partial x_j\,\partial x_k}\right|^2 dx. \tag{28.10}$$

Теперь из неравенства (28.10) и очевидного неравенства

$$\|Au\|^2 \leqslant c_2\,\|Bu\|^2 \leqslant c_3 \int_\Omega \sum_{j,\,k=1}^{m} \left|\frac{\partial^2 u}{\partial x_j\,\partial x_k}\right|^2 dx, \qquad c_2,\ c_3 = \text{const},$$

вытекает, что оператор (28.9) родствен оператору (28.1).

Из теоремы 23.1 вытекает теперь, что если при решении задачи $Au = f(x)$, где A — определенный выше оператор (28.1), заданный на множестве функций $\overset{0}{W}{}_2^{(2)}(\Omega)$, принять собственные функции оператора Лапласа за координатные, то при $n \to \infty$ невязка $Au_n - f \to 0$; здесь u_n — приближенное решение уравнения $Au = f$ по Ритцу. Отсюда в свою очередь вытекает, что при указанном выборе координатных функций вторые производные приближенных по Ритцу решений сходятся в метрике $L_2(\Omega)$ ко вторым производным точного решения. Действительно, если u_n — приближенное решение по Ритцу, а u_0 — точное решение, то из неравенства (28.3) следует

$$\int_\Omega \left| \frac{\partial^2 u_n}{\partial x_j \partial x_k} - \frac{\partial^2 u_0}{\partial x_j \partial x_k} \right|^2 dx \leqslant \frac{1}{c} \| Au_n - Au_0 \|^2 = \\ = \frac{1}{c} \| Au_n - f \|^2 \xrightarrow[n \to \infty]{} 0.$$

Напомним, что при произвольном выборе координатной системы можно утверждать только, что

$$|u_n - u_0|_A^2 = \\ = \int_\Omega \left\{ \sum_{j,k=r}^{m} A_{jk} \left(\frac{\partial u_n}{\partial x_j} - \frac{\partial u_0}{\partial x_j} \right) \left(\frac{\partial u_n}{\partial x_k} - \frac{\partial u_0}{\partial x_k} \right) + C |u_n - u_0|^2 \right\} dx \xrightarrow[n \to \infty]{} 0,$$

откуда вытекает только, что первые производные приближенных решений сходятся в той же метрике $L_2(\Omega)$ к первым производным точного решения.

ГЛАВА V

О РАЦИОНАЛЬНОМ ВЫБОРЕ КООРДИНАТНОЙ СИСТЕМЫ

§ 29. Общие замечания

Из результатов предшествующих глав, в особенности глав II и IV, вытекает ряд следствий, которые можно охарактеризовать как соображения относительно рационального выбора координатных элементов для процесса Ритца. В настоящей главе мы детально изложим эти соображения, имея в виду главным образом энергетический метод для положительно определенных задач теории дифференциальных уравнений[1]; в конце главы мы кратко рассмотрим вопрос о выборе координатной системы и для метода наименьших квадратов, в частности, для приближенного решения интегральных уравнений.

Вопросами рационального выбора координатной системы мы будем отчасти заниматься также в гл. VI в применении к положительным, но не положительно определенным задачам теории дифференциальных уравнений.

Итак, пусть в сепарабельном гильбертовом пространстве H дано уравнение

$$Au = f, \tag{29.1}$$

где A — положительно определенный оператор.

Имея в виду применить процесс Ритца, выберем систему координатных элементов

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, \tag{29.2}$$

которые прежде всего должны удовлетворять трем условиям § 6. Напомним их: 1) $\varphi_n \in H_A$ ($n = 1, 2, \ldots$); 2) элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы при любом n; 3) система (29.2) полна в H_A. Приближенное решение по Ритцу строится в виде

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k. \tag{29.3}$$

---

[1]) Ряд результатов этого рода содержится в статье автора [18].

Если n велико, то, как мы видели в гл. II, процесс Ритца может оказаться неустойчивым; впрочем, «большое» n может оказаться на самом деле и небольшим: так, в примере § 8 явления неустойчивости делаются достаточно заметными уже при $n=4$. Поэтому перечисленные выше три условия практически позволяют получить приемлемое приближенное решение только при малых n. Если же n сравнительно велико, то следует потребовать, чтобы координатная система была сильно минимальна в метрике H_A. Тут же следует оговориться, что это требование желательно усилить. Действительно, если координатная система сильно минимальна, то процесс Ритца устойчив по отношению к малым погрешностям, допущенным при *составлении* системы Ритца. Однако при *решении* этой системы неизбежные ошибки округления могут привести к большим погрешностям в коэффициентах Ритца. Поэтому, как было выяснено в § 11, желательно, чтобы координатная система была в метрике H_A почти ортонормированной; если координатная система только сильно минимальна, то систему Ритца надо решать с особо большой точностью, что делает вычислительный процесс более трудоемким.

Наконец, как мы видели в гл. IV, в качестве координатной системы имеет смысл выбирать систему собственных элементов (если она полна) оператора B, родственного A, — тогда стремится к нулю невязка $Au_n - f$. Если A — дифференциальный оператор, то, как мы видели в гл. IV, в ряде случаев это означает следующее: старшие входящие в дифференциальное уравнение производные от приближенных решений стремятся (в метрике H) к аналогичным производным от точного решения. При произвольном выборе координатной системы верно только, что $|u_n - u_0|_A \to 0$; в случае дифференциального уравнения это означает сходимость в метрике H производных, порядок которых не превосходит половины порядка уравнения.

Важно подчеркнуть еще одно обстоятельство: если собственные элементы оператора B нормировать в метрике H_B, то полученная система будет почти ортонормированной в H_A. Действительно, упомянутая система ортогональна в H_B (ВМ, стр. 189) и после нормировки в H_B станет в этом пространстве ортонормированной; сходные операторы A и B также и полусходны, и по следствию 4.3 наша система будет в H_A почти ортонормированной.

Можно нормировать собственные элементы оператора B и в метрике H_A; нетрудно видеть, что таким способом мы опять получим систему, почти ортонормированную в H_A. Такая нормировка имеет свои преимущества: величины $|\varphi_k|_A^2$, где φ_k — координатные элементы, все равно приходится вычислять — эти величины суть диагональные элементы матрицы Ритца. Поэтому, если величины $|\varphi_k|_B$ не вычисляются достаточно просто, то нормировка в H_A, по-видимому, более целесообразна.

Таким образом, намечается следующая схема выбора координатной системы для энергетического метода в случае положительно определенного оператора; этой схемы мы будем придерживаться в последующих параграфах настоящей главы, посвященных положительно определенным задачам.

1. Прежде всего будем пытаться строить оператор B со следующими свойствами: 1) он должен быть родственным оператору A; 2) спектр оператора B — точечный. К этому следует добавить, что собственные элементы оператора B должны легко определяться и быть достаточно простыми по своей структуре.

Если всем этим требованиям удалось удовлетворить, то за координатную систему можно взять систему собственных элементов оператора B, нормированную в метрике H_B (или в метрике H_A). При этом: а) процесс Ритца и приближенное решение по Ритцу устойчивы; б) число обусловленности матрицы Ритца ограничено независимо от ее порядка; отсюда, между прочим, следует, что систему Ритца удобно решать итерациями (см. § 12); в) невязка в уравнении стремится к нулю.

2. Если требованиям п. 1 удовлетворить не удается, то можно пытаться строить координатную систему, почти ортонормированную в H_A. При этом будут иметь место утверждения а) и б) п. 1. Так, если B — полусходный с A оператор, то такой будет любая полная ортонормированная в H_B система (см. следствие 4.3). В частности, если B имеет точечный спектр, то можно воспользоваться системой собственных элементов оператора B, нормированной в метрике H_B.

3. Если построение достаточно простой почти ортонормированной в H_A координатной системы окажется затруднительным, то можно пытаться строить координатную систему, сильно минимальную в H_A; при этом будет иметь место утверждение а) п. 1. Сильно минимальную координатную систему можно строить, например, так: если пространство H_A вкладывается в некоторое пространство $\mathfrak{H}$ (§ 4), а система (29.2) полна в H_A и сильно минимальна в $\mathfrak{H}$, то она будет сильно минимальной координатной системой в H_A (см. теорему 4.2). Например, если система (29.2) полна в H_A и ортонормирована в H, то она является сильно минимальной координатной системой в H_A.

Систему, сильно минимальную в энергетической метрике, иногда удается построить совсем просто. Пусть речь идет о первой краевой задаче для самосопряженного невырождающегося эллиптического уравнения, скажем, второго порядка, и пусть область Ω гомеоморфна шару. Пусть $\omega(x_1, x_2, \ldots, x_m) = 0$ — уравнение поверхности S, ограничивающей область Ω; примем, что $\omega \in C^{(2)}(\bar{\Omega})$, $\omega > 0$ внутри Ω и $\frac{\partial\omega}{\partial\nu} > 0$ на S; здесь ν — внешняя нормаль к S. Тогда система функций $\omega \prod_{k=1}^{m} x_k^{n_k} (n_1, n_2, \ldots, n_m = 0, 1, 2, 3, \ldots)$ полна в соответствующем энергетическом пространстве[1]); отсюда легко вытекает, что в том же пространстве полна система

$$\omega \prod_{k=1}^{m} p_{n_k}(x_k) \qquad (n_1, n_2, \ldots, n_m = 0, 1, 2, \ldots), \tag{29.4}$$

где $p_n(x)$ — нормированные полиномы Лежандра.

Внутри Ω возьмем некоторый куб Q; поместим начало координат в центре этого куба, оси координат направим параллельно его ребрам и единицу длины выберем так, чтобы ребро куба имело длину 2. В пространстве $L_2(Q)$ оператор умножения на ω положительно определен и ограничен, поэтому его энергетическое пространство совпадает с $L_2(Q)$. По следствию 4.2, система (29.4) сильно минимальна в $L_2(Q)$.

Каждой функции из $L_2(\Omega)$ приведем в соответствие ту же функцию, определенную только в Q; этим определено вложение $L_2(\Omega)$ в $L_2(Q)$, причем, очевидно, $\|u\|_{L_2(\Omega)} \geqslant \|u\|_{L_2(Q)}$. По теореме 4.2 система (29.4) сильно минимальна в $L_2(\Omega)$; по той же теореме она сильно минимальна и в энергетическом пространстве нашей задачи. Разумеется, оператор этой задачи предполагается положительно определенным[2]).

Замечание. К сказанному в настоящем параграфе необходимо добавить следующее: наряду с отмеченными выше положительными результатами, выбор системы собственных элементов родственного оператора в качестве координатной приводит в ряде случаев и к некоторым нежелательным последствиям. Рассмотрим этот вопрос подробнее.

При указанном выборе координатной системы невязка $Au_n - f$, где u_n — приближенное по Ритцу решение уравнения $Au = f$, стремится к нулю в метрике того пространства, в котором задача рассматривается. Такой выбор дает, следовательно, сходимость, улучшенную по сравнению с той, которую могла бы дать произвольно выбранная координатная система. Однако дальнейшее улучшение сходимости на этом пути, по-видимому, невозможно или, по крайней мере, весьма затруднено. Поясним это следующими соображениями.

Пусть A — дифференциальный оператор, заданный на функциях, определенных в конечной области Ω и удовлетворяющих некоторым однородным краевым условиям; пусть этот оператор положительно определен в пространстве $L_2(\Omega)$. Если координатные функции суть

[1]) См. Л. В. Канторович и В. И. Крылов [1], стр. 294—295.

[2]) Изложенным здесь замечанием автор обязан М. Н. Яковлеву.

собственные функции родственного оператора, то невязка стремится к нулю в среднем. Однако в общем случае она не будет стремиться к нулю, например, равномерно; убедиться в этом можно на следующем простом примере.

Предположим, что данное уравнение имеет вид

$$Au = -\Delta u + C(x)u = f(x), \qquad u|_S = 0, \tag{29.5}$$

и пусть Ω — конечная область с достаточно гладкой границей S, а $C(x) \geqslant 0$. Родственным оператору A является оператор B, где

$$Bu = -\Delta u, \qquad u|_S = 0. \tag{29.6}$$

Спектр оператора B дискретен; пусть

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots$$

— полная система его ортонормированных в $L_2(\Omega)$ собственных функций и пусть

$$u_n(x) = \sum_{k=1}^{n} a_k^{(n)} \varphi_k(x)$$

есть приближенное по Ритцу решение задачи (29.5). По теореме 23.1 $Au_n \to f$ в метрике $L_2(\Omega)$; если обозначить через $u_0(x)$ точное решение задачи (29.5), то

$$-\Delta u_n + C(x)u_n \xrightarrow[L_2(\Omega)]{} -\Delta u + C(x)u,$$

и так как $u_n \xrightarrow[L_2(\Omega)]{} u_0$, то и $\Delta u_n \xrightarrow[L_2(\Omega)]{} \Delta u_0$. В то же время невозможно, чтобы в общем случае Δu_n стремилось к Δu_0 равномерно в $\bar{\Omega} = \Omega \cup S$. Действительно, если бы такая сходимость имела место, то было бы справедливо и такое соотношение:

$$\Delta u_n|_S \rightrightarrows \Delta u_0|_S,$$

где знак $\rightrightarrows$ означает равномерную сходимость. Но

$$\Delta u_n = \sum_{k=1}^{n} a_k^{(n)} \Delta\varphi_k = -\sum_{k=1}^{n} a_k^{(n)} \mu_k \varphi_k,$$

где μ_k — собственные числа оператора (29.6). В таком случае $\Delta u_n|_S = 0$; в то же время из уравнений (29.5) следует, что

$$\Delta u_0|_S = -f(x)|_S + C(x)u_0|_S = -f(x)|_S,$$

а заданная функция $f(x)$ вовсе не обязана обращаться в нуль на границе области. Из доказанного следует, что в рассматриваемой задаче производные второго и более высоких порядков от приближенных решений не будут, вообще говоря, стремиться равномерно к соответствующим производным точного решения. Можно высказать и более общее утверждение: если процессом Ритца решается дифференциальное уравнение (или система таких уравнений) порядка k и за координатные функции взяты собственные функции родственного оператора, то в общем случае производные порядка $\geqslant k$ от приближенных решений не сходятся равномерно в $\overline{\Omega}$ к соответствующим производным точного решения.

По поводу теоремы 23.1 следует сделать еще одно замечание: если в качестве координатных взяты собственные функции родственного оператора, то получающиеся приближения могут не очень быстро сходиться к точному решению. Так, рассмотрим задачу

$$Au = -\frac{d}{dk}\left(p(x)\frac{du}{dx}\right) + q(x)u = f(x), \qquad u(0) = u(\pi) = 0, \tag{29.7}$$

где $p(x)$, $p'(x)$, $q(x)$ — непрерывные на отрезке $[0, \pi]$ функции, причем $p(x) > 0$, $q(x) \geqslant 0$. Оператору задачи (29.7) родствен оператор B:

$$Bu = -\frac{d^2u}{dx^2}, \qquad u(0) = u(\pi) = 0,$$

собственные функции которого суть $\varphi_k(x) = \sqrt{\frac{2}{\pi}} \sin kx$. Если

$$u_n(x) = \sqrt{\frac{2}{\pi}} \sum_{k=1}^{n} a_k^{(n)} \sin k\pi x$$

есть приближенное по Ритцу решение задачи (29.7), а $u_0(x)$ — ее точное решение, то[1] $|u_n - u_0| = O\left(n^{-\frac{3}{2}}\right)$; улучшить эту оценку можно только в том случае, когда $f(x)$ — свободный член уравнения (29.7) — удовлетворяет некоторым специальным условиям, в частности условиям $f(0) = f(\pi) = 0$. Если же в качестве координатных функций взяты полиномы, то, как показал И. К. Даугавет [1][2], оценка величины $|u_n - u_0|$ улучшается с улучшением аналитических свойств решения и функций p, q, f: если эти последние функции достаточно гладкие, а решение $u_0(x)$ имеет r непрерывных производных, последняя из которых удовлетворяет условию Липшица с показателем α, то

$$\left|\frac{d^s(u_n - u_0)}{dx^s}\right| = \begin{cases} O\left(n^{-r-\alpha+\frac{3}{2}} \ln n\right), & s = 1, \\ O\left(n^{-r-\alpha-\frac{5}{2}+4s}\right), & s \geqslant 2. \end{cases}$$

[1]) См. ПМ, § 38 и цитированную там литературу.

[2]) В этой работе рассмотрены обыкновенные дифференциальные уравнения без вырождения не только второго, но и любого четного порядка при весьма общих краевых условиях. Некоторые новые результаты содержатся в статье И. К. Даугавета [2].

Отсюда следует, что при достаточно большом r и достаточно гладких p, q, f величина $|u_n - u_0|$ убывает как $n^{-r-\alpha+\frac{3}{2}}\ln n$, а при $r \geqslant 6$ невязка стремится к нулю, и притом равномерно.

Изложенные здесь соображения выдвигают задачу, которая кажется автору достаточно важной: найти метод подбора таких координатных функций, которые, обеспечивая выполнение требований а) — в) настоящего параграфа, в то же время давали бы сходимость производных по возможности высокого порядка и возможно более быструю сходимость в энергетической метрике. Как мы отметили выше, для обыкновенных дифференциальных уравнений такими координатными функциями, согласно результатам, полученным И. К. Даугаветом, являются полиномы, удовлетворяющие всем краевым условиям задачи.

Вместе с тем, автор полагает, что метод выбора координатных функций, основанный на теореме 23.1, имеет и свои преимущества. Этот метод — довольно общий и может быть применен ко многим классам дифференциальных уравнений как обыкновенных, так и в частных производных; он требует минимальных ограничений на коэффициенты уравнения, на его свободный член и на область интегрирования.

§ 30. Обыкновенные дифференциальные уравнения второго порядка

Всюду в этом параграфе $H = L_2(0, 1)$.

1. Пусть оператор A определяется соотношениями:

$$Au = -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u, \qquad 0 < x < 1, \tag{30.1}$$

$$u(0) = u(1) = 0. \tag{30.2}$$

Примем пока, как и в § 25, что функции $p(x)$, $p'(x)$, $q(x)$ непрерывны на отрезке $0 \leqslant x \leqslant 1$ и что $p(x) \geqslant p_0$, где p_0 — положительная постоянная, а $q(x) \geqslant -\mu_1$, где μ_1 — наименьшее собственное значение оператора

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) \tag{30.3}$$

при краевых условиях (30.2). Как было установлено в § 25, оператор B, где

$$Bu = -\frac{d^2u}{dx^2}, \qquad u(0) = u(1) = 0, \tag{30.4}$$

— родственный оператору A.

Для задачи (30.1) — (30.2) можно выбрать в качестве координатных собственные функции оператора B, нормированные в H_B.

Напомним, что метрика в H_B определяется равенствами

$$[u, v]_B = \int_0^1 u'(x)\overline{v'(x)}\,dx, \qquad |u|_B^2 = \int_0^1 |u'(x)|^2\,dx.$$

Эти функции (см. формулу (25.8)) имеют вид $\varphi_n(x) = c_n \sin n\pi x$; коэффициент c_n определяется из условия

$$|\varphi_n|_B^2 = \int_0^1 [\varphi_n'(x)]^2\,dx = cn^2n^2\pi^2 \int_0^1 \cos^2 n\pi x\,dx = 1.$$

Отсюда

$$c_n = \frac{\sqrt{2}}{n\pi}$$

и

$$\varphi_n(x) = \frac{\sqrt{2}}{n\pi}\sin n\pi x \qquad (n = 1, 2, \ldots). \tag{30.5}$$

При таком выборе координатных функций имеют место утверждения а) — в) § 29.

Для задачи (30.1) — (30.2) можно указать еще одну довольно простую и удобную систему координатных функций; это полиномы

$$Q_n(x) = c_n \int_0^1 P_n(2t-1)\,dt \qquad (n = 1, 2, \ldots),$$

где P_n — полиномы Лежандра, а c_n определяются из условия $|Q_n|_B = 1$. Прежде всего, $Q_n \in D(A)$. Действительно, эти функции имеют производные всех порядков и удовлетворяют краевым условиям (30.2): равенство $Q_n(0) = 0$ очевидно, а равенство

$$0 = Q_n(1) = c_n \int_0^1 P_n(2t-1)\,dt = \frac{1}{2}c_n \int_{-1}^1 P_n(z)\,dz$$

следует из того, что на отрезке $[-1, 1]$ полиномы Лежандра $P_n(z)$, $n > 0$, ортогональны к полиному $P_0(z) \equiv 1$. Далее, полиномы $Q_n(x)$ ортонормированы в H_B. То, что они нормированы, определяется выбором постоянной c_n, а ортогональность следует из цепочки равенств

$$[Q_n, Q_m]_B = \int_0^1 Q_n'(x)\overline{Q_m'(x)}\,dx = c_n c_m \int_0^1 P_n(2x-1)P_m(2x-1)\,dx =$$

$$= \frac{c_n c_m}{2}\int_{-1}^1 P_n(z)P_m(z)\,dz = \frac{c_n c_m}{2n+1}\delta_{nm};$$

отсюда же находим и значение c_n:

$$c_n = \sqrt{2n+1},$$

и следовательно,

$$Q_n(x) = \sqrt{2n+1} \int_0^x P_n(2t-1)\,dt. \tag{30.6}$$

Как это следует из определения нормы в H_B, сходимость в этом пространстве означает сходимость производных в $L_2(0, 1)$. Отсюда (ср. § 8) легко вытекает, что система (30.6) полна в H_B.

Из всего сказанного вытекает, что система (30.6) полна и почти ортонормирована в метрике H_A; если указанную систему выбрать в качестве координатной для задачи (30.1) — (30.2), то будут иметь место утверждения а) и б) § 29. Из приведенных в том же § 29 результатов И. К. Даугавета следует, что для системы (30.6) верно и утверждение в) о сходимости невязки, если только решение задачи (30.1) — (30.2) имеет 6 непрерывных производных, последняя из которых удовлетворяет условию Липшица с положительным показателем. Заметим еще, что множитель $\sqrt{2n+1}$ в формуле (30.6) можно заменить на $[\sqrt{n}]$.

2. По-прежнему будем рассматривать задачу (30.1) — (30.2), но при более общих предположениях относительно коэффициентов $p(x)$ и $q(x)$. Будем считать, что функция $p(x)$ непрерывна и положительно ограничена снизу: $p(x) \geqslant p_0 = \text{const} > 0$, но ее производная может иметь конечное число разрывов первого рода. Отсюда, между прочим, следует, что $p(x)$ удовлетворяет условию Липшица с показателем единица и, следовательно, абсолютно непрерывна на отрезке $0 \leqslant x \leqslant 1$. Относительно $q(x)$ допустим, что эта функция измерима и ограничена и что $q(x) \geqslant -\mu_1 + \delta$, $\delta = \text{const} > 0$.

Проследив за рассуждениями § 25, легко убедиться, что соотношение (25.6) остается справедливым, отсюда следует, что произведение $p(x)\,u_0'(x)$, где $u_0'(x)$ — обобщенное решение задачи (30.1) — (30.2), абсолютно непрерывно; функция $p(x)$ не обращается в нуль и абсолютно непрерывна, поэтому[1]) абсолютно непрерывна производная $u_0'(x)$. Теперь из формулы (25.6) вытекает, что вторая производная $u_0''(x)$ суммируема на отрезке [0, 1] с квадратом.

Таким образом, в условиях настоящего пункта область определения оператора A (точнее, его расширения по Фридрихсу) такая же, как и в п. 1, и оператор B по-прежнему родствен оператору A. Отсюда следует, что все выводы п. 1 остаются в силе и в настоящем случае.

---

[1]) См. И. П. Натансон [2].

3. Положение меняется, если коэффициент $p(x)$ разрывен. Примем, например, все допущения п. 2, кроме допущения о непрерывности $p(x)$; пусть, для простоты, этот коэффициент разрывен в одной точке a, $0<a<1$, где он терпит скачок первого рода.

Если функции, входящие в область определения оператора A, удовлетворяют только краевым условиям (30.2), то этот оператор просто несимметричен. Действительно, допуская, что $u(0)=u(1)=v(0)=v(1)=0$ и что проводимые ниже выкладки законны, имеем:

$$
\begin{gathered}
(Au,\ v)=\int_0^1 \overline{v(x)}\,\{-(p(x)\,u'(x))'+q(x)\,u(x)\}\,dx= \\
=-\int_0^a \overline{v(x)}\,(p(x)\,u'(x))'\,dx-\int_a^1 \overline{v(x)}\,(p(x)\,u'(x))'\,dx+ \\
+\int_0^1 q(x)\,u(x)\,\overline{v(x)}\,dx=-\,p(x)\,u'(x)\,\overline{v(x)}\Big|_0^a-p(x)\,u'(x)\,\overline{v(x)}\Big|_a^1+ \\
+\int_0^1 \{p(x)\,u'(x)\,\overline{v'(x)}+q(x)\,u(x)\,\overline{v(x)}\}\,dx. \qquad (30.7)
\end{gathered}
$$

Приняв во внимание условия (30.2), получим:

$$
\begin{gathered}
(Au,\ v)=u'(a)\,\overline{v(a)}\,[p(a+0)-p(a-0)]+ \\
+\int_0^1 \{p(x)\,u'(x)\,\overline{v'(x)}+q(x)\,u(x)\,\overline{v(x)}\}\,dx;
\end{gathered}
$$

ясно, что последнее выражение несимметрично относительно u и v.

Оператор A, действующий по формуле (30.1), станет симметричным, если задать его, например, на множестве функций, удовлетворяющих следующим условиям: 1) они непрерывны на отрезке $0\leqslant x\leqslant 1$ и обращаются в нуль на его концах; 2) их первые производные терпят разрыв первого рода в точке a, так что выполняется равенство

$$
p(a-0)\,u'(a-0)=p(a+0)\,u'(a+0); \qquad (30.8)
$$

3) если доопределить функцию $u'(x)$, положив ее в точке a равной $u'(a-0)$, то эта функция станет абсолютно непрерывной на отрезке $0\leqslant x\leqslant a$; если же за значение функции $u'(x)$ в точке a принять величину $u'(a+0)$, то $u'(x)$ станет абсолютно непрерывной на отрезке $a\leqslant x\leqslant 1$; 4) вторая производная $u''(x)$ суммируема с квадратом на отрезке $0\leqslant x\leqslant 1$.

Действительно, в этом случае формула (30.7) дает:

$$(Au,\ v) = [p(a+0)\,u'(a+0) - p(a-0)\,u'(a-0)]\,\overline{v(a)} + \int_0^1 \{p(x)\,u'(x)\,\overline{v'(x)} + q(x)\,u(x)\,\overline{v(x)}\}\,dx = \int_0^1 \{p(x)\,u'(x)\,\overline{v'(x)} + q(x)\,u(x)\,\overline{v(x)}\}\,dx, \tag{30.9}$$

что уже симметрично относительно u и v. Более того, из формулы (30.9) следует, что

$$(Au,\ u) = \int_0^1 \{p(x)|u'(x)|^2 + q(x)|u(x)|^2\}\,dx \geqslant \delta \int_0^1 |u(x)|^2\,dx = \delta\|u\|^2,$$

т. е. что так заданный оператор A — положительно определенный. Повторяя рассуждения § 25, нетрудно убедиться, что этот оператор и самосопряженный. Легко построить сравнительно простой сходный с A оператор: таким будет, например, оператор B_1, область определения которого $D(B_1) = D(A)$ и который действует по формуле

$$B_1 u = \begin{cases} p(a-0)\,u'', & 0 < x < a, \\ p(a+0)\,u'', & a < x < 1. \end{cases}$$

Те же соображения, что и в § 25, показывают, что оператор B_1 родствен оператору A.

Можно, следовательно, для решения уравнения (30.1) при краевых условиях (30.2) и (30.8) использовать в качестве координатных собственные функции оператора B_1. Это, однако, затруднительно, так как определение этих функций связано с решением довольно сложного трансцендентного уравнения. Покажем это. Указанные функции удовлетворяют краевым условиям (30.2) и (30.8) и дифференциальным уравнениям

$$\left.\begin{aligned} p(a-0)\,u'' + \lambda u = 0, &\quad 0 < x < a, \\ p(a+0)\,u'' + \lambda u = 0, &\quad a < x < 1; \end{aligned}\right\}$$

кроме того, эти функции непрерывны при $x = a$.

Полагая для удобства $\lambda = \mu^2$, $p(a-0) = r^{-2}$, $p(a+0) = s^{-2}$, найдем:

$$u = \begin{cases} C_1 \cos r\mu x + C_2 \sin r\mu x, & 0 < x < a, \\ C_3 \cos s\mu x + C_4 \sin s\mu x, & a < x < 1. \end{cases}$$

Условия (30.2) и (30.8), а также условие непрерывности $u(x)$ в точке a дают:

$$\begin{aligned}
&C_1 = 0,\\
&C_3 \cos s\mu + C_4 \sin s\mu = 0,\\
&C_2 \sin r\mu a - C_3 \cos s\mu a - C_4 \sin s\mu a = 0,\\
&C_2 r^{-1}\mu \cos r\mu a + C_3 s^{-1}\mu \sin s\mu a - C_4 s^{-1}\mu \cos s\mu a = 0.
\end{aligned}$$

Отсюда для μ получается уравнение

$$\begin{vmatrix} 0 & \cos s\mu & \sin s\mu \\ \sin r\mu a & -\cos s\mu a & -\sin s\mu a \\ r^{-1}\mu \cos r\mu a & s^{-1}\mu \sin s\mu a & -s^{-1}\mu \cos s\mu a \end{vmatrix} = 0.$$

Вычисление корней этого уравнения представляет собой задачу, достаточно трудоемкую. Нам поэтому кажется целесообразным в данном случае отказаться от использования собственных функций родственного оператора и взять в качестве координатной систему, полную и ортонормированную в энергетической метрике какого-нибудь полусходного оператора. В данном случае, как легко видеть, полусходным с A будет оператор B (в формуле (30.4)), и в качестве координатной можно взять, например, любую из систем (30.5) или (30.6); при этом будут иметь место утверждения а) и б) § 29.

4. Пусть теперь оператор A определяется дифференциальным выражением (30.1) при краевых условиях

$$u(0) = 0, \qquad u'(1) = 0; \tag{30.10}$$

мы сохраняем предположения п. 1, но под μ_1 будем понимать наименьшее собственное значение оператора (30.3) при краевых условиях (30.10). В данном случае сходным с оператором A будет оператор $B_2 = -\dfrac{d^2}{dx^2}$ при краевых условиях (30.10); его собственные функции, нормированные в H_{B_2}, суть

$$\varphi_n(x) = \frac{2\sqrt{2}}{(2n-1)\pi} \sin \frac{2n-1}{2} \pi x \qquad (n = 1,\ 2,\ \ldots). \tag{30.11}$$

Легко доказать (ср. § 25), что оператор B_2 родствен оператору задачи (30.1), (30.10), поэтому если при решении задачи (30.1), (30.10) взять функции (30.11) за координатные, то будут иметь место утверждения а) — в) § 29.

5. Сохраним предположения п. 1, кроме одного: будем считать, что $q(x) \geqslant q_0 = \mathrm{const} > 0$, и пусть оператор A определяется дифференциальным выражением (30.1) и краевыми условиями

$$u'(0) = u'(1) = 0. \tag{30.12}$$

Сходным с A в данном случае будет оператор B_3, определяемый дифференциальным выражением $B_3u = -\frac{d^2u}{dx^2} + u$ и краевыми условиями (30.12); его нормированные в H_{B_3} собственные функции суть

$$\varphi_0(x) = 1, \quad \varphi_n(x) = \sqrt{\frac{2}{n^2\pi^2+1}} \cos n\pi x \qquad (n = 1, 2, \ldots). \tag{30.13}$$

На данный случай переносится без изменений все, сказанное в конце предшествующего п. 4.

6. Рассмотрим теперь случай, когда оператор A определяется дифференциальным выражением (30.1) и краевыми условиями

$$u'(0) - \alpha u(0) = 0, \qquad u'(1) + \beta u(1) = 0, \qquad \alpha,\ \beta > 0; \tag{30.14}$$

примем, что $p(x)$ и $q(x)$ удовлетворяют условиям п. 1, где на этот раз μ_1 есть наименьшее собственное значение оператора (30.3) при краевых условиях (30.14). В рассматриваемом случае родственным будет оператор B_4, определяемый дифференциальным выражением $B_4 = -\frac{d^2}{dx^2}$ и краевыми условиями (30.14).

Однако построение собственных функций этого оператора затруднительно, так как оно связано с решением некоторого трансцендентного уравнения. Но в этом случае нетрудно построить полусходный с A оператор и ортонормированную в метрике этого оператора полную систему. Так, полусходным с оператором A настоящего примера является оператор B_3 примера п. 5: соответствующие пространства H_A и H_{B_3} состоят из абсолютно непрерывных на отрезке $0 \leqslant x \leqslant 1$ функций, производные которых на этом отрезке квадратично суммируемы. При решении задачи (30.1), (30.14) можно использовать в качестве координатных функций (30.13); при этом имеют место утверждения а) и б) § 29.

В качестве полусходного можно использовать также оператор $\tilde{B}$, определяемый соотношениями

$$\tilde{B}u = -\frac{d^2u}{dx^2}, \qquad u'(0) - u(0) = 0, \qquad u'(1) = 0;$$

ему соответствуют энергетические произведения и норма

$$\left.\begin{aligned} [u, v]_{\tilde{B}} &= u(0)\overline{v(0)} + \int_0^1 u'(x)\overline{v'(x)}\,dx, \\ |u|_{\tilde{B}} &= |u(0)|^2 + \int_0^1 |u'(x)|^2\,dx; \end{aligned}\right\} \tag{30.15}$$

из формул (30.15) сразу видно, что пространства H_A и $H_{\tilde{B}}$ состоят из одних и тех же элементов. Нетрудно построить полную и ортонормированную в $H_{\tilde{B}}$ систему функций — такова, например, система

$$1,\ x,\ \frac{\sqrt{2}}{\pi}\sin\pi x,\ \frac{\sqrt{2}}{2\pi}\sin 2\pi x,\ \ldots,\ \frac{\sqrt{2}}{n\pi}\sin n\pi x,\ \ldots \tag{30.16}$$

Очевидно, что в метрике (30.15) эта система ортонормирована; надо доказать только ее полноту. Для этого выясним, какие функции ортогональны в метрике (30.15) к функциям

$$\frac{\sqrt{2}}{n\pi}\sin n\pi x \qquad (n=1,\ 2,\ \ldots).$$

Пусть $\omega(x)$ — такая функция. Тогда

$$\left[\omega(x),\ \frac{\sqrt{2}}{n\pi}\sin n\pi x\right]_{\tilde{B}} = \sqrt{2}\int_0^1 \omega'(x)\cos n\pi x\,dx = 0 \quad (n=1,\ 2,\ \ldots).$$

Отсюда $\omega'(x)=c$ и $\omega(x)=cx+c_1$, где c, $c_1=\mathrm{const}$, и полнота системы (30.15) доказана. Использование системы (30.16) в качестве координатной также приводит к устойчивому приближенному решению задачи (30.1) и (30.14) и к матрице Ритца с ограниченным числом обусловленности.

7. Аналогично тому, как это было сделано в пп. 2 и 3, можно рассмотреть краевые задачи последующих пунктов в предположении разрывности тех или иных из функций $p(x)$, $p'(x)$, $q(x)$. Этот анализ мы предоставляем сделать читателю.

Замечание. Пусть A и B — полусходные операторы. Как мы знаем, система элементов, почти ортонормированная в метрике H_B, обладает тем же свойством и в метрике H_A. В настоящем параграфе мы действовали так: исходя из системы, ортогональной в H_B, мы ее нормировали в метрике H_B умножением на подходящие множители; это давало нам систему, ортонормированную в H_B и, следовательно, почти ортонормированную в H_A. Однако для достижения последней цели нет необходимости добиваться равенства $|\varphi_n|_B=1$; достаточно, чтобы было $a\leqslant|\varphi_n|_B\leqslant b$, где a и b — положительные постоянные. Тогда система $\{\varphi_n\}$, ортогональная в H_B, станет почти ортонормированной в H_B, а следовательно, и в H_A. Это простое замечание позволяет упростить нормирующие множители, а с ними и строение координатных функций, что в свою очередь уменьшает объем вычислительной работы без ущерба для качества результата. Так, в системе (30.5) можно заменить нормирующий множитель $\frac{\sqrt{2}}{n\pi}$ через $\frac{1}{n}$, что приведет нас к несколько более простой почти ортонормированной системе

$$\varphi_n(x)=\frac{1}{n}\sin n\pi x \qquad (n=1,\ 2,\ \ldots). \tag{30.5$'$}$$

Аналогично можно заменить системы (30.6), (30.11), (30.13), (30.16) следующими, более простыми системами, которые также почти ортонормированы в соответствующих пространствах H_A:

$$Q_n(x) = \sqrt{n}\int_0^x P_n(2t-1)\,dt \qquad (n=1,\ 2,\ \ldots), \tag{30.6'}$$

$$\varphi_n(x) = \frac{1}{2n-1}\sin\frac{2n-1}{2}\pi x \qquad (n=1,\ 2,\ \ldots), \tag{30.11'}$$

$$\varphi_0(x)=1, \qquad \varphi_n(x) = \frac{1}{n}\cos n\pi x \qquad (n=1,\ 2,\ \ldots), \tag{30.13'}$$

$$1,\ x,\ \sin\pi x,\ \frac{1}{2}\sin 2\pi x,\ \ldots,\ \frac{1}{n}\sin n\pi x,\ \ldots \tag{30.16'}$$

В формуле (30.6′) можно также заменить $\sqrt{n}$ его целой частью.

Заметим еще следующее: мы удовлетворим неравенству

$$a \leqslant |\varphi_n|_B \leqslant b,$$

если будем нормировать элементы φ_n в метрике пространства H_A.

§ 31. Случай вырождающегося уравнения

Ограничимся некоторыми простейшими задачами; рассмотрение других краевых условий, как нам кажется, не будет связано с серьезными затруднениями.

1. Простой и важный пример вырождения обыкновенного дифференциального уравнения доставляет нам § 26. С этого примера мы и начнем.

Итак, рассмотрим уравнение

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)\,u = f(x), \qquad 0 < x < 1, \tag{31.1}$$

где

$$p(x) = \tilde{p}(x)\,x^{\alpha}, \qquad 0 < \alpha < 2, \qquad \tilde{p}(x) \geqslant p_0 = \text{const} > 0; \tag{31.2}$$

мы сохраним здесь и другие предположения § 26 о свойствах функций $\tilde{p}(x)$ и $q(x)$.

Уравнение (31.1) будем решать при следующих краевых условиях:

$$u(1) = 0, \tag{31.3}$$

и если $0 < \alpha < 1$, то

$$u(0) = 0. \tag{31.4}$$

Из результатов § 26 вытекает, что в качестве координатных целесообразно выбрать собственные функции оператора B, определяемого дифференциальным выражением

$$Bu = -\frac{d}{dx}\left(x^{\alpha}\frac{du}{dx}\right) \tag{31.5}$$

при тех же краевых условиях (31.3), (31.4). Как мы видели, эти функции суть (формула (26.11))

$$\varphi_n(x) = c_n x^{\frac{1-\alpha}{2}} J_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right) \qquad (n = 1,\ 2,\ \ldots). \tag{31.6}$$

Постоянную c_n выберем так, чтобы $a \leqslant |\varphi_n|_B < b$, где a и b — положительные постоянные[1]); тогда будут иметь место утверждения а) — в) § 29.

Проще всего положить

$$c_n = \left\{\int_0^1 x^{\alpha}\left[\frac{d}{dx}\left(x^{\frac{1-\alpha}{2}} J_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right)\right)\right]^2 dx\right\}^{-\frac{1}{2}} \qquad (n = 1,\ 2,\ \ldots), \tag{31.7}$$

тогда $|\varphi_n|_B = 1$ и система (31.6) ортонормирована в H_B. Можно избежать вычисления интеграла, входящего в формулу (31.7), и заменить множитель (31.7) более простым; для этого мы предварительно выясним асимптотику интеграла

$$K = \int_0^1 x^{\alpha}\left[\frac{d}{dx}\left(x^{\frac{1-\alpha}{2}} J_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right)\right)\right]^2 dx$$

при $n \to \infty$.

Выполнив дифференцирование под знаком интеграла K, представим его в виде

$$K = K_1 + K_2 + K_3,$$

где

$$K_1 = \frac{(2-\alpha)^2}{4}\gamma_{\nu,\,n}^2 \int_0^1 x^{1-\alpha}\left[J'_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right)\right]^2 dx,$$

$$K_2 = \frac{(2-\alpha)(1-\alpha)}{2}\gamma_{\nu,\,n} \int_0^1 x^{-\frac{\alpha}{2}} J'_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right) J_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right) dx,$$

$$K_3 = \frac{(1-\alpha)^2}{4} \int_0^1 x^{-1}\left[J_\nu\left(\gamma_{\nu,\,n} x^{\frac{2-\alpha}{2}}\right)\right]^2 dx.$$

Найдем асимптотическое выражение интеграла K_1 при больших n; будем ниже писать γ вместо $\gamma_{\nu,\,n}$. Положив $\gamma x^{\frac{2-\alpha}{2}} = z$, получим

[1]) См. замечание в конце предшествующего параграфа.

(ε — произвольное число, $0 < \varepsilon < 1$):

$$K_1 = \frac{2-\alpha}{2} \int_0^{\gamma} z\left[J'_\nu(z)\right] dz = \frac{2-\alpha}{2} \int_0^{\gamma^{\frac{1-\varepsilon}{2}}} z\left[J'_\nu(z)\right]^2 dz + \\ + \frac{2-\alpha}{2} \int_{\gamma^{\frac{1-\varepsilon}{2}}}^{\gamma} z\left[J'_\nu(z)\right]^2 dz = K'_1 + K''_1.$$

Воспользуемся следующими формулами [1]):

$$2J'_\nu(x) = J_{\nu-1}(x) - J_{\nu+1}(x), \tag{31.8}$$

$$J_\nu(x) = \frac{1}{\pi} \int_0^{\pi} \cos(\nu\theta - x\sin\theta)\, d\theta - \frac{\sin\nu\pi}{\pi} \int_0^{\infty} e^{-\nu t - x \operatorname{sh} t}\, dt, \qquad x > 0, \tag{31.9}$$

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) + O\left(x^{-\frac{3}{2}}\right), \qquad x > 0. \tag{31.10}$$

Из формул (31.8) и (31.9) вытекают оценки

$$|J_\nu(x)| \leqslant \sigma, \qquad |J'_\nu(x)| \leqslant \sigma, \qquad \sigma = 1 + \frac{1}{\pi\nu}; \tag{31.11}$$

формулы (31.8) и (31.10) дают асимптотическое выражение производной $J'_\nu(x)$ при больших положительных x:

$$J'_\nu(x) = \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{(\nu-1)\pi}{2} - \frac{\pi}{4}\right) + O\left(x^{-\frac{3}{2}}\right). \tag{31.12}$$

По неравенству (31.11),

$$K'_1 \leqslant \frac{2-\alpha}{2} \sigma^2 \int_0^{\gamma^{\frac{1-\varepsilon}{2}}} z\, dz = o(\gamma).$$

Заменив $J'_\nu(z)$ по формуле (31.12), получим:

$$K''_1 = \frac{2-\alpha}{\alpha} \frac{2}{\pi} \int_{\gamma^{\frac{1-\varepsilon}{2}}}^{\gamma} \left[\cos^2\left(x - \frac{(\nu-1)\pi}{2} - \frac{\pi}{4}\right) + O(x^{-1})\right] dx = \\ = \frac{2-\alpha}{2\pi}\gamma + o(\gamma)$$

[1]) См. Г. Н. Ватсон [1], стр. 56, 195, 218.

и окончательно

$$K_1 = \frac{2-\alpha}{2\pi}\gamma + o(\gamma). \tag{31.13}$$

Если $\alpha = 1$, то $K_2 = K_3 = 0$, поэтому последующие оценки достаточно провести при $\alpha \neq 1$ и, следовательно, $\nu > 0$.

Уже использованная подстановка $z = \gamma x^{\frac{2-\alpha}{2}}$ дает:

$$K_3 = \frac{(1-\alpha)^2}{2(2-\alpha)} \int_0^\gamma z^{-1} [J_\nu(z)]^2 \, dz =$$

$$= \frac{(1-\alpha)^2}{2(2-\alpha)} \left\{ \int_0^1 z^{-1} [J_\nu(z)]^2 \, dz + \int_1^\gamma z^{-1} [J_\nu(z)]^2 \, dz \right\}.$$

Первый интеграл справа — постоянная, а второй, в силу неравенства (31.11), оценивается так:

$$\int_1^\gamma z^{-1} [J_\nu(z)]^2 \, dz \leqslant \sigma^2 \ln \gamma.$$

Теперь

$$K_3 = o(\gamma). \tag{31.14}$$

Интеграл K_2 оценивается совсем просто:

$$|K_2| \leqslant 2K_1^{\frac{1}{2}} K_3^{\frac{1}{2}} = o(\gamma). \tag{31.15}$$

Из формул (31.13) — (31.15) вытекает нужный нам результат:

$$K = \frac{2-\alpha}{2\pi}\gamma_{\nu,n} + o(\gamma_{\nu,n}). \tag{31.16}$$

Воспользовавшись асимптотической формулой для корней функций Бесселя[1]), из которой следует, что

$$\gamma_{\nu,n} = \pi n + o(n),$$

получим окончательно:

$$K = \frac{2-\alpha}{2} n + o(n).$$

Теперь ясно, что в формуле (31.6) можно положить $c_n = \sqrt{n}$ или $c_n = [\sqrt{n}]$, где на этот раз квадратные скобки обозначают целую часть записанного в скобках числа. Тогда система (31.6) будет почти ортонормированной в H_A и будут верны утверждения а) — в) § 29.

---

[1]) Г. Н. Ватсон [1], стр. 558.

2. Рассмотрим теперь случай, когда уравнение вырождается на обоих концах промежутка. В пространстве $L_2(-1, 1)$ рассмотрим оператор

$$Au = -\frac{d}{dx}\left[(1-x)^{\alpha}(1+x)^{\beta}\tilde{p}(x)\frac{du}{dx}\right] + q(x)u = f(x). \tag{31.17}$$

Для простоты допустим, что $\tilde{p}(x)$, $\tilde{p}'(x)$ и $q(x)$ непрерывны при $-1 \leqslant x \leqslant 1$ и что $\tilde{p}(x) \geqslant p_0$, $q(x) \geqslant 0$, где p_0 — положительная постоянная. Показатели α и β будем считать такими, что $0 \leqslant \alpha < 2$, $0 \leqslant \beta < 1$, но исключим из рассмотрения тривиальный случай $\alpha = \beta = 0$, когда вырождения нет. Заметим, что при $\alpha > 0$, $\beta = 0$ получается случай п. 1 настоящего параграфа.

Оператор (31.17) зададим на функциях $u(x)$, подчиненных следующим условиям:

1) $u(x)$ непрерывна, а $(1-x)^{\alpha}(1+x)^{\beta}u'(x)$ абсолютно непрерывна на отрезке $-1 \leqslant x \leqslant 1$;

2) $u(1) = 0$; (31.18)

3) если $\alpha < 1$, то

$$u(0) = 0. \tag{31.19}$$

При таком задании оператор A — положительно определенный, причем

$$|u|_A^2 = \int_{-1}^{1} \left\{(1-x)^{\alpha}(1+x)^{\beta}\tilde{p}(x)|u'(x)|^2 + q(x)|u(x)|^2\right\}dx. \tag{31.20}$$

Наряду с A рассмотрим положительно определенный оператор B, заданный на том же множестве функций и действующий по формуле

$$Bu = -\frac{d}{dx}\left[(1-x)^{\alpha}(1+x)^{\beta}\frac{du}{dx}\right], \tag{31.21}$$

так что

$$|u|_B^2 = \int_{-1}^{1} (1-x)^{\alpha}(1+x)^{\beta}|u'(x)|^2\,dx. \tag{31.22}$$

Повторяя рассуждения § 26, можно доказать, что операторы A и B — самосопряженные и, следовательно, сходные. Нетрудно также доказать, что оператор B — родственный оператору A. Однако построение собственных функций оператора B затруднительно, и мы ограничимся тем, что укажем систему, ортонормированную в метрике (31.22) и, следовательно, почти ортонормированную в метрике (31.20). Это — система полиномов

$$q_n^{(\alpha,\beta)}(x) = -\int_x^1 p_n^{(\alpha,\beta)}(t)\,dt, \tag{31.23}$$

где $p_n^{(\alpha, \beta)}(x)$ — нормированные полиномы Якоби[1]). Выбор функций $q_n^{(\alpha, \beta)}(x)$ в качестве координатных обеспечивает справедливость утверждений а) и б) § 29.

§ 32. Обыкновенные дифференциальные уравнения четвертого порядка

Пусть дано уравнение

$$\frac{d^2}{dx^2}\left(p_2(x)\frac{d^2u}{dx^2}\right)-\frac{d}{dx}\left(p_1(x)\frac{du}{dx}\right)+p_0u=f(x), \tag{32.1}$$

коэффициенты p_0, p_1, p_2 подчинены условиям § 27.

Уравнение (32.1) рассмотрим при нескольких простых типах краевых условий.

1. Начнем с условий

$$u(0)=u(1)=u''(0)=u''(1)=0. \tag{32.2}$$

Из результатов § 27 вытекает, что в данном случае целесообразно взять в качестве координатных функций $\varphi_n(x)=c_n \sin n\pi x$ $(n=1, 2, \ldots)$, причем коэффициент c_n можно определить, например, из условия $|\varphi_n|_B=1$, где B — оператор (27.4). Это дает значение $c_n=\frac{\sqrt{2}}{n^2\pi^2}$.

Постоянный множитель $\frac{\sqrt{2}}{\pi^2}$ можно отбросить и мы получаем

$$\varphi_n(x)=\frac{\sin n\pi x}{n^2} \qquad (n=1, 2, \ldots). \tag{32.3}$$

При таком выборе координатной системы имеют место утверждения а) — в) § 29.

2. Рассмотрим теперь уравнение (32.1) при краевых условиях

$$u(0)=u(1)=u'(0)=u'(1)=0. \tag{32.4}$$

Оператор (32.1), (32.4) — сходный и, тем более, полусходный с оператором $B_1=\frac{d^4}{dx^4}$, заданном на функциях, удовлетворяющих условиям (32.4). Энергетическая метрика оператора B_1 определяется формулой

$$|u|_{B_1}^2=\int_0^1 |u''(x)|^2\,dx, \tag{32.5}$$

---

[1]) См., например, Д. Джексон [1].

причем входящие в пространство H_{B_1} функции удовлетворяют условиям (32.4). Легко убедиться, что система

$$\varphi_n(x)=\sqrt{2n+1}\int_0^x dt\int_0^t P_n(2\tau-1)\,d\tau=$$
$$=\sqrt{2n+1}\int_0^x (x-t)P_n(2t-1)\,dt \qquad (n=2,\ 3,\ \ldots), \tag{32.6}$$

где P_n — полиномы Лежандра, ортонормирована и полна в H_{B_1}; если ее взять за координатную, то справедливы утверждения а), б) § 29. Множитель $\sqrt{2n+1}$ можно заменить на $\sqrt{n}$ или на $[\sqrt{n}]$.

3. Поставим теперь следующие краевые условия

$$u(0)=u'(0)=u(1)=u''(1)=0. \tag{32.7}$$

Сходным и, следовательно, полусходным с оператором задачи (32.1), (32.7) является оператор $B_2=\frac{d^4}{dx^4}$, заданный на функциях, удовлетворяющих условиям (32.7). Энергетическая метрика этого оператора определяется той же формулой (32.5), но функции из H_{B_2} удовлетворяют только условиям $u(0)=u'(0)=u(1)=0$, условие $u''(0)=0$ — естественное[1]). Система (32.6) ортонормирована, но не полна в H_{B_2}; выясним, чем ее следует пополнить. Пусть функция $g(x)\in H_{B_2}$ ортогональна (в метрике этого пространства) к системе (32.6). Это значит, что

$$\int_0^1 g''(x)P_n(x)\,dx=0 \qquad (n=2,\ 3,\ \ldots). \tag{32.8}$$

Отсюда следует, что $g''(x)$ — полином первой степени и, следовательно, $g(x)$ — полином третьей степени. Но $g\in H_{B_2}$, поэтому $g(0)=g'(0)=g(1)=0$, и $g(x)$ только постоянным множителем отличается от функции $x^2(1-x)$. Таким образом, полной и почти ортонормированной в энергетической метрике оператора (32.1), (32.7) является система функций

$$\varphi_1(x)=x^2(1-x), \qquad \varphi_n(x)=[\sqrt{n}]\int_0^x (x-t)P_n(2t-1)\,dt, \quad n>1. \tag{32.9}$$

4. Рассмотрим теперь случай, когда оба условия на конце $x=1$ естественные. Пусть краевые условия имеют вид

$$u(0)=u'(0)=0, \qquad u''(1)=0, \qquad (p_2(x)\,u''(x))'\,|_{x=1}=0; \tag{32.10}$$

[1]) См. ВМ, § 17; ПМ, § 35.

в задаче об изгибе балки (ВМ, § 21) эти условия означают, что левый конец балки жестко закреплен, а правый конец свободен. Полусходным с оператором задачи (32.1), (32.10) является оператор $B_3 = \frac{d^4}{dx^4}$, заданный на функциях, удовлетворяющих первым трем условиям (32.10) и еще условию $u'''(1) = 0$. Функции из H_{B_3} удовлетворяют условиям $u(0) = u'(0) = 0$; метрика в этом пространстве определяется формулой (32.5). Система (32.6) ортонормирована, но не полна в H_{B_3}. Рассуждая, как в п. 3, найдем опять, что функции, ортогональные в H_{B_3} к системе (32.6), суть полиномы третьей степени; таких полиномов, удовлетворяющих условиям $u(0) = u'(0) = 0$ и линейно независимых между собой, — два: x^2 и x^3; полной и почти ортонормированной в метрике оператора (32.1), (32.10) является система функций

$$\varphi_0(x) = x^2, \qquad \varphi_1(x) = x^3,$$

$$\varphi_n(x) = [n] \int_0^x (x - t) P_n(2t - 1)\, dt \qquad (n = 2,\ 3,\ \ldots). \tag{32.11}$$

5. По образцу пп. 2 – 4 могут быть построены почти ортонормированные координатные системы и для некоторых других типов краевых условий, а также для невырождающихся обыкновенных дифференциальных уравнений более высоких порядков.

§ 33. Двумерные эллиптические уравнения; первая краевая задача

В этом параграфе мы всюду полагаем $H = L_2(\Omega)$, где Ω — конечная плоская область; ее границу мы, как обычно, обозначим через S.

Рассмотрим невырождающееся эллиптическое уравнение

$$-\sum_{j,\ k=1}^{2} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_j}\right) + Cu = f(x_1,\ x_2), \quad x_1 = x, \quad x_2 = y, \tag{33.1}$$

при краевом условии

$$u|_S = 0. \tag{33.2}$$

Через μ_1 обозначим наименьшее собственное значение оператора

$$-\sum_{j,\ k=1}^{2} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_j}\right) \tag{33.3}$$

при краевом условии (33.2) и будем считать, что $C > -\mu_1$. Примем также, что A_{jk}, $\frac{\partial A_{jk}}{\partial x_l}$, C непрерывны в замкнутой области $\overline{\Omega} = \Omega \cup S$.

Допустим, что контур S достаточно гладкий и что нам удалось найти преобразование

$$x' = \varphi(x, y), \qquad y' = \psi(x, y), \tag{33.4}$$

которое взаимно однозначно отображает замкнутую область $\overline{\Omega}$ на круг $x'^2 + y'^2 \leqslant 1$, причем в области $\overline{\Omega}$ функции φ и ψ дважды непрерывно дифференцируемы и якобиан

$$J = \frac{D(x, y)}{D(x', y')}$$

положительно ограничен сверху и снизу. Преобразование (33.4) переводит задачу (33.1) — (33.2) в следующую:

$$-\sum_{j,\,k=1}^{2} \frac{\partial}{\partial x'_j}\left(A'_{jk}\frac{\partial u}{\partial x'_k}\right) + CJu = fJ, \qquad x'_1 = x', \qquad x'_2 = y', \tag{33.5}$$

$$u|_\Gamma = 0. \tag{33.6}$$

Здесь Γ — окружность $x'^2 + y'^2 = 1$ и

$$A'_{jk} = J \sum_{r,\,s=1}^{2} A_{rs} \frac{\partial x'_j}{\partial x_r} \frac{\partial x'_k}{\partial x_s}.$$

Обозначим через A оператор, определяемый левой частью уравнения (33.5) и краевым условием (33.6). Как это следует из результатов § 28, родственным к A является оператор

$$B = -\Delta = -\frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial y'^2}$$

при краевом условии (33.6). Собственные функции оператора B суть

$$\varphi_{k,\,n}(x', y') = c_{k,\,n} J_k(\gamma_{k,\,n} r) \begin{matrix}\cos\\ \sin\end{matrix} k\theta \qquad (k, n = 0, 1, 2, \ldots), \tag{33.7}$$

где $x' = r\cos\theta$, $y' = r\sin\theta$ и коэффициент $c_{k,\,n}$ таков, что $|\varphi_{k,\,n}|_B = 1$. Взятые в качестве координатных при решении задачи (33.5) — (33.6), эти функции обеспечивают справедливость утверждений а) — в) § 29; при этом имеет место утверждение § 28 о сходимости в метрике $L_2(\Omega)$ вторых производных от приближенного решения по Ритцу к соответствующим вторым производным точного решения. В силу теорем вложения С. Л. Соболева отсюда следует, что первые производные сходятся в $L_p(\Omega)$ при любом $p > 1$, а сами приближенные решения сходятся равномерно в Ω.

Пусть область Ω заключена между прямыми $y = \pm b$, которые касаются контура S, и пусть любая расположенная между ними прямая $y=\text{const}$ пересекает S только в двух точках, абсциссы которых обозначим через $\alpha(y)$ и $\beta(y)$. Преобразование (33.4) дается формулами

$$x'=\mu(y)x+\nu(y), \quad y'=\frac{1}{b}y, \tag{33.8}$$

где

$$\mu(y)=\frac{2\sqrt{b^2-y^2}}{b[\beta(y)-\alpha(y)]}, \qquad \nu(y)=\frac{\beta(y)+\alpha(y)}{\beta(y)-\alpha(y)}\frac{\sqrt{b^2-y^2}}{b}. \tag{33.9}$$

Это преобразование пригодно для наших целей, если функции $\mu(y)$ и $\nu(y)$ имеют непрерывные вторые производные на отрезке $|y|\leqslant b$.

Пример 33.1. Пусть S — эллипс

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

и Ω — его внутренность. В данном случае

$$\alpha(y)=-\frac{a}{b}\sqrt{b^2-y^2}, \qquad \beta(y)=\frac{a}{b}\sqrt{b^2-y^2}$$

и, следовательно,

$$\mu(y)=\frac{1}{a}, \qquad \nu(y)=0,$$

и преобразование (33.8) имеет вид

$$x'=\frac{x}{a}, \qquad y'=\frac{y}{b}.$$

Может случиться, что преобразование (33.4) дает взаимно однозначное и достаточно гладкое отображение области Ω на круговой сектор с углом раствора $\beta\leqslant\pi$ или на прямоугольник. Упомянутый выше оператор B (в формуле (33.6) под Γ будем в данном случае понимать контур сектора или соответственно прямоугольника) является во всяком случае полусходным с оператором A; собственные функции оператора B суть

$$\varphi_{k,n}(x', y')=c_{k,n}J_{\frac{k\pi}{\beta}}\left(\gamma_{\frac{k\pi}{\beta},n}r\right)\sin\frac{k\pi\theta}{\beta} \qquad (k, n=1, 2, \ldots), \tag{33.10}$$

в случае сектора $0<r<1$, $0<\theta<\beta$; если $\beta=\pi$, так что Ω допускает достаточно гладкое отображение на полукруг, то

$$\varphi_{k,n}(x', y')=c_{k,n}J_k(\gamma_{k,n}r)\sin k\theta \qquad (k, n=1, 2, \ldots). \tag{33.10'}$$

Далее,

$$\varphi_{k,n}(x', y')=\frac{2}{\pi}\left(\frac{k^2}{a^2}+\frac{m^2}{b^2}\right)^{-\frac{1}{2}}\sin\frac{k\pi x'}{a}\sin\frac{m\pi y'}{b} \qquad (k, n=1, 2, \ldots), \tag{33.11}$$

в случае прямоугольника $0<x'<a$, $0<y'<b$. Коэффициент $c_{k,n}$

в формуле (33.10) определяется из условия

$$|\varphi_{k,n}|_B^2 = \int_0^1 r\,dr \int_0^\beta \left[\left(\frac{\partial\varphi_{k,n}}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial\varphi_{k,n}}{\partial\theta}\right)^2\right] d\theta =$$

$$= \frac{\beta}{2}\int_0^1 \left[\gamma^2_{\frac{k\pi}{\beta},n} J'^2_{\frac{k\pi}{\beta}}\left(\gamma_{\frac{k\pi}{\beta},n} r\right) + \frac{k^2\pi^2}{\beta^2} J^2_{\frac{k\pi}{\beta}}\left(\gamma_{\frac{k\pi}{\beta},n} r\right)\right] r\,dr = 1.$$

Если функции (33.10), соответственно (33.11), использовать в качестве координатных при решении задачи (33.5) — (33.6), то будут иметь место утверждения а) и б) § 29.

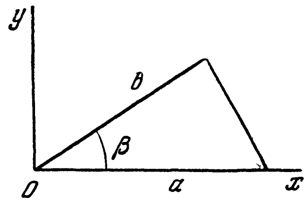

Рис. 2.

Пример 33.2. Пусть Ω — прямолинейный треугольник с острым углом β; размеры и положение треугольника показаны на рис. 2. Отображение этого треугольника на круговой сектор $0 < r < 1$, $0 < \vartheta < \beta$ дается формулами

$$x' = \frac{1}{b}\left\{y \operatorname{ctg}\beta + \frac{b\sin\beta}{a}(x - y\operatorname{ctg}\beta)\frac{\sqrt{b^2-y^2} - y\operatorname{ctg}\beta}{b\sin\beta - y}\right\}, \quad y' = \frac{1}{b}y; \tag{33.12}$$

легко видеть, что это преобразование бесконечно дифференцируемо и что его якобиан положительно ограничен сверху и снизу. Действительно, мы докажем утверждение о дифференцируемости, если установим, что дробь

$$\varphi(y) = \frac{\sqrt{b^2-y^2} - y\operatorname{ctg}\beta}{b\sin\beta - y}$$

бесконечно дифференцируема по y в промежутке $0 \leqslant y \leqslant b\sin\beta$. Положим $y = b\sin\lambda$, $0 \leqslant \lambda \leqslant \beta$. Тогда

$$\varphi(y) = \frac{1}{\sin\beta}\,\frac{\cos\dfrac{\beta-\lambda}{2}}{\cos\dfrac{\beta+\lambda}{2}}$$

и очевидно, что при $0 \leqslant \lambda \leqslant \beta$ эта функция имеет все производные по λ, а следовательно, и по y. Далее,

$$\frac{D(x', y')}{D(x, y)} = \frac{\sin\beta}{ab}\,\frac{\sqrt{b^2-y^2} - y\operatorname{ctg}\beta}{b\sin\beta - y} = \frac{1}{ab}\,\frac{\cos\dfrac{\beta-\lambda}{2}}{\cos\dfrac{\beta+\lambda}{2}}.$$

Отсюда

$$\frac{1}{ab} \leqslant \frac{D(x', y')}{D(x, y)} \leqslant \frac{1}{ab\cos\beta}.$$

§ 34. Двумерные эллиптические уравнения; задачи с естественными краевыми условиями

Дифференциальное уравнение (33.1) рассмотрим теперь при краевом условии

$$\left[\sum_{j,\ k=1}^{2} A_{jk}\frac{\partial u}{\partial x_k}\cos(\nu,\ x_k)+\sigma u\right]_S=0. \tag{34.1}$$

Здесь ν — внешняя нормаль к контуру S, σ — ограниченная измеримая неотрицательная функция точки S; функции σ и C считаем такими, что оператор, порожденный дифференциальным выражением (33.1) при краевом условии (34.1), — положительно определенный. Упомянутый только что оператор (точнее, его самосопряженное расширение по Фридрихсу) будем обозначать через A. В некоторых, впрочем, довольно редких случаях удается построить сравнительно простой оператор, родственный оператору A. Пусть, например, уравнение (33.1) имеет вид

$$-\Delta u+C(x,\ y)\,u=f(x,\ y),$$

а в условии (34.1) $\sigma\equiv 0$, так что это условие сводится к следующему:

$$\left.\frac{\partial u}{\partial \nu}\right|_S=0. \tag{34.1$'$}$$

Из результатов § 24, п. 1 следует, что самосопряженный оператор B, определяемый дифференциальным выражением $Bu=-\Delta u+u$ и краевым условием (34.1′), родственен оператору A, если только контур S — достаточно гладкий; собственные функции оператора B легко строятся, если, например, S есть единичная окружность, именно,

$$\varphi_{k,\ n}=c_{k,\ n}J_k(\tilde{\gamma}_{k,\ n}r)\begin{matrix}\cos\\ \sin\end{matrix}\,k\theta \tag{34.2}$$

$$(k=0,\ 1,\ 2,\ \ldots;\quad n=1,\ 2,\ \ldots).$$

Здесь $\tilde{\gamma}_{k,\ n}$ означает n-й положительный корень производной $J_k'(x)$. Как обычно, коэффициент $c_{k,\ n}$ определяется из условия $|\varphi_{k,\ n}|_B=1$ или из более общего условия, чтобы величина $|\varphi_{k,\ n}|_B$ была положительно ограничена сверху и снизу.

Если функции (34.2) взяты за координатные, то имеют место утверждения а) — в) § 29.

В общем случае мы допустим, что преобразованием вида (33.4) можно отобразить область Ω на круг, круговой сектор или прямоугольник; мы сохраняем условия § 33 относительно преобразования (33.4).

Уравнение (33.1) преобразуется при этом в уравнение (33.5), а краевое условие (34.1) приводится к виду

$$\left[\sum_{j,\ k=1}^{2} A'_{jk}\frac{\partial u}{\partial x'_k}\cos(\nu',\ x_j)+\sigma' u\right]_\Gamma=0. \tag{34.3}$$

Здесь Γ — контур преобразованной области, ν' — внешняя нормаль к Γ,

$$\sigma'=\sigma\sqrt{E\cos^2(\nu',\ y')-2F\cos(\nu',\ x')\cos(\nu',\ y')+G\cos^2(\nu',\ x')},$$

где

$$E=\left(\frac{\partial x}{\partial x'}\right)^2+\left(\frac{\partial x}{\partial y'}\right)^2,\quad F=\frac{\partial x}{\partial x'}\frac{\partial x}{\partial y'}+\frac{\partial y}{\partial x'}\frac{\partial y}{\partial y'},\quad G=\left(\frac{\partial y}{\partial x'}\right)^2+\left(\frac{\partial y}{\partial y'}\right)^2.$$

Оператор B, где

$$Bu=-\frac{\partial^2 u}{\partial x'^2}-\frac{\partial^2 u}{\partial y'^2}+u,\qquad \left.\frac{\partial u}{\partial \nu'}\right|_\Gamma=0,$$

— полусходный с оператором A', порожденным уравнением (33.5) и краевым условием (34.3). Для круга собственные функции оператора B определяются формулой (34.2), для сектора $0<r<1$, $0<\vartheta<\beta$ имеем:

$$\varphi_{k,\ n}=c_{k,\ n}J_{\frac{k\pi}{\beta}}\left(\gamma_{\frac{k\pi}{\beta},\ n}r\right)\cos\frac{k\pi\vartheta}{\beta}\qquad (k=0,\ 1,\ 2,\ \ldots;\ n=1,\ 2,\ \ldots), \tag{34.4}$$

в частном случае полукруга ($\beta=\pi$)

$$\varphi_{k,\ n}=c_{k,\ n}J_k(\tilde{\gamma}_{k,\ n}r)\cos k\vartheta\quad (k=0,\ 1,\ 2,\ \ldots;\ n=1,\ 2,\ \ldots). \tag{34.5}$$

Наконец, для прямоугольника $0<x'<a$, $0<y'<b$

$$\varphi_{k,\ n}=c_{k,\ n}\cos\frac{k\pi x'}{a}\cos\frac{n\pi y'}{b}\qquad (k,\ n=0,\ 1,\ 2,\ \ldots). \tag{34.6}$$

Коэффициент $c_{k,\ n}$ выбирается из условия, которое неоднократно упоминалось выше, что $|\varphi_{k,n}|_B$ положительно ограничена сверху и снизу; в частности, можно потребовать, чтобы $|\varphi_{k,\ n}|_B=1$. Для функций (34.6) можно положить, например,

$$c_{k,\ n}=\frac{1}{\pi}\left(\frac{k^2}{a^2}+\frac{n^2}{b^2}\right)^{-\frac{1}{2}}$$

или проще:

$$c_{k,\ n}=\frac{1}{[\sqrt{k^2+n^2}]}.$$

Выбрав в качестве координатной одну из систем (34.4) — (34.6) в соответствии с характером области, мы обеспечим справедливость утверждений а) и б) § 29.

Некоторые дополнительные подробности читатель найдет в книге Х. Л. Смолицкого и автора [1].

§ 35. Трехмерные задачи

Мы не станем изучать уравнения с произвольным числом независимых переменных и ограничимся случаем $m=3$ как наиболее интересным для приложений.

Будем рассматривать невырождающееся эллиптическое уравнение

$$-\sum_{j,\,k=1}^{3} \frac{\partial}{\partial x_j}\left(A_{jk}(x)\frac{\partial u}{\partial x_j}\right)+C(x)u=f(x), \quad x=(x_1,\, x_2,\, x_3)=(x,\, y,\, z), \tag{35.1}$$

при краевом условии одного из видов

$$u|_S=0, \tag{35.2}$$

$$\left[\sum_{j,\,k=1}^{3} A_{jk}\frac{\partial u}{\partial x_j}\cos(\nu,\, x_k)+\sigma u\right]_S=0; \tag{35.3}$$

ν — внешняя нормаль к S. Будем считать, что оператор A, порожденный левой частью уравнения (35.1) и краевым условием (35.2) или (35.3), положительно определенный в $L_2(\Omega)$, где Ω — конечная область, ограниченная поверхностью S.

Ограничимся теми случаями, когда область Ω допускает достаточно гладкое отображение с положительным якобианом на одну из следующих областей: 1) шар, 2) прямоугольный параллелепипед, 3) прямой круговой цилиндр, 4) часть такого цилиндра, вырезанную из него двумя полуплоскостями, проходящими через его ось, 5) сферический конус — так мы называем область, ограниченную прямым круговым конусом и частью сферы, центр которой совпадает с вершиной конуса. Более того, будем считать, что такое преобразование уже выполнено, так что Ω просто совпадает с одной из перечисленных областей.

Наряду с определенным выше оператором A рассмотрим оператор B, который действует по формуле

$$Bu=-\Delta u+u\,^{1)} \tag{35.4}$$

[1]) Слагаемое u в формуле (35.4) можно опустить во всех случаях, кроме следующего: ставится краевое условие (35.3), в котором $\sigma(x)\equiv 0$.

и задан на множестве функций, удовлетворяющих условию (35.2), если уравнение (35.1) решается при этом же условии, и условию

$$\left.\frac{\partial u}{\partial \nu}\right|_S = 0, \tag{35.5}$$

если уравнение (35.1) решается при условии (35.3). Будем считать, что оператор B расширен по Фридрихсу до самосопряженного.

Если $A_{jk}(x) = \delta_{jk}$, так что уравнение (35.1) имеет вид $-\Delta u + C(x)u = f(x)$, если, далее, область Ω есть шар и поставлено краевое условие (35.2), то оператор B родствен оператору A, — это вытекает из результатов § 28. Выбрав собственные функции оператора B в качестве координатных, мы обеспечим справедливость утверждений а) — в) § 29. Во всех остальных случаях, которые мы изучаем в настоящем параграфе, операторы A и B по крайней мере полусходные; тот же выбор координатной системы обеспечивает справедливость утверждений а) и б) § 29.

Ниже мы приводим собственные функции оператора B[1]) (мы будем обозначать их через φ_{klm}) для перечисленных ранее областей как в случае условия (35.2) («задача Дирихле»), так и в случае условия (35.5) («задача Неймана»).

1. Шар радиуса единица. Введем сферические координаты r, θ, φ, с центром в центре шара. Разделением переменных легко находим: для задачи Дирихле

$$\varphi_{klm} = c_{klm} \frac{1}{\sqrt{r}} J_{k+\frac{1}{2}}\left(\gamma_{k+\frac{1}{2},\, l}\, r\right) P_k^{(m)}(\cos\theta) \begin{matrix}\cos m\varphi \\ \sin m\varphi\end{matrix} \tag{35.6}$$

$$(k = 0,\ 1,\ 2,\ \ldots;\quad l = 1,\ 2,\ \ldots;\quad m = 0,\ 1,\ 2,\ \ldots,\ k),$$

для задачи Неймана

$$\varphi_{klm} = c_{klm} \frac{1}{\sqrt{r}} J_{k+\frac{1}{2}}\left(\overset{*}{\gamma}_{k+\frac{1}{2},\, l}\, r\right) P_k^{(m)}(\cos\theta) \begin{matrix}\cos m\varphi \\ \sin m\varphi\end{matrix} \tag{35.7}$$

$$(k = 0,\ 1,\ 2,\ \ldots;\quad l = 1,\ 2,\ \ldots;\quad m = 0,\ 1,\ 2,\ \ldots,\ k).$$

В формулах (35.6) и (35.7) $P_k^{(m)}(\cos\theta)$ означает присоединенную функцию Лежандра с индексами k и m, $\gamma_{k+\frac{1}{2},\, l}$ есть l-й положительный корень функции Бесселя $J_{k+\frac{1}{2}}(x)$, а $\overset{*}{\gamma}_{k+\frac{1}{2},\, l}$ есть l-й положительный корень функции

$$J'_{k+\frac{1}{2}}(x) - \frac{1}{2x} J_{k+\frac{1}{2}}(x) = 0.$$

[1]) То есть собственные функции оператора Лапласа.

2. Параллелепипед $0 < x < a$, $0 < y < b$, $0 < z < c$. Разделением переменных в декартовых координатах сразу получаем для задачи Дирихле:

$$\varphi_{klm} = c_{klm} \sin \frac{k\pi x}{a} \sin \frac{l\pi y}{b} \sin \frac{m\pi z}{c} \qquad (k,\ l,\ m = 1,\ 2,\ \ldots). \tag{35.8}$$

Требование $|\varphi_{klm}|_B = 1$ дает:

$$c_{klm} = \frac{1}{\pi}\left(\frac{k^2}{a^2} + \frac{l^2}{b^2} + \frac{m^2}{c^2} + 1\right)^{-1}. \tag{35.9}$$

Если ограничиться требованием, чтобы величина $|\varphi_{klm}|_B$ была положительно ограничена сверху и снизу, то достаточно положить

$$c_{klm} = \frac{1}{[\sqrt{k^2 + l^2 + m^2}]}. \tag{35.10}$$

Для задачи Неймана

$$\varphi_{klm} = c_{klm} \cos \frac{k\pi x}{a} \cos \frac{l\pi y}{b} \cos \frac{m\pi z}{c} \qquad (k,\ l,\ m = 0,\ 1,\ 2,\ \ldots). \tag{35.11}$$

В качестве c_{klm} можно взять значения, даваемые формулой (35.9) или (35.10).

3. Цилиндр, определяемый в цилиндрических координатах r, φ, z неравенствами $0 < r < 1$, $0 \leqslant z \leqslant b$. Для задачи Дирихле

$$\varphi_{klm} = c_{klm} J_k(\gamma_{k,\,l} r) \sin \frac{m\pi z}{b} \begin{matrix}\cos\\ \sin\end{matrix} k\varphi \tag{35.12}$$
$$(k = 0,\ 1,\ 2,\ \ldots;\ l,\ m = 1,\ 2,\ 3,\ \ldots).$$

Для задачи Неймана

$$\varphi_{klm} = c_{klm} J_k(\tilde{\gamma}_{k,l} r) \cos \frac{m\pi z}{b} \begin{matrix}\cos\\ \sin\end{matrix} k\varphi \tag{35.13}$$
$$(k,\ m = 0,\ 1,\ 2,\ \ldots;\quad l = 1,\ 2,\ 3,\ \ldots)$$

через $\gamma_{k,\,l}$ и $\tilde{\gamma}_{k,\,l}$ обозначены l-е положительные корни функций $J_k(x)$ и $J_k'(x)$. Коэффициенты c_{klm} в формулах (35.12) и (35.13) можно определить из условия $|\varphi_{klm}|_A = 1$.

4. Часть кругового цилиндра, определяемая в цилиндрической системе координат неравенствами $0 < r < 1$, $0 < \varphi < \beta$ и $0 < z < b$. Для задачи Дирихле

$$\varphi_{klm} = c_{klm} J_{\frac{k\pi}{\beta}}\left(\gamma_{\frac{k\pi}{\beta},\,l} r\right) \sin \frac{k\pi\varphi}{\beta} \sin \frac{k\pi z}{b} \qquad (k,\ l,\ m = 1,\ 2,\ 3,\ \ldots). \tag{35.14}$$

Для задачи Неймана

$$\varphi_{klm} = c_{klm} J_{\frac{k\pi}{\beta}}\left(\tilde{\gamma}_{\frac{k\pi}{\beta},\,l} r\right) \cos \frac{k\pi\varphi}{\beta} \cos \frac{k\pi z}{b} \qquad (k,\ l,\ m = 0,\ 1,\ 2,\ \ldots). \tag{35.15}$$

Как и выше, коэффициенты c_{klm} можно определить из условия $|\varphi_{klm}|_A = 1$.

5. Сферический конус, определяемый в сферических координатах неравенствами $0 < r < 1$, $0 < \vartheta < \beta$.

Задача Дирихле. Разделение переменных в сферических координатах опять приводит к выражению (35.6) для собственных функций с той, однако, существенной разницей, что индекс k принимает, вообще, не целые значения.

Небесполезно заметить, что так как m есть целое неотрицательное число, то[1])

$$P_k^{(m)}(\cos\theta) = (-1)^m \sin^m\theta \frac{d^m P_k(\cos\theta)}{(d\cos\theta)^m},$$

где $P_k(x)$ — обобщенная функция Лежандра первого рода и индекса k. Функция (35.6) обращается в нуль при $r = 1$; необходимо еще, чтобы она обращалась в нуль при $\theta = \beta$, что дает уравнение для k:

$$P_k^{(m)}(\cos\beta) = 0. \tag{35.16}$$

Докажем, что при любом вещественном m уравнение (35.16) имеет счетное множество положительных корней. Если k удовлетворяет уравнению (35.16), то $P_k^{(m)}(x)$ есть собственная функция оператора

$$Pu = -\frac{d}{dx}\left[(1-x^2)\frac{du}{dx}\right] + \frac{m^2}{1-x^2}u, \qquad u(\cos\beta) = 0; \tag{35.17}$$

эта функция соответствует собственному числу $k(k+1)$. Нам достаточно доказать поэтому, что спектр оператора (35.17) дискретен. Интеграл

$$\int_{-1}^{\cos\beta} \frac{1+x}{1-x}\,dx$$

сходится, поэтому оператор

$$P_0(u) = -\frac{d}{dx}\left[(1-x^2)\frac{du}{dx}\right], \qquad u(\cos\beta) = 0 \tag{35.18}$$

имеет дискретный спектр[2]). Далее, оператор $m^2(1-x^2)^{-1}u$ — неотрицательный[3]) и, следовательно, оператор P не меньше оператора P_0: это значит, что

$$(Pu, u) \geqslant (P_0 u, u).$$

Повторяя рассуждения замечания 1 § 23, мы найдем, что спектр оператора P дискретен. В таком случае уравнение (35.16) имеет

[1]) См. А. Эрдейи и др. [1], т. 1, стр. 148.
[2]) См. статью автора [9].
[3]) При $m^2 > 0$ он положительно определенный.

счетное множество корней, положительных и отрицательных; каждому положительному корню k соответствует отрицательный корень $-k-1$.

Обозначим через $k_{n,m}$ $(n=1, 2, 3, \ldots)$ положительные корни уравнения (35.16), тогда собственные функции задачи Дирихле для сферического конуса можно представить в виде:

$$\varphi_{nlm}=c_{nlm}J_{k_{n,m}+\frac{1}{2}}\left(\gamma_{k_{n,m}+\frac{1}{2},l}\,r\right)P_{k_{n,m}}^{(m)}(\cos\theta)\begin{matrix}\cos m\varphi\\ \sin m\varphi\end{matrix} \tag{35.19}$$

$$(l,\ n=1,\ 2,\ 3,\ \ldots;\ m=0,\ 1,\ 2,\ \ldots),$$

коэффициент c_{nlm} можно определить из условия $|\varphi_{nlm}|_A=1$.

Задача Неймана. Аналогичные рассуждения дают в этом случае

$$\varphi_{nlm}=c_{nlm}J_{q_{n,m}+\frac{1}{2}}\left(\tilde{\gamma}_{q_{n,m}+\frac{1}{2},l}\,r\right)P_{q_{n,m}}^{(m)}(\cos\theta)\begin{matrix}\cos m\varphi\\ \sin m\varphi\end{matrix} \tag{35.20}$$

$$(l,\ n=1,\ 2,\ 3,\ \ldots;\ m=0,\ 1,\ 2,\ \ldots),$$

где $q_{n,m}$ суть корни уравнения

$$\left[\frac{d}{dx}P_q^{(m)}(x)\right]_{x=\cos\beta}=0. \tag{35.21}$$

Отметим случай $\beta=\frac{\pi}{2}$, когда сферический конус превращается в полушар. В этом случае, как легко видеть,

$$k_n=m+2n-1, \qquad n=1,\ 2,\ 3,\ \ldots \tag{35.22}$$

Чтобы в этом убедиться, заметим, что при таком определении k_n пробегает все целые значения, большие m и разной четности с m, а тогда функция $P_{k_n}^{(m)}(x)$ — нечетная и обращается в нуль при $x=0$, т. е. эта функция удовлетворяет уравнению (35.16) при $\beta=\frac{\pi}{2}$. Других значений k_n не существует. Действительно, в противном случае система

$$P_{m+2n-1}^{(m)}(x) \qquad (n=1,\ 2,\ 3,\ \ldots) \tag{35.23}$$

была бы неполной в промежутке $0<x<1$. Пусть функция $\omega(x)\in L_2(0,\ 1)$ ортогональна к функциям (35.23). Продолжив эту функцию нечетным образом на промежуток $-1<x<0$, мы нашли бы, что система (35.23) неполна на подпространстве нечетных функций пространства $L_2(-1,\ 1)$, а это неверно.

В случае задачи Неймана для полушара

$$q_n=m+2n \qquad (n=1,\ 2,\ 3,\ \ldots). \tag{35.24}$$

§ 36. Системы обыкновенных дифференциальных уравнений

Для простоты ограничимся системами второго порядка.

1. Первая краевая задача. Рассмотрим систему уравнений

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)+q(x)\,u(x)=f(x), \qquad 0<x<1; \quad (36.1)$$

при краевых условиях

$$u(0)=u(1)=0. \quad (36.2)$$

Здесь $u(x)$ и $f(x)$ — s-компонентные вектор-функции, $p(x)$ и $q(x)$ — эрмитовы матрицы порядка s. Матрицу $p(x)$ будем считать для простоты непрерывно дифференцируемой и положительно определенной на отрезке $0\leqslant x\leqslant 1$. В таком случае существует такая положительная постоянная p_0, что наименьшее собственное число матрицы $p(x)$ не меньше, чем p_0. Примем далее, что матрица $q(x)$ измерима, ограничена и неотрицательна. Левая часть уравнения (36.1) и краевые условия (36.2) порождают оператор, который мы обозначим через A и который будем считать заданным на множестве $D(A)$ функций, обладающих следующими свойствами: эти функции и их первые производные абсолютно непрерывны на отрезке $0\leqslant x\leqslant 1$; сами функции обращаются в нуль на концах этого отрезка; их вторые производные на том же отрезке суммируемы с квадратом. Легко убедиться, что оператор A — положительно определенный и что его спектр дискретен [1]).

Область $D(A)$ задания оператора A не зависит от специального выбора матриц $p(x)$ и $q(x)$, поэтому, если B есть оператор, заданный дифференциальным выражением

$$Bu=-\frac{d}{dx}\left(p_1(x)\frac{du}{dx}\right)+q_1(x)\,u \quad (36.3)$$

на том же множестве функций $D(A)$, причем матрицы s-го порядка $p_1(x)$ и $q_1(x)$ удовлетворяют перечисленным выше условиям для $p(x)$ и $q(x)$, то $D(B)=D(A)$, и операторы A и B — сходные. Докажем, что *если матрицы $p(x)$ и $p_1(x)$ перестановочны, то операторы A и B — родственные.*

---

[1]) Свойства положительной определенности и дискретности оператора A сохраняются и тогда, когда матрица $q(x)$ не является обязательно неотрицательной, а удовлетворяет следующему, более общему требованию: ее наименьшее собственное число не меньше, чем $\varepsilon-\lambda_1$, где ε — какое-либо положительное число, а λ_1 — наименьшее собственное число оператора

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right),$$

заданного на том же множестве $D(A)$.

Имеем

$$(Au,\ (B+kI)\,u)=\int_0^1 pu''\cdot p_1u''\,dx+\int_0^1 pu''\cdot\left(p_1'u'-q_1u\right)dx+\int_0^1 (p'u'-qu)\cdot p_1'u''\,dx+\int_0^1 (p'u'-qu)\cdot\left(p_1'u'-q_1u\right)dx+k\,|\,u\,|_A^2=J_1+J_2+J_3+J_4+k\,|u|_A^2; \tag{36.4}$$

точкой обозначено скалярное умножение векторов в унитарном s-мерном пространстве. Оценим каждый из интегралов в (36.4). Начнем с интеграла

$$J_1=\int_0^1 pu''\cdot p_1u''\,dx=\int_0^1 p_1pu''\cdot u''\,dx.$$

Докажем, что матрица p_1p — положительно определенная. Прежде всего она симметрична. Далее, обе матрицы одним и тем же унитарным преобразованием приводятся к диагональной форме[1]); существует, следовательно, такая унитарная матрица $U(x)$, что

$$\left.\begin{aligned} p(x)&=U^{-1}(x)\,[\mu_1(x),\ \mu_2(x),\ \ldots,\ \mu_s(x)]\,U(x),\\ p_1(x)&=U^{-1}(x)\,[\nu_1(x),\ \nu_2(x),\ \ldots,\ \nu_s(x)]\,U(x),\end{aligned}\right\} \tag{36.5}$$

где квадратные скобки обозначают диагональную матрицу, а $\mu_j(x)$ и $\nu_j(x)$ суть (расположенные в порядке возрастания) собственные числа матриц $p(x)$ и $p_1(x)$ соответственно.

Отсюда

$$p_1(x)\,p(x)=U^{-1}(x)\,[\mu_1(x)\,\nu_1(x),\ \mu_2(x)\,\nu_2(x),\ \ldots,\ \mu_s(x)\,\nu_s(x)]\,U(x),$$

и ясно, что матрица $p_1(x)\,p(x)$ — положительно определенная, причем ее наименьшее собственное число $\mu_1(x)\,\nu_1(x)$ не меньше положительной постоянной $\tilde{p}=p_0p_1$, где $0<p_1\leqslant\nu_1(x)$.

Теперь

$$J_1\geqslant\tilde{p}\int_0^1 |\,u''\,|^2\,dx. \tag{36.6}$$

[1]) См. В. И. Смирнов [2], стр. 151—152.

Перейдем к интегралу J_2. Матрицы p, p_1', q_1 ограничены на отрезке $0 \leqslant x \leqslant 1$, поэтому

$$|J_2| = \left| \int_0^1 pu'' \cdot (p_1'u' - q_1 u)\, dx \right| \leqslant c_1 \int_0^1 |u''| \{ |u'| + |u| \} dx \leqslant$$

$$\leqslant \frac{c_1\varepsilon}{2} \int_0^1 |u''|^2 dx + \frac{c_1}{2\varepsilon} \int_0^1 \{ |u'| + |u| \}^2 dx \leqslant$$

$$\leqslant \frac{c_1\varepsilon}{2} \int_0^1 |u''|^2 dx + \frac{c_1}{\varepsilon} \int_0^1 \{ |u'|^2 + |u|^2 \} dx;$$

здесь c_1 — некоторая определенная, а ε — произвольно малая постоянная. Второй интеграл справа легко оценивается через $|u|_A^2$ и окончательно

$$|J_2| \leqslant \frac{c_1\varepsilon}{2} \int_0^1 |u''|^2 dx + \frac{c^2}{\varepsilon} |u|_A^2, \qquad c_2 = \text{const} > 0. \tag{36.7}$$

Нетрудно видеть, что этому же неравенству удовлетворяет и интеграл J_3, если только постоянные c_1 и c_2 взяты достаточно большими:

$$|J_3| \leqslant \frac{c_1\varepsilon}{2} \int_0^1 |u''|^2 dx + \frac{c_2}{\varepsilon} |u|_A^2. \tag{36.8}$$

Наконец, интеграл J_4 просто оценивается через $|u|_A^2$ (ср. оценку интеграла J_2):

$$|J_4| \leqslant c_3 |u|_A^2, \qquad c_3 = \text{const}. \tag{36.8'}$$

Из формул (36.4) и (36.6) — (36.8′) следует:

$$|(Au,\ (B + kI)\,u)| \geqslant (\tilde{p} - c_1\varepsilon) \int_0^1 |u''|^2 dx + \left(k - \frac{2c_2}{\varepsilon} - c_3\right) |u|_A^2.$$

Возьмем $\varepsilon \leqslant \frac{\tilde{p}}{2c_1}$ и $k \geqslant \frac{2c_2}{\varepsilon} + c_3$, тогда

$$|(Au,\ (B + kI)\,u)| \geqslant \frac{1}{2} \tilde{p} \int_0^1 |u''|^2 dx.$$

Оператор $-\frac{d^2}{dx^2}$, заданный на множестве $D(A)$, — самосопряженный и положительно определенный и, следовательно, сходный с A; по теореме 3.1 существует такая положительная постоянная (обозначим ее c_4), что

$$\int_0^1 |u''|^2\,dx \geqslant c_4 \|Au\|^2.$$

Окончательно

$$|(Au,\ (B+kI)\,u)| \geqslant a\|Au\|^2, \qquad a = \frac{\tilde{p}c_4}{2},$$

что и требовалось доказать.

Положим теперь $p_1(x) = I_s$, $q_1(x) = 0$ (I_s — единичная матрица порядка s). Тогда

$$Bu = -\frac{d^2u}{dx^2}, \qquad u(0) = u(1) = 0.$$

Собственные векторы оператора B, нормированные в метрике H_B, суть

$$\varphi_{nl}(x) = \left(0,\ \ldots,\ 0, \frac{\sqrt{2}}{\pi n}\sin n\pi x,\ 0,\ \ldots,\ 0\right) \tag{36.9}$$
$$(l = 1,\ 2,\ \ldots,\ s;\ n = 1,\ 2,\ 3,\ \ldots),$$

где отличный от нуля элемент занимает l-е место. Множитель $\frac{\sqrt{2}}{\pi}$ можно отбросить и положить

$$\varphi_{nl}(x) = \left(0,\ \ldots,\ 0, \frac{1}{n}\sin n\pi x,\ 0,\ \ldots,\ 0\right) \tag{36.9$'$}$$
$$(l = 1,\ 2,\ \ldots,\ s;\ n = 1,\ 2,\ 3,\ \ldots).$$

Если векторы (36.9) или (36.9′) взяты в качестве координатных при решении задачи (36.1) — (36.2), то имеют место утверждения а) — в) § 29.

Можно воспользоваться также системой

$$\varphi_{nl}(x) = (0,\ \ldots,\ 0,\ Q_n(x),\ 0,\ \ldots,\ 0) \tag{36.10}$$
$$(l = 1,\ 2,\ \ldots,\ s;\ n = 1,\ 2,\ \ldots),$$

где $Q_n(x)$ определены формулой (30.6); эта система гарантирует справедливость утверждений а) и б) § 29.

2. Краевая задача с естественными условиями. Систему уравнений (36.1) рассмотрим при краевых условиях

$$u'(0) - M_0 u(0) = 0, \qquad u'(1) + M_1 u(1) = 0. \tag{36.11}$$

Сохраним все предположения п. 1 о матрице $p(x)$; матрицу $q(x)$ для простоты будем считать непрерывной и неотрицательной. Если через A обозначить теперь оператор левой части уравнения (36.1) при краевых условиях (36.11), то

$$(Au, u) = p(0)\, M_0 u(0) \cdot \overline{u(0)} + p(1)\, M_1 u(1) \cdot u(1) + \\ + \int_0^1 \{p(x)\, u'(x) \cdot u'(x) + q(x)\, u(x) \cdot u(x)\}\, dx \qquad (36.12)$$

и ясно, что оператор A будет положительно определенным, если матрицы $p(0)\, M_0$ и $p(1)\, M_1$ будут неотрицательными, причем хотя бы одна из них будет положительной. Обозначим теперь через B оператор

$$Bu = -\frac{d^2u}{dx^2} + u \qquad (36.13)$$

при краевых условиях

$$u'(0) = u'(1) = 0. \qquad (36.14)$$

Нетрудно видеть, что операторы A и B полусходные. Собственные векторы оператора B, нормированные в H_B, суть

$$\varphi_{nl} = \left\{0, \ldots, 0, \sqrt{\frac{2}{n^2\pi^2+1}} \cos n\pi x, 0, \ldots, 0\right\} \qquad (36.15)$$

$$(l = 1, 2, \ldots, s; \quad n = 0, 1, 2, \ldots);$$

отличная от нуля составляющая стоит на l-м месте, при $n = 0$ нормирующий множитель следует положить равным единице. Нормирующий множитель $\sqrt{\frac{2}{n^2\pi^2+1}}$, $n > 0$, можно заменить более простым $\frac{1}{n}$. Мы получим тогда систему

$$\varphi_{nl}(x) = \left(0, \ldots, 0, \frac{1}{n} \cos n\pi x, 0, \ldots, 0\right) \qquad (36.16)$$

$$(l = 1, 2, \ldots, s; \quad n = 0, 1, 2, \ldots);$$

при $n = 0$ следует заменить $\frac{1}{n}$ на 1. Система (36.16) почти ортонормирована в H_A; принятая за координатную для задачи (36.1), (36.11), она обеспечивает справедливость утверждений а) и б) § 29.

Аналогично тому, как это было сделано в § 30 для одного уравнения, можно и здесь построить отличные от (36.15) системы, ортонормированные в H_B. Мы не будем на этом останавливаться.

§ 37. Системы уравнений в частных производных

Как и в предшествующем параграфе, мы ограничимся системой второго порядка.

Рассмотрим систему s уравнений второго порядка следующего вида

$$-\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right)+Cu=f(x). \tag{37.1}$$

Здесь $u(x)$ и $f(x)$ — s-компонентные вектор-функции, определенные почти всюду в некоторой конечной области Ω пространства координат $(x_1, x_2, \ldots, x_m)$, A_{jk} и C — матрицы s-го порядка, элементы которых суть функции от x, определенные и для простоты непрерывные в замкнутой области $\overline{\Omega}=\Omega\cup S$. Относительно матриц A_{jk} примем еще, что они кусочно дифференцируемы в Ω и удовлетворяют следующим соотношениям:

а) $$A_{jk}(x)=\overset{*}{A}_{kj}(x), \tag{37.2}$$

где звездочка означает эрмитово-сопряженную матрицу;

б) каковы бы ни были s-компонентные векторы $t_1, t_2, \ldots, t_m$, имеет место неравенство

$$\sum_{j,\,k=1}^{m} A_{jk}(x)\,t_k\cdot t_j \geqslant \mu_0\sum_{k=1}^{m}\|t_k\|^2, \qquad \mu_0=\mathrm{const}>0; \tag{37.3}$$

в формуле (37.3) точка означает обычное скалярное умножение s-компонентных векторов, а знак $\|\ \ \|$ — длину такого вектора.

Границу S будем считать кусочно гладкой; систему (37.1) будем рассматривать при краевых условиях одного из двух видов

$$u|_S=0 \tag{37.4}$$

или

$$\left[\sum_{j,\,k=1}^{m} A_{jk}\frac{\partial u}{\partial x_k}\cos(\nu,\ x_j)+\sigma u\right]_S=0, \tag{37.5}$$

где σ — измеримая ограниченная неотрицательная матрица, заданная на S.

Рассмотрим оператор

$$-\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right) \tag{37.6}$$

при одном из краевых условий (37.4) или (37.5). Спектр этого оператора дискретен и положителен во всех случаях, кроме одного, когда ставится условие (37.5) и $\sigma\equiv 0$; в этом частном случае наименьшее собственное число оператора (37.6) равно нулю. В общем случае обозначим наименьшее собственное число оператора (37.6) через $\lambda_1(x)$.

Примем следующие допущения о матрице C: а) эта матрица эрмитова; б) ее наименьшее собственное число $\gamma_1(x)$ удовлетворяет неравенству $\gamma_1(x) \geqslant -\lambda_1(x) + \varepsilon$, где ε — положительная постоянная.

При перечисленных условиях оператор каждой из задач (37.1), (37.4) или (37.1), (37.5) положительно определенный; обозначая этот оператор через A, имеем:

$$|u|_A = \int_\Omega \left\{ \sum_{j,\,k=1}^{m} A_{jk} \frac{\partial u}{\partial x_k} \cdot \frac{\partial u}{\partial x_j} + Cu \cdot u \right\} dx$$

в случае условия (37.4) и

$$|u|_A^2 = \int_\Omega \left\{ \sum_{j,\,k=1}^{m} A_{jk} \frac{\partial u}{\partial x_k} \cdot \frac{\partial u}{\partial x_j} + Cu \cdot u \right\} dx + \int_S \sigma u \cdot u\, ds$$

в случае условия (37.5). Из соотношения (37.3) и из непрерывности матриц A_{jk} легко вытекает, что оператор A — полусходный с оператором B, который определяется дифференциальным выражением

$$Bu = -\Delta u + u$$

и которому сопутствует краевое условие (37.4), если это условие сопутствует уравнению (37.1), или краевое условие

$$\left. \frac{\partial u}{\partial \nu} \right|_S = 0, \tag{37.4$'$}$$

если уравнению (37.1) сопутствует условие (37.5); в случае условия (37.4) можно положить $Bu = -\Delta u$.

Пусть $\varphi_n(x)$ $(n = 1, 2, 3, \ldots)$ — собственные функции скалярного оператора Лапласа при краевом условии (37.4) или (37.4′) соответственно. Тогда собственные векторы оператора B суть

$$\varphi_{nl}(x) = \{0, \ldots, 0, \varphi_n(x), 0, \ldots, 0\} \tag{37.7}$$
$$(l = 1, 2, \ldots, s; \quad n = 1, 2, 3, \ldots);$$

составляющая $\varphi_n(x)$ занимает l-е место. Если функции $\varphi_n(x)$ ортонормированы в метрике интеграла Дирихле, т. е.

$$\int_\Omega \sum_{k=1}^{m} \frac{\partial \varphi_n}{\partial x_k} \frac{\partial \varphi_p}{\partial x_k} dx = \delta_{np},$$

то векторы (37.7) ортонормированы в метрике H_B и, следовательно, почти ортонормированы в метрике H_A. Функции $\varphi_n(x)$ сравнительно просто вычисляются в случаях, указанных в §§ 33—35; если область Ω может быть преобразована в одну из простых областей, о которых идет речь в упомянутых параграфах, то можно предварительно преобразовать к такой области данную задачу.

Само собой разумеется, в качестве $\varphi_n(x)$ можно брать не собственные функции оператора Лапласа, а любые функции, ортонормированные в метрике интеграла Дирихле. Заметим еще, что соотношение (37.3) не необходимо; достаточно, чтобы

$$|u|_A \geqslant C|u|_B, \qquad C = \text{const} > 0.$$

Такое соотношение имеет место, например, для основных задач статической теории упругости в силу так называемого *неравенства Корна* (ПМ, гл. IV).

§ 38. Координатные системы для метода наименьших квадратов

Будем предполагать, как обычно в методе наименьших квадратов, что входящий в уравнение

$$Au = f \tag{38.1}$$

оператор A — замкнутый (не обязательно симметричный), действующий из гильбертова сепарабельного пространства H в сепарабельное же гильбертово пространство H_1 и что обратный оператор A^{-1} определен на всем пространстве H_1 и, следовательно, ограничен. Напомним[1]), что метод наименьших квадратов состоит в следующем: выбираем координатную систему φ_n ($n = 1, 2, 3, \ldots$), подчиненную следующим условиям: 1) $\varphi_n \in D(A)$; 2) элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы при любом n; 3) система $\{A\varphi_k\}$ полна в H_1, и применяем процесс Ритца к функционалу $\|Au - f\|^2$. Это приводит к системе уравнений

$$\sum_{k=1}^{n} \left(A\varphi_k, A\varphi_j\right) a_k^{(n)} = \left(f, A\varphi_j\right) \qquad (j = 1, 2, \ldots, n), \tag{38.2}$$

матрицу системы (38.2) будем называть *матрицей метода наименьших квадратов*.

Как мы видели в § 6, в данном случае пространство H_A состоит из элементов множества $D(A)$, на которых скалярное произведение и норма задаются формулами

$$[u, v]_A = (Au, Av), \qquad |u|_A = \|Au\|. \tag{38.3}$$

Из общих результатов §§ 9—11 вытекает следующее:

1) Для того чтобы процесс наименьших квадратов[2]) и полученное этим методом приближенное решение были устойчивы, необхо-

---

[1]) См. ВМ, гл. X.

[2]) То есть процесс Ритца, примененный к функционалу метода наименьших квадратов.

димо и достаточно, чтобы координатная система была сильно минимальна в метрике (38.3).

2) Для того чтобы число обусловленности матрицы метода наименьших квадратов было ограничено независимо от ее порядка, необходимо и достаточно, чтобы координатная система была почти ортонормирована в метрике (38.3).

Заметим еще, что при любом выборе координатной системы, удовлетворяющей перечисленным выше условиям 1) — 3), невязка приближенного решения, построенного по методу наименьших квадратов, стремится к нулю (см. ВМ, стр. 412).

Пространство H_A вкладывается в H. Действительно, если $u \in H_A$, то $u \in D(A)$ и тем более $u \in H$. Далее, из ограниченности оператора A^{-1} вытекает существование такой постоянной k, что $\|Au\| \geqslant k\|u\|$ или, что то же, $\|u\| \leqslant k^{-1}|u|_A$, и неравенство (4.1) выполнено с постоянной $K = k^{-1}$.

Из результатов § 4 теперь вытекает следующее:

1. Если координатная система сильно минимальна (например, ортонормирована) в данном пространстве H, то она сильно минимальна в пространстве H_A (т. е. в метрике (38.3)).

2. Расширим понятие сходного оператора, именно, два *замкнутых* оператора A и B, действующих из H в H_1, будем называть *сходными*, если: 1) $D(A) = D(B)$; 2) $R(A) = R(B) = H_1$, где $R(A)$ означает множество значений оператора A; 3) операторы A^{-1} и B^{-1} существуют — тогда они ограничены. Как легко видеть, при таком определении остается в силе теорема 3 1. Нетрудно также убедиться, что пространства H_A и H_B вкладываются одно в другое и, следовательно, система, почти ортонормированная и полная в метрике H_B, т. е. в метрике, определяемой формулами

$$[u, v]_B = (Bu, Bv), \qquad |u|_B = \|Bu\|,$$

будет почти ортонормированной и полной в метрике (38.3); в качестве координатной системы можно взять, например, любую полную ортонормированную в H_B систему. В частности, если $H_1 = H$, а оператор B — самосопряженный, с точечным спектром, то можно за координатную выбрать систему собственных элементов $\{\varphi_n\}$ этого оператора, нормированных равенством $\|B\varphi_n\| = 1$.

Рассмотрим один важный пример сильно минимальной системы, связанной с методом наименьших квадратов. Пусть Γ — замкнутая кривая на комплексной плоскости; мы предположим, что эта кривая достаточно гладкая и что она ограничивает некоторую односвязную область D. Начало координат поместим внутри Γ. Как обычно, декартовы координаты на комплексной плоскости обозначим через x и y и положим $z = x + iy$, $\bar{z} = x - iy$; комплексные координаты (аффиксы) переменных точек на кривой Γ мы будем обозначать также буквами ζ и t. Введем в рассмотрение гильбертово пространство $L_2(\Gamma)$ функций, определенных почти всюду на Γ и суммируемых с квадратом вдоль Γ. Легко доказать, что $L_2(\Gamma)$ есть прямая сумма двух подпространств, которые мы обозначим H_1 и H_2; первое образовано функциями из $L_2(\Gamma)$, которые являются предельными значениями функций, голоморфных внутри Γ, второе — функциями из $L_2(\Gamma)$, которые являются предельными значениями функций, комплексно сопряженных к функциям, голоморфным внутри Γ и равным нулю в начале координат. Действительно, сказанное означает, что если $f(\zeta) = L_2(\Gamma)$, то, во-первых,

$$f(\zeta) = f_1(\zeta) + \overline{f_2(\zeta)}, \tag{38.4}$$

где $f_1(\zeta)$, $f_2(\zeta) \in L_2(\Gamma)$ и существуют аналитические продолжения $f_1(z)$ и $f_2(z)$ внутрь контура Γ, причем $f_2(0) = 0$, и, во-вторых, что разложение (38.4) — единственное. Допуская, что разложение (38.4) возможно, умножим обе его части, а также обе части равенства $\overline{f(\zeta)} = \overline{f_1(\zeta)} + f_2(\zeta)$, полученного переходом к сопряженным величинам на ядро Шварца[1]) $T(z;\ \zeta)$ области D, которое мы нормируем условием

$$\operatorname{Im}\{T(0,\ \zeta)\} = 0;$$

результат умножения проинтегрируем по Γ.

Приняв во внимание, что $f_2(0) = 0$, получим:

$$f_1(z) - \frac{1}{2} f_1(0) = \frac{1}{4\pi} \int\limits_{\Gamma} f(\zeta)\, T(z;\ \zeta)\, d\sigma,$$

$$f_2(z) + \frac{1}{2} \overline{f_1(0)} = \frac{1}{4\pi} \int\limits_{\Gamma} \overline{f(\zeta)}\, T(z;\ \zeta)\, d\sigma, \qquad d\sigma = |d\zeta|.$$

Полагая $z = 0$ в первом из этих равенств, получим:

$$\frac{1}{2} f_1(0) = \frac{1}{4\pi} \int\limits_{\Gamma} f(\zeta)\, T(0,\ \zeta)\, d\sigma$$

и, следовательно,

$$f_1(z) = \frac{1}{4\pi} \int\limits_{\Gamma} f(\zeta)\, [T(z;\ \zeta) + T(0,\ \zeta)]\, d\sigma, \tag{38.5}$$

$$f_2(z) = \frac{1}{4\pi} \int\limits_{\Gamma} \overline{f(\zeta)}\, [T(z;\ \zeta) - T(0,\ \zeta)]\, d\sigma. \tag{38.6}$$

Отсюда видно, что разложение (38.4) единственно.

---

[1]) См. книгу автора [3], § 41, особенно формулы (22), (23) и (27).

Докажем теперь, что формулы (38.5) и (38.6) действительно реализуют это разложение. Из формулы (38.6) сразу вытекает, что $f_2(0)=0$.

Далее, формулы (38.5) и (38.6) дают:

$$f_1(z)+\overline{f_2(z)}=\frac{1}{2\pi}\int\limits_{\Gamma} f(\zeta)\operatorname{Re}\{T(z;\ \zeta)\}\,d\sigma=\frac{1}{2\pi}\int\limits_{\Gamma} f(\zeta)\frac{\partial G}{\partial \nu}\,d\sigma,$$

где G — функция Грина области D и ν — внутренняя нормаль к Γ. Полагая $z\to t\in\Gamma$, мы в силу известных свойств функции Грина найдем, что $f_1(t)+\overline{f_2(t)}=f(t)$.

Имеет место представление

$$\frac{1}{4\pi}T(z;\ \zeta)\,d\sigma=\frac{1}{2\pi i}\,\frac{d\zeta}{\zeta-z}+P(z;\ \zeta)\,d\sigma,$$

где $P(z;\ \zeta)$ непрерывна, когда ζ пробегает контур Γ, а z — замкнутую область $\bar{D}=D\cup\Gamma$. Пользуясь хорошо известной предельной формулой для интегралов типа Коши, получим:

$$f_1(t)=\frac{1}{2}f(t)+\frac{1}{2\pi i}\int\limits_{\Gamma} f(\zeta)\frac{d\zeta}{\zeta-t}+\frac{1}{4\pi}\int\limits_{\Gamma}[P(t,\zeta)-T(0,\zeta)]\,d\sigma,\quad t\in\Gamma.$$

Первый интеграл справа следует рассматривать как сингулярный, т. е. как главное значение по Коши. Хорошо известно[1]), что этот интеграл представляет собой оператор над f, ограниченный в $L_2(\Gamma)$; отсюда легко усмотреть, что $f_1\in L_2(\Gamma)$.

Аналогично находим, что $f_2\in L_2(\Gamma)$, и наше утверждение доказано.

Из хорошо известных теорем теории функций комплексного переменного вытекает, что в подпространствах H_1 и H_2 полны соответственно системы z^n $(n=0,\ 1,\ 2,\ \ldots)$ и $\bar{z}^n$ $(n=1,\ 2,\ 3,\ \ldots)$. Обозначим $\rho_1=\inf\limits_{\zeta\in\Gamma}|\zeta|$ и пусть $0<\rho_0<\rho_1$. Докажем, что *системы*

$$\left(\frac{z}{\rho_0}\right)^n\qquad(n=0,\ 1,\ 2,\ \ldots)\tag{38.7}$$

и

$$\left(\frac{\bar{z}}{\rho_0}\right)^n\qquad(n=1,\ 2,\ 3,\ \ldots)\tag{38.8}$$

в соответствующих подпространствах H_1 и H_2 сильно минимальны.

Достаточно, очевидно, провести рассуждения для системы (38.7). Заменой переменных можно сделать $\rho_0=1$; тогда $|\zeta|>1$, $\zeta\in\Gamma$.

---

[1]) См., например, А. Зигмунд [1], стр. 404, где приведено доказательство более общего утверждения.

Рассмотрим матрицу Грама элементов z^k $(k=0, 1, 2, \ldots, n)$ и ее наименьшее собственное число λ_1:

$$\lambda_1^{(n)} = \inf \frac{\left\| \sum\limits_{k=0}^{n} a_k \zeta^k \right\|^2}{\sum\limits_{k=0}^{n} |a_k|^2} = \inf \left\{ \sum_{k=0}^{n} |a_k|^2 \right\}^{-1} \int_{\Gamma} |\Phi(\zeta)|^2 \, d\sigma, \qquad (38.9)$$

где

$$\Phi(\zeta) = \sum_{k=0}^{n} a_k \zeta^k.$$

Круг $|z| \leqslant 1$ целиком расположен внутри Γ (рис. 3). Докажем, что

$$\int_{\Gamma} |\Phi(\zeta)|^2 \, d\sigma \geqslant c \int_{\Gamma_1} |\Phi(\zeta)|^2 \, d\sigma, \qquad c = \text{const} > 0, \qquad (38.10)$$

где Γ_1 — окружность $|\zeta| = 1$. Конформно отобразим область D на единичный круг плоскости τ; при этом Γ перейдет в единичную

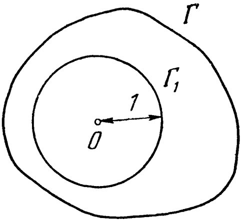

Рис. 3.

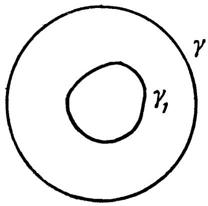

Рис. 4.

окружность γ, а окружность Γ_1 — в некоторую кривую γ_1, расположенную внутри γ (рис. 4). Отображающую функцию обозначим через $\omega(\tau)$; так как Γ — достаточно гладкая кривая, то производная $\omega'(\tau)$ непрерывна и отлична от нуля при $|\tau| \leqslant 1$. Обозначим еще $\Phi(z) = \Phi(\omega(\tau)) = \varphi(\tau)$. Пусть

$$\varphi(\tau) = \sum_{n=0}^{\infty} \alpha_n \tau^n,$$

тогда

$$\int_{\gamma} |\varphi(\tau)|^2 \, |d\tau| = 2\pi \sum_{n=0}^{\infty} |\alpha_n|^2.$$

Если $|\tau| < 1$, то по неравенству Коши

$$|\varphi(\tau)|^2 \leqslant \sum_{n=0}^{\infty} |\alpha_n|^2 \sum_{n=0}^{\infty} |\tau|^n = \frac{1}{2\pi(1-|\tau|)} \int_{\gamma} |\varphi(\tau)|^2 \, d\tau$$

и, следовательно,

$$\int_{\gamma_1} |\varphi(\tau)|^2 |d\tau| \leqslant \frac{l}{2\pi(1-\tau_1)} \int_{\gamma} |\varphi(\tau)|^2 d\tau;$$

здесь l — длина кривой γ_1 и $\tau_1 = \max|\tau|$, $\tau \in \gamma_1$. Возвращаясь к переменной ζ, имеем:

$$\int_{\Gamma_1} |\Phi(\zeta)|^2 \frac{d\sigma}{|\omega'(\tau)|} \leqslant \frac{l}{2\pi(1-\tau_1)} \int_{\Gamma} |\Phi(\zeta)|^2 \frac{d\sigma}{|\omega'(\tau)|}.$$

Величина $|\omega'(\tau)|$ заключена между двумя положительными постоянными; пусть $0 < \alpha \leqslant |\omega'(\tau)| \leqslant \beta < \infty$, тогда из последнего неравенства легко вытекает неравенство (38.10) со значением

$$c = \frac{2\pi\alpha(1-\tau_1)}{\beta l}.$$

Остается теперь заметить, что

$$\int_{\Gamma_1} |\Phi(\zeta)|^2 d\zeta = 2\pi \sum_{k=0}^{n} |a_k|^2, \tag{38.11}$$

и неравенство (38.9) дает $\lambda_1^{(n)} \geqslant 2\pi c$; наше утверждение доказано.

Теперь нетрудно доказать следующую теорему.

Теорема 38.1. *Система, полученная объединением систем* (38 7) *и* (38.8), *сильно минимальна в* $L_2(\Gamma)$.

Как и выше, будем считать, что $\rho_0 = 1$.

Наименьшее собственное число $\lambda_{m,n}^{(1)}$ матрицы Грама функций $1, z, \ldots, z^m, \bar{z}, \ldots, \bar{z}^n$ равно

$$\lambda_{m,n}^{(1)} = \inf \frac{\|\Phi(\zeta) + \overline{\Psi(\zeta)}\|^2}{\sum_{k=0}^{m} |a_k|^2 + \sum_{k=1}^{n} |b_k|^2}, \tag{38.12}$$

где

$$\Phi(\zeta) = \sum_{k=0}^{m} a_k \zeta^k, \qquad \Psi(\zeta) = \sum_{k=1}^{n} b_k \zeta^k.$$

Введем опять функцию $\omega(\tau)$, конформно отображающую область D на единичный круг; потребуем, чтобы $\omega(0) = 0$. Положим $\varphi(\tau) = \Phi(\omega(\tau))$, $\psi(\tau) = \Psi(\omega(\tau))$ На единичной окружности γ функции $\varphi(\tau)$ и $\overline{\psi(\tau)}$ ортогональны, так как $\psi(0) = \Psi(0) = 0$. Поэтому

$$\int_{\gamma} |\varphi(\tau) + \overline{\psi(\tau)}|^2 |d\tau| = \int_{\gamma} |\varphi(\tau)|^2 |d\tau| + \int_{\gamma} |\psi(\tau)|^2 |d\tau|.$$

или, если вернуться к переменной ζ и заменить $|\omega'(\tau)|$ справа его наименьшим значением, а слева — наибольшим,

$$\|\Phi(\zeta)+\overline{\Psi(\zeta)}\|^2 \geqslant \frac{\alpha}{\beta}\{\|\Phi(\zeta)\|^2+\|\Psi(\zeta)\|^2\};$$

норма берется в метрике пространства $L_2(\Gamma)$.

Выше мы видели (формулы (38.10) и (38.11)), что

$$\|\Phi(\zeta)\|^2 \geqslant 2\pi c \sum_{k=0}^{m} |a_k|^2.$$

Аналогично

$$\|\Psi(\zeta)\|^2 \geqslant 2\pi c \sum_{k=1}^{n} |b_k|^2.$$

По формуле (38.12)

$$\lambda_{m,n}^{(1)} \geqslant \frac{2\pi c\alpha}{\beta}.$$

Правая часть последнего неравенства положительна и не зависит от m и n. Теорема доказана.

Следствие 38.1. *Пусть в пространстве* $L_2(\Gamma)$ *задан замкнутый оператор* A, *имеющий ограниченный обратный. Если* $z^k \in D(A)$, $\bar{z}^k \in D(A)$ $(k=0, 1, 2, \ldots)$, *то система, полученная объединением систем* (38.7) *и* (38.8), *сильно минимальна в пространстве* H_A, *в котором метрика определяется формулами*

$$[u, v]_A = (Au, Av), \qquad |u|_A = \|Au\|.$$

Из сказанного вытекает, что система (38.7) — (38.8) сильно минимальна в пространствах H_A, соответствующих внутренним задачам Дирихле, Неймана и смешанной для уравнения Лапласа на плоскости[1]).

Пусть теперь

$$\rho^{(0)} > \max_{\zeta\in\Gamma} |\zeta|.$$

Аналогично предыдущему легко доказать, что система функций

$$\left(\frac{\rho^0}{z}\right)^k \quad (k=0, 1, 2, \ldots), \qquad \left(\frac{\rho}{\bar{z}}\right)^k \quad (k=1, 2, 3, \ldots) \tag{38.13}$$

полна и сильно минимальна в $L_2(\Gamma)$; она, следовательно, сильно минимальна в пространствах H_A, соответствующих внешним задачам Дирихле, Неймана и смешанной, для уравнения Лапласа на плоскости.

---

[1]) См. ВМ, § 87 и § 90; в случае задачи Неймана из системы (38.7) — (38.8) следует исключить единицу.

Обратимся теперь к случаю многосвязной[1]) области D. Пусть контур Γ состоит из замкнутых достаточно гладких кривых Γ_0, $\Gamma_1, \ldots, \Gamma_s$; примем, что D — конечная область и что кривая Γ_0 ограничивает ее извне. Начало координат поместим внутри D; внутри каждой из кривых Γ_j ($j = 1, 2, \ldots, s$) выберем точку z_j.

Пусть

$$\rho_0 < \min_{\zeta \in \Gamma_0} |\zeta|, \quad \rho^{(j)} > \max_{\zeta \in \Gamma_j} |\zeta - z_j| \qquad (j = 1, 2, \ldots, s).$$

Тогда система (38.7) — (38.8) сильно минимальна в $L_2(\Gamma_0)$, а система

$$\left(\frac{\rho^{(i)}}{\zeta - z_j}\right)^k \quad (k = 0, 1, 2, \ldots), \qquad \left(\frac{\rho^{(j)}}{\bar\zeta - \bar z_j}\right)^k \quad (k = 1, 2, \ldots) \qquad (38.14)$$

сильно минимальна в $L_2(\Gamma_j)$ $(j = 1, 2, \ldots, s)$.

Нетрудно видеть, что пространство $L_2(\Gamma)$ есть прямая сумма пространств $L_2(\Gamma_0)$, $L_2(\Gamma_1)$, ..., $L_2(\Gamma_s)$; отсюда легко следует, что объединение систем (38.7), (38.8), (38.14) сильно минимально в $L_2(\Gamma)$. Теперь ясно, что для основных задач теории гармонических функций на плоскости можно использовать указанное объединение в качестве координатной системы в методе наименьших квадратов, и при этом как вектор коэффициентов $a_k^{(n)}$, так и само приближенное решение будут устойчивы.

§ 39. Интегральные уравнения

Если в уравнении

$$Au = f \tag{39.1}$$

оператор A ограничен, то подбор координатной системы сильно упрощается. Так, например, если уравнение (39.1) решается по методу наименьших квадратов, то любая ортонормированная в данном гильбертовом пространстве система будет почти ортонормирована в метрике $|u|_A = \|Au\|$; мы предполагаем при этом, разумеется, что существует ограниченный обратный оператор A^{-1}. То же будет, если A — ограниченный самосопряженный положительно определенный оператор и $|u|_A^2 = \left\| A^{\frac{1}{2}} u \right\|^2 = (Au, u)$. Эти соображения мы поясняем ниже на простейших типах интегральных уравнений.

1. Уравнения Фредгольма. Если оператор T вполне непрерывен в сепарабельном пространстве H, и уравнение

$$u + Tu = f, \qquad f \in H \tag{39.2}$$

имеет в H единственное решение, то это уравнение можно приближенно решать, применяя процесс Бубнова — Галёркина. Как было

[1]) См. статью автора [2].

выяснено в § 14, этот процесс устойчив, если координатная система сильно минимальна в H. Такую систему в принципе построить нетрудно: достаточно взять любую полную ортонормированную в H систему. Эту последнюю желательно выбирать так, чтобы выражения $T\varphi_n$, где φ_n — координатные элементы, вычислялись по возможности проще. Заметим, что для дифференциальных операторов аналогичное требование не играет особой роли, но вполне непрерывные операторы обычно суть операторы интегральные, и если Ω — множество m измерений, а

$$Tu = \int_\Omega K(x,\ y)\, u(y)\, dy,$$

то, например, выражение

$$(T\varphi_k,\ \varphi_j) = \int_\Omega \int_\Omega K(x,\ y)\, \varphi_k(y)\, \overline{\varphi_j(x)}\, dy$$

представляет собой интеграл кратности $2m$; если функции

$$T\varphi_k = \int_\Omega K(x,\ y)\, \varphi_k(y)\, dy$$

удается просто вычислить, то скалярные произведения $(T\varphi_k,\ \varphi_j)$ выразятся через интегралы m-й кратности

Поясним сказанное примером. Пусть требуется найти функцию $u(x,\ y)$, гармоническую в круге $x^2 + y^2 < 1$ и удовлетворяющую на окружности этого круга условию

$$\left[\frac{\partial u}{\partial \rho} + \sigma(\theta)\, u\right]_{\rho=1} = f(\theta), \tag{39.3}$$

где ρ и θ — полярные координаты точки $(x,\ y)$, а $\sigma(\theta)$ — неотрицательная функция, отличная от тождественного нуля. Решение будем искать в виде потенциала простого слоя

$$u(x,\ y) = \int_0^{2\pi} \mu(\theta') \ln\frac{1}{r}\, d\theta', \tag{39.4}$$

где r — расстояние между точками $(\rho,\ \theta)$ и $(1,\ \theta')$. Подстановка в условие (39.3) дает интегральное уравнение

$$\mu(\theta) - \frac{1}{2\pi}\int_0^{2\pi} \mu(\theta')\, d\theta' + \frac{\sigma(\theta)}{\pi}\int_0^{2\pi} \mu(\theta') \ln\frac{1}{r}\, d\theta' = \frac{1}{\pi} f(\theta). \tag{39.5}$$

Система функций

$$\begin{matrix}\sin\\ \cos\end{matrix}\, n\theta \qquad (n = 0,\ 1,\ 2,\ \ldots)$$

ортонормирована и полна в $L_2(\theta, 2\pi)$. Вычислим интегралы

$$\frac{1}{\pi}\int_0^{2\pi} \cos n\theta' \ln\frac{1}{r}\,d\theta', \qquad \frac{1}{\pi}\int_0^{2\pi} \sin n\theta' \ln\frac{1}{r}\,d\theta'.$$

В уравнении (39.5) r есть расстояние между точками $(1, \theta')$ и $(1, \theta)$, поэтому $r = 2\left|\sin\frac{\theta-\theta'}{2}\right|$ и

$$\ln\frac{1}{r} = \ln\frac{1}{2\left|\sin\frac{\theta'-\theta}{2}\right|} = \sum_{k=1}^{\infty}\frac{\cos k(\theta'-\theta)}{k}.$$

Отсюда сразу находим:

$$\frac{1}{\pi}\int_0^{2\pi} \ln\frac{1}{r}\,d\theta' = 0,$$

$$\frac{1}{\pi}\int_0^{2\pi} \cos n\theta' \ln\frac{1}{r}\,d\theta' = \frac{\cos n\theta}{2n},$$

$$\frac{1}{\pi}\int_0^{2\pi} \sin n\theta' \ln\frac{1}{r}\,d\theta' = \frac{\sin n\theta}{2n}.$$

Теперь нахождение элементов матрицы и свободных членов системы Бубнова — Галёркина для уравнения (39.5) сводится к вычислению однократных интегралов.

Замечание. К уравнению (39.2) можно применить и метод наименьших квадратов (ВМ, § 84) Как и выше, достаточно взять координатную систему ортонормированной (даже только почти ортонормированной) в пространстве H; если оператор $(I+T)^{-1}$ существует и ограничен, то координатная система будет почти ортонормированной и в метрике

$$|u|^2 = \|u + Tu\|^2.$$

В этом случае процесс построения коэффициентов приближенного решения, а также процесс построения самого приближенного решения будет устойчивым, а матрица, к которой приводит метод наименьших квадратов, будет иметь ограниченное число обусловленности. Это замечание относится и к сингулярным интегральным уравнениям, речь о которых будет идти ниже.

2. Одномерные сингулярные уравнения. Ограничимся случаем замкнутого контура интегрирования.

Рассмотрим сингулярное уравнение вида

$$Au = a(t)\,u(t) + \frac{b(t)}{\pi i}\int_{\Gamma}\frac{u(\tau)}{\tau - t}\,d\tau + \int_{\Gamma} K(t, \tau)\,u(\tau)\,d\tau = f(t), \quad t \in \Gamma. \tag{39.6}$$

Здесь Γ — контур, ограничивающий область, одно- или многосвязную, на плоскости комплексной переменной. Будем считать, что Γ есть объединение простых замкнутых ляпуновских кривых Γ_0, $\Gamma_1, \ldots, \Gamma_m$, из которых Γ_0 ограничивает упомянутую область извне (если эта область — бесконечная, то кривая Γ_0 отсутствует), а остальные кривые — изнутри. Далее допустим, что коэффициенты $a(t)$ и $b(t)$ непрерывны на Γ, а ядро $K(t, \tau)$ и свободный член $f(t)$ суммируемы с квадратом

$$\int_{\Gamma}\int_{\Gamma} |K(t, \tau)|^2\,|dt|\,|d\tau| < \infty, \qquad \int_{\Gamma} |f(t)|^2\,|dt| < \infty.$$

Наконец, потребуем, чтобы

$$a^2(t) - b^2(t) \neq 0, \qquad t \in \Gamma. \tag{39.7}$$

Перечислим некоторые факты, на которые мы будем опираться[1]: 1) оператор A в левой части уравнения (39.1) ограничен в пространстве $L_2(\Gamma)$; 2) каждое из уравнений $Au = 0$ и $A^*v = 0$ имеет только конечное число линейно независимых решений; 3) оператор A нормально разрешим, иначе говоря, уравнение (39.6) имеет решение тогда и только тогда, когда функция $f(t)$ ортогональна ко всем решениям уравнения $A^*v = 0$.

Будем считать, что $f(t)$ удовлетворяет упомянутым условиям ортогональности. Допустим сначала, что уравнение $Au = 0$ имеет только тривиальное решение. Тогда решение уравнения (39.6) существует и единственно, и это решение можно приближенно построить, применяя процесс Ритца к функционалу метода наименьших квадратов.

Операторы A и A^{-1} ограничены, поэтому любая система, сильно минимальная или почти ортонормированная в $L_2(\Gamma)$, будет обладать тем же свойством и в метрике

$$|u|_A = \|Au\|. \tag{39.8}$$

Пространство $L_2(\Gamma)$ есть ортогональная сумма подпространств функций, отличных от нуля только на одном из контуров Γ_j; такое подпространство естественным образом можно отождествить с пространством $L_2(\Gamma_j)$. Отсюда следует, что полную почти ортонормированную (соответственно, сильно минимальную) систему в $L_2(\Gamma)$ можно получить, объединив такого рода системы для отдельных

[1]) Подробнее об одномерных сингулярных уравнениях в пространстве $L_2(\Gamma)$ см. статью автора [1].

пространств $L_2(\Gamma_j)$. Так, можно объединить системы $\{\omega_{jk}(t)\}$ $(0 \leqslant j \leqslant m,\ -\infty < k < \infty)$,

$$\omega_{jk}(t) = \frac{(t-a_j)^k}{\rho_j^k},$$

где ρ_j имеет то же значение, что и в § 38. Здесь a_j $(j = 1, 2, \ldots, m)$ — фиксированная точка внутри Γ_j. Указанное объединение приведет нас к системе функций

$$\varphi_{jk}(t) = \begin{cases} \omega_{jk}(t), & t \in \Gamma_j, \\ 0, & t \in \Gamma_l,\ l \neq j. \end{cases}$$

Система $\{\varphi_{jk}(t)\}$ почти ортонормирована в метрике (39.8).

Важно отметить, что сингулярные интегралы

$$\frac{1}{\pi i} \int\limits_{\Gamma} \frac{\varphi_{jk}(\tau)}{\tau - t}\, d\tau$$

просто вычисляются. Действительно,

$$\frac{1}{\pi i} \int\limits_{\Gamma} \frac{\varphi_{jk}(\tau)}{\tau - t}\, d\tau = \frac{1}{\pi i} \int\limits_{\Gamma_j} \frac{\omega_{jk}(\tau)}{\tau - t}\, d\tau.$$

Пусть сначала $j = 0$. Обозначая через z произвольную точку внутри Γ_0, имеем по интегральной формуле Коши

$$\frac{1}{2\pi i} \int\limits_{\Gamma_0} \frac{\omega_{0k}(\tau)}{\tau - z}\, d\tau = \begin{cases} \omega_{0k}(z), & k \geqslant 0, \\ 0, & k < 0. \end{cases}$$

Полагая $z \to t$, $t \in \Gamma_0$, и пользуясь известной теоремой о предельных значениях интеграла типа Коши, получим:

$$\frac{1}{\pi i} \int\limits_{\Gamma_0} \frac{\omega_{0k}(\tau)}{\tau - t}\, d\tau = \begin{cases} \omega_{0k}(t), & t \in \Gamma_0, \quad k > 0, \\ 0, & t \in \Gamma_0, \quad k < 0. \end{cases}$$

Аналогично найдем

$$\frac{1}{\pi i} \int\limits_{\Gamma_j} \frac{\omega_{jk}(\tau)}{\tau - t}\, d\tau = \begin{cases} -\omega_{jk}(t), & t \in \Gamma_j \quad (j = 1, 2, \ldots, m), \quad k < 0, \\ 0, & \qquad\qquad\qquad\qquad\qquad\qquad k \geqslant 0. \end{cases}$$

Уже упомянутая интегральная формула Коши дает:

$$\frac{1}{\pi i} \int\limits_{\Gamma_0} \frac{\omega_{0k}(\tau)}{\tau - t}\, d\tau = \begin{cases} 2\omega_{0k}(t), & t \in \Gamma_l, \quad l \neq 0, \quad k \geqslant 0, \\ 0, & t \in \Gamma_l, \quad l \neq 0, \quad k < 0, \end{cases}$$

$$\frac{1}{\pi i} \int\limits_{\Gamma_j} \frac{\omega_{jk}(\tau)}{\tau - t}\, d\tau = \begin{cases} -2\omega_{jk}(t), & t \in \Gamma_l, \quad l \neq j, \quad k < 0, \\ 0, & t \in \Gamma_l, \quad l \neq j, \quad k \geqslant 0. \end{cases}$$

Если уравнение $Au = 0$ имеет нетривиальные решения и эти решения известны, то можно применять метод наименьших квадратов так, как это указано в ВМ, § 84. Если решения уравнения $Au = 0$ неизвестны, то их можно приближенно определить процессом Ритца как собственные функции самосопряженного оператора A^*A, соответствующие собственному числу $\lambda = 0$.

3. Многомерные сингулярные интегральные уравнения. Будем рассматривать уравнение[1]) с неизвестной скалярной функцией $u(x)$

$$Au = a(x)u(x) + \int\limits_{E_m} \frac{f(x, \theta)}{r^m} u(y)\,dy + Tu = g(x). \qquad (39.9)$$

Здесь E_m — евклидово пространство m измерений, x и y — точки этого пространства $r = |y - x|$, $\theta = \frac{y-x}{r}$. Для упрощения формулировок будем считать, что «характеристика» $f(x, \theta)$ и ее производные достаточно высокого порядка по декартовым координатам точки θ равномерно непрерывны, когда θ пробегает единичную сферу, а x — пространство E_m с присоединенной к нему бесконечно удаленной точкой («расширенное евклидово пространство E_m»). В том же расширенном пространстве E_m будем считать непрерывным и коэффициент $a(x)$. Далее, пусть $g(x) \in L_2(E_m)$, и пусть оператор T вполне непрерывен в том же пространстве E_m. Наконец, потребуем, чтобы символ уравнения (39.9) нигде не обращался в нуль. При перечисленных условиях оператор в левой части уравнения (39.9) ограничен в $L_2(E_m)$, и для этого уравнения верны основные теоремы теории Фредгольма.

Будем рассматривать далее случай, когда однородное уравнение

$$a(x)u(x) + \int\limits_{E_m} \frac{f(x, \theta)}{r^m} u(y)\,dy + Tu = 0 \qquad (39.10)$$

имеет только тривиальное решение $u(x) \equiv 0$. Тогда уравнение (39.9) всегда разрешимо, и притом единственным образом, и оператор A^{-1} ограничен.

Уравнение (39.9) можно приближенно решить по методу наименьших квадратов; соответствующий вычислительный процесс будет устойчив, если координатная система будет сильно минимальна в $L_2(E_m)$. Еще лучше, если она будет ортонормирована, или хотя бы почти ортонормирована, в $L_2(E_m)$.

Весьма важно было бы так подобрать координатные функции $\varphi_n(x)$, чтобы сингулярные интегралы

$$\psi_n(x) = \int\limits_{E_m} \frac{f(x, \theta)}{r^m} \varphi_n(y)\,dy \qquad (39.11)$$

---

[1]) Необходимые сведения по теории многомерных сингулярных интегральных уравнений даны в книге автора [20].

просто вычислялись. Это можно сделать, если характеристика (а следовательно, и символ) сингулярного интеграла в (39.9) есть сферический полином. Для простоты ограничимся случаем $m=2$. Пусть характеристика имеет вид

$$f(x, \theta)=\sum_{k=-N}^{N}{}' a_k(x) e^{ik\theta}.$$

Тогда символ интеграла (39.11) равен

$$\Phi(x, \theta)=\sum_{k=-N}^{N}{}' b_k(x) e^{ik\theta}, \qquad b_k=\frac{2\pi i^{|k|}}{|k|} a_k. \tag{39.12}$$

Пусть сначала $N=1$; в этом случае символ (39.12) приводится к виду

$$\Phi(x, \theta)=\beta_1(x)\cos\theta+\beta_2(x)\sin\theta,$$

а характеристика — к виду

$$f(x, \theta)=\alpha_1(x)\cos\theta+\alpha_2(x)\sin\theta, \quad \alpha_j(x)=\frac{\beta_j(x)}{2\pi i} \qquad (j=1,\ 2).$$

Интеграл (39.11) можно тогда записать в такой форме:

$$\psi_n(x)=\alpha_1(x)\frac{\partial}{\partial x_1}\int\limits_{E_2}\frac{\varphi_n(y)}{r}\,dy+\alpha_2(x)\frac{\partial}{\partial x_2}\int\limits_{E_2}\frac{\varphi_n(y)}{r}\,dy, \tag{39.13}$$

где x_1 и x_2 — декартовы координаты точки x.

Введем в рассмотрение полупространство

$$-\infty<x_1<\infty, \quad -\infty<x_2<\infty, \quad 0<x_3<\infty$$

трехмерного пространства E_3; обозначим это пространство через E_3^+. Выберем последовательность функций $\{v_n(x_1, x_2, x_3)\}$, гармонических в E_3^+, равных нулю на бесконечности и непрерывных вместе с их первыми производными в полупространстве E_3^+ вместе с его границей. Положим

$$\varphi_n(y)=\left.\frac{\partial v_n(y_1, y_2, y_3)}{\partial y_3}\right|_{y_3=0}, \tag{39.14}$$

где y_1 и y_2 суть декартовы координаты точки y, и вычислим интеграл (39.13).

На границе $x_3=0$ полупространства E_3^+ функция $v_n(x_1, x_2, x_3)$ удовлетворяет условию задачи Неймана

$$\frac{\partial v_n}{\partial \nu}=-\varphi_n(x),$$

где ν — внешняя нормаль к плоскости $x_3 = 0$ — направлена против оси x_3. Используя известное решение задачи Неймана для полупространства, найдем:

$$v_n(x_1, x_2, x_3) = \frac{1}{2\pi} \int\limits_{E_2} \frac{\varphi_n(y)}{R}\, dy,$$

$$R^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2 + x_3^2.$$

Отсюда легко следует, что

$$\left.\begin{aligned} \frac{\partial}{\partial x_1} \int\limits_{E_2} \frac{\varphi_n(y)}{r}\, dy &= 2\pi \frac{\partial v_n(x_1, x_2, 0)}{\partial x_1}, \\ \frac{\partial}{\partial x_2} \int\limits_{E_2} \frac{\varphi_n(y)}{r}\, dy &= 2\pi \frac{\partial v_n(x_1, x_2, 0)}{\partial x_2}, \end{aligned}\right\} \tag{39.15}$$

и интеграл (39.13) вычислен.

Перейдем к общему случаю полиномиальной характеристики. Положим

$$hu = \frac{1}{2\pi i} \int\limits_{E_2} \frac{e^{i\theta}}{r^2}\, u(y)\, dy.$$

Тогда, как это следует из формулы (39.12), интеграл (39.11) можно представить в форме

$$\int\limits_{E_2} \frac{f(x, \theta)}{r^2}\, u(y)\, dy = \sum_{k=-N}^{N} b_k(x)\, h^k u,$$

и дело сводится к такому подбору функций $\varphi_n(x)$, чтобы выражения $h^k\varphi_n$ легко вычислялись. Как и выше, введем последовательность функций $v_n(x_1, x_2, x_3)$, гармонических в E_3^+ и равных нулю на бесконечности; на этот раз потребуем только, чтобы в замкнутом полупространстве $x_3 \geqslant 0$ эти функции были N раз непрерывно дифференцируемы. Функции $\varphi_n(x)$ определим по-прежнему формулой (39.14).

Положим

$$w_{n1}(x_1, x_2, x_3) = 2\pi \left[\frac{\partial v_n}{\partial x_1} + i \frac{\partial v_n}{\partial x_2}\right].$$

Очевидно, функция w_{n1} гармонична в E_3^+ и $h\varphi_n = w_{n1}(x_1, x_2, 0)$. Построим гармоническую в E_3^+ функцию $\omega_{n1}(x_1, x_2, x_3)$ такую, что $\omega_{n1} = 0$ на бесконечности и

$$\left.\frac{\partial \omega_{n1}}{\partial x_3}\right|_{x_3=0} = w_{n1}(x_1, x_2, 0).$$

Легко видеть, что

$$\omega_{n1} = \int_{+\infty}^{x_3} w_{n1}(x_1, x_2, \xi)\, d\xi.$$

Теперь аналогично предыдущему

$$h^2\varphi_n = h w_{n1}(x_1, x_2, 0) = 2\pi \left[\frac{\partial \omega_{n1}}{\partial x_1} + i \frac{\partial \omega_{n1}}{\partial x_2}\right]_{x_3=0}.$$

Полагаем теперь

$$w_{n2}(x_1, x_2, x_3) = 2\pi \left[\frac{\partial \omega_{n1}}{\partial x_1} + i \frac{\partial \omega_{n2}}{\partial x_2}\right]$$

и повторяем процесс.

В качестве функций v_n можно взять, например, преобразование Кели гармонических полиномов.

Распространение приема на многократные сингулярные интегралы очевидно.

ГЛАВА VI

СЛУЧАЙ БЕСКОНЕЧНОЙ ОБЛАСТИ И ДРУГИЕ СИНГУЛЯРНЫЕ ЗАДАЧИ

§ 40. Предварительные замечания

Если область Ω, в которой ищется решение краевой задачи, бесконечна, то в обычно встречающихся случаях, таких, например, как случай задачи Дирихле для уравнения Пуассона, соответствующий оператор часто оказывается не положительно определенным, а только положительным; в некоторых случаях оператор остается положительно определенным, но его спектр перестает быть дискретным. Спектр только положительного оператора также недискретен (см. ниже). Напомним, что в построениях гл. V мы широко использовали дискретность спектров изучаемых операторов; потеря этого свойства естественным образом затрудняет подбор координатной системы.

Свойство дискретности спектра оператора или его положительной определенности может потеряться и тогда, когда область Ω конечна, но дифференциальное уравнение, будучи эллиптическим внутри Ω, вырождается на всей границе или на ее части.

В настоящей главе мы и будем заниматься построением координатных систем для операторов с недискретным спектром. Одновременно мы подробнее изучим задачи вида

$$Au = f, \qquad f \in H, \tag{40.1}$$

где оператор A предполагается только положительным, но не положительно определенным, в данном гильбертовом пространстве H. На множестве $D(A)$ — области определения оператора A — выполняется неравенство

$$(Au, u) > 0, \quad u \neq 0, \tag{40.2}$$

но при этом

$$\inf_{u \in D(A)} \frac{(Au, u)}{(u, u)} = 0. \tag{40.3}$$

Положительный оператор, очевидно, полуограничен снизу и потому допускает самосопряженное расширение по Фридрихсу. Может случиться при этом, что расширенный оператор (обозначим его $\tilde{A}$) потеряет свойство положительности и будет удовлетворять не неравенству (40 2), а более слабому соотношению $(\tilde{A}u, u) \geqslant 0$, причем знак равенства будет достигаться хотя бы для одного элемента[1]) $u_0 \neq 0$. Этот случай мы исключим из рассмотрения и будем впредь требовать, чтобы расширение по Фридрихсу положительного оператора также было положительным. Ниже будем считать, что в уравнении (40.1) оператор A — самосопряженный.

Как и в случае оператора, положительно определенного, можно ввести энергетические произведение и норму[2])

$$[u, v] = [u, v]_A = (Au, v), \quad |u|^2 = |u|_A^2 = (Au, u) \tag{40.4}$$

и построить энергетическое пространство H_A как замыкание множества $D(A)$ в метрике (40.4). Если оператор A только положителен, то H_A не вкладывается в H — среди идеальных элементов пространства H_A имеются и такие, которым не соответствуют никакие элементы пространства H. Пусть оператор A положительный. Тогда уравнение (40.1) имеет в H не более одного решения; если оно разрешимо в H, то его решение реализует минимум функционала

$$F(u) = |u|_A^2 - (f, u) - (u, f) \tag{40.5}$$

на множестве $D(A)$; обратно, если функционал $F(u)$ достигает минимума на множестве $D(A)$, то элемент, реализующий этот минимум, удовлетворяет[3]) уравнению (40.1).

Элемент $u_0 \in H_A$, реализующий минимум функционала (40.5) в пространстве H_A (если этот элемент существует), в общем случае не принадлежит исходному пространству H, но для этого элемента величина $|u_0|_A^2$ конечна. С физической точки зрения это можно истолковать так: состояние, описываемое элементом u_0, имеет конечную энергию. Такое толкование дает повод назвать элемент u_0 ***решением с конечной энергией*** для уравнения (40.1) Из сказанного в § 6 вытекает[4]), что уравнение (40.1) имеет решение с конечной энергией тогда и только тогда, когда скалярное произведение (u, f) есть ограниченный в метрике H_A функционал.

---

[1]) В этом случае $\tilde{A}u_0 = 0$. Действительно, пусть v — произвольный элемент области $D(\tilde{A})$. Тогда при любом вещественном ε

$$(\tilde{A}(u_0 + \varepsilon v), u_0 + \varepsilon v) = (\tilde{A}u_0, u_0) + 2\varepsilon \operatorname{Re}(\tilde{A}u_0, v) + \varepsilon^2 (\tilde{A}v, v) \geqslant 0$$

и так как $(\tilde{A}u_0, u_0) = 0$, то необходимо $\operatorname{Re}(\tilde{A}u_0, v) = 0$. Заменив v через iv, найдем, что и $\operatorname{Im}(\tilde{A}u_0, v) = 0$ и, следовательно, $(\tilde{A}u_0, v) = 0$. Будучи ортогональным к плотному множеству $D(\tilde{A})$, элемент $\tilde{A}u_0 = 0$.

[2]) См. ПМ, § 6.

[3]) Доказательство перечисленных выше утверждений см. ПМ, § 6.

[4]) Подробнее об этом см. ПМ, § 6.

Можно рассматривать более общую вариационную задачу о минимуме функционала

$$F(u) = |u|_A^2 - lu - \overline{lu}, \qquad u \in H_A, \tag{40.5'}$$

где lu — линейный функционал. Оказывается также, что эта новая задача имеет решение (которое мы, по-прежнему, назовем решением с конечной энергией) тогда и только тогда, когда функционал lu ограничен в метрике H_A.

Часто оказывается полезной такая постановка задачи. Пусть линейный функционал lu определен первоначально не на всем пространстве H_A, а только на некотором его плотном множестве (например, на области определения оператора A). Если на этом множестве функционал l ограничен в метрике H_A, то этот функционал расширяется по непрерывности до некоторого нового функционала l_1, также ограниченного в H_A и определенного на всем этом пространстве. Задача о минимуме функционала

$$F_1(u) = |u|_A^2 - l_1 u - \overline{l_1 u}, \qquad u \in H_A,$$

имеет решение, которое мы и будем трактовать как решение с конечной энергией для вариационной задачи (40.5′).

Если энергетическое пространство H_A сепарабельно, а уравнение (40.1) имеет решение с конечной энергией, то это решение можно построить с помощью процесса Ритца. Докажем, что если сепарабельно исходное пространство H, то энергетическое пространство H_A также сепарабельно. Оператор $B = A + I$, где I — тождественный оператор, самосопряжен и положительно определен:

$$|u|_B^2 = (Bu, u) = |u|_A^2 + \|u\|^2 \geqslant \|u\|^2 \qquad u \in D(B) = D(A).$$

В таком случае пространство H_B сепарабельно [1]). А тогда в метрике этого пространства сепарабельно плотное в нем множество $D(B) = D(A)$: существует последовательность $v_k \in D(A)$ $(k = 1, 2, \ldots)$ такая, что каковы бы ни были элемент $u \in D(A)$ и число $\varepsilon > 0$, найдется элемент v_k, удовлетворяющий неравенству $|u - v_k|_B < \varepsilon$. Тем более, $|u - v_k|_A \leqslant |u - v_k|_B < \varepsilon$; множество $D(A)$ сепарабельно и в метрике H_A. Но это множество плотно в пространстве H_A, поэтому если w — произвольный элемент из H_A, то можно найти элемент $u \in D(A)$ такой, что $|w - u|_A < \varepsilon$, и элемент v_k такой, что $|u - v_k|_A < \varepsilon$. Теперь $|w - v_k|_A < 2\varepsilon$, и пространство H_A сепарабельно.

Отметим хорошо известную, но важную особенность самосопряженных только положительных [2]) операторов: их спектры недискретны.

---

[1]) См. ПМ, § 8.

[2]) Мы называем оператор «только положительным», если он положительный, но не положительно определенный.

Действительно, пусть спектр самосопряженного только положительного оператора A дискретен. Тогда если λ_1 — наименьшее собственное число этого оператора, то, по формуле (40.3),

$$\lambda_1 = \inf_{u \in D(A)} \frac{(Au, u)}{\|u\|^2} = 0.$$

Обозначим через u_1 нормированный собственный элемент, отвечающий собственному числу $\lambda_1 = 0$. Тогда

$$(Au_1, u_1) = 0, \qquad \|u_1\| = 1,$$

что противоречит неравенству (40.2).

Если A и K — самосопряженные операторы, причем A — только положительный, а K — ограниченный оператор, то сумма $A+K$ может оказаться положительно определенной (например, при $K=I$), но спектр такой суммы обязательно недискретен. Это обстоятельство затрудняет выбор координатных элементов для задач, содержащих операторы указанного вида.

§ 41. Эллиптические уравнения второго порядка в бесконечной области

Пусть Ω — бесконечная область в пространстве E_m декартовых координат $x_1, x_2, \ldots, x_m$, ограниченная кусочно гладкой поверхностью S. Нас будут в основном интересовать два типа бесконечных областей: 1) области первого типа, содержащие куб со сколь угодно длинным ребром; 2) области второго типа, которые отличаются следующим свойством: поверхность, ограничивающую такую область, можно заключить между двумя бесконечными или полубесконечными призмами, каждая из которых в подходящим образом выбранной системе координат может быть описана неравенствами (случай бесконечных призм)

$$\begin{aligned} &0 \leqslant x_k \leqslant a, \quad k < m, \quad -\infty < x_m < +\infty; \\ &0 \leqslant x_k \leqslant b, \quad k < m, \quad -\infty < x_m < +\infty, \quad a > b, \end{aligned}$$

или (случай полубесконечных призм)

$$\begin{aligned} &0 \leqslant x_k \leqslant a, \quad k < m, \quad x_m \geqslant \alpha; \\ &0 \leqslant x_k \leqslant b, \quad k < m, \quad x_m \geqslant \beta, \quad a > b, \quad \beta > \alpha. \end{aligned}$$

К первому типу относятся области, лежащие вне конечных границ, а также области, лежащие внутри или вне параболоида, гиперболоида, конуса, внешность кругового цилиндра и т. д. Ко второму типу относится, например, внутренность кругового бесконечного или полубесконечного цилиндра.

В пространстве $L_2(\Omega)$, где Ω — бесконечная область, рассмотрим дифференциальный оператор

$$Au = -\sum_{j,\, k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}(x)\frac{\partial u}{\partial x_k}\right), \qquad A_{jk} = \overline{A}_{kj} \tag{41.1}$$

при одном из обычных краевых условий

$$u|_S = 0 \tag{41.2}$$

или

$$\left[\sum_{j,\, k=1}^{m} A_{jk}\frac{\partial u}{\partial x_j}\cos(\nu,\ x_k) + \sigma u\right]_S = 0, \qquad \sigma \geqslant 0; \tag{41.3}$$

смысл обозначений тот же, что и в предшествующих главах. Будем считать, что собственные числа матрицы коэффициентов A_{jk} положительно ограничены сверху и снизу, так что

$$\mu_0 \sum_{k=1}^{m} |t_k|^2 \leqslant \sum_{j,\, k=1}^{m} A_{jk}(x)\overline{t}_j t_k \leqslant M_0 \sum_{k=1}^{m} |t_k|^2, \tag{41.4}$$

где μ_0 и M_0 — положительные постоянные. Примем еще, что $A_{jk} \in C^1(\Omega)$.

За область $D(A)$ определения оператора A примем множество финитных функций из $C^2(\overline{\Omega})$, которые удовлетворяют соответствующему краевому условию (41.2) или (41.3); в соответствии с общепринятым определением мы называем функцию *финитной*, если она гождественно равна нулю вне некоторой сферы.

Нетрудно видеть, что так определенный оператор A положителен. Действительно, так как функция $u \in D(A)$ финитна, то в формуле

$$(Au,\ u) = -\int_{\Omega} \overline{u} \sum_{j,\, k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right) dx$$

интегрировать надо только по конечной области, и можно применить формулу интегрирования по частям; учтя граничное условие, получим

$$(Au,\ u) = \int_{\Omega} \sum_{j,\, k=1}^{m} A_{jk}\frac{\partial \overline{u}}{\partial x_j}\frac{\partial u}{\partial x_k}\, dx \tag{41.5}$$

в случае условия (41.2) и

$$(Au,\ u) = \int_{\Omega} \sum_{j,\, k=1}^{m} A_{jk}\frac{\partial \overline{u}}{\partial x_j}\frac{\partial u}{\partial x_k}\, dx + \int_{S} \sigma |u|^2 dS \tag{41.6}$$

в случае условия (41.3). Отсюда ясно, что $(Au,\ u) \geqslant 0$. При этом, если $(Au,\ u) = 0$, то, очевидно, $u(x) \equiv C = \text{const}$; будучи равной нулю вне некоторой сферы, постоянная $C = 0$.

Замкнув область $D(A)$ в метрике, порождаемой интегралом (41.5) или (41.6), мы получим соответствующее пространство H_A. Нетрудно доказать, что это пространство состоит из функций, которые: 1) имеют обобщенные первые производные, суммируемые с квадратом в Ω, 2) допускают в метрике интеграла Дирихле аппроксимацию финитными непрерывно дифференцируемыми функциями; таким образом, если $u \in H_A$, то существует последовательность финитных непрерывно дифференцируемых в $\bar{\Omega}$ функций $u_n(x)$ таких, что

$$\lim_{n\to\infty} \int_\Omega |\operatorname{grad}(u-u_n)|^2\,dx = 0;$$

3) в случае краевого условия (41.2) функции из H_A обращаются в нуль на S, а в случае условия (41.3) они квадратично суммируемы на S с весом σ. Из теоремы вложения С. Л. Соболева вытекает, что в любой конечной подобласти Ω и на любой конечной измеримой части поверхности S эти функции квадратично суммируемы.

Энергетические произведение и норма определяются формулами

$$\left.\begin{aligned} [u, v]_A &= \int_\Omega \sum_{j,k=1}^m A_{jk} \frac{\partial \bar{v}}{\partial x_j} \frac{\partial u}{\partial x_k}\,dx, \\ |u|_A^2 &= \int_\Omega \sum_{j,k=1}^m A_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k}\,dx \end{aligned}\right\} \tag{41.7}$$

в случае условия (41.2), и формулами

$$\left.\begin{aligned} [u, v]_A &= \int_\Omega \sum_{j,k=1}^m A_{jk} \frac{\partial \bar{v}}{\partial x_j} \frac{\partial u}{\partial x_k}\,dx + \int_S \sigma u \bar{v}\,dS, \\ |u|_A^2 &= \int_\Omega \sum_{j,k=1}^m A_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k}\,dx + \int_S \sigma |u|^2\,dS \end{aligned}\right\} \tag{41.8}$$

в случае условия (41.3).

Теорема 41.1. *Оператор A — только положительный, если Ω — область первого типа, и положительно определенный, если Ω — область второго типа, а краевое условие имеет вид* (41.2).

Пусть Ω — область первого типа. Поместим в нее куб Q со стороной a и направим оси координат по ребрам куба так, чтобы внутри него все координаты были положительны. Функция

$$u_a(x) = \begin{cases} \prod_{k=1}^m \sin^3 \dfrac{\pi x_k}{a}, & x \in Q, \\ 0, & x \bar{\in} Q, \end{cases}$$

принадлежит области $D(A)$ определения оператора A. В силу неравенства (41.4).

$$\frac{(Au_a, u_a)}{\| u_a \|^2} = \frac{| u_a |_A^2}{\| u_a \|^2} \leqslant \frac{M_0}{\| u_a \|^2} \int\limits_\Omega \sum_{k=1}^m \left(\frac{\partial u_a}{\partial x_k} \right)^2 dx =$$

$$= \frac{M_0}{\| u_a \|^2} \int\limits_Q \sum_{k=1}^m \left(\frac{\partial u_a}{\partial x_k} \right)^2 dx = \frac{c}{a^2} \underset{a \to \infty}{\longrightarrow} 0;$$

здесь c — некоторая постоянная. Отсюда следует, что

$$\inf_{u \in D(A)} \frac{(Au, u)}{\| u \|^2} = 0.$$

Не существует элемента $u_0 \neq 0$, для которого $(Au_0, u_0) = 0$: если бы такой элемент существовал, то было бы $| u_0 |_A = 0$. Из формулы (41.7) следует тогда, что $u_0 = \text{const}$, и так как $u_0 \in D(A)$ и потому одновременно $u_0 \in H$, а мера области Ω бесконечна, то $u_0 \equiv 0$. Отсюда следует, что A — только положительный оператор.

Пусть теперь Ω — область второго типа, а краевое условие имеет вид (41.2). Допустим, например, что ее можно поместить внутрь полубесконечной призмы

$$0 \leqslant x_k \leqslant a, \quad k < m; \quad 0 \leqslant x_m < \infty,$$

где число a на этот раз фиксировано; упомянутую призму обозначим через Q. Введем в рассмотрение оператор B, действующий в пространстве $L_2(Q)$ по формуле $Bu = -\Delta u$ и заданный на финитных функциях из $C^{(2)}(\overline{Q})$, которые обращаются в нуль на границе призмы Q. Любую функцию $u \in H_A$ можно рассматривать и как элемент пространства H_B: достаточно для этого доопределить функцию $u(x)$, положив ее равной нулю, если $x \in Q \setminus \Omega$. При этом, в силу неравенства (41.4),

$$| u |_A^2 \geqslant \mu_0 \int\limits_\Omega \sum_{k=1}^m \left| \frac{\partial u}{\partial x_k} \right|^2 dx = \mu_0 \int\limits_Q \sum_{k=1}^m \left| \frac{\partial u}{\partial x_k} \right|^2 dx = \mu_0 | u |_B^2,$$

и нам достаточно будет доказать положительную определенность оператора B.

Пусть $u(x)$ — произвольная функция из области определения оператора B. Эта функция финитна, поэтому если число h достаточно велико и $x_m > h$, то $u(x) \equiv 0$, и

$$\frac{| u |_B^2}{\| u \|^2} = \frac{\int\limits_0^a \cdots \int\limits_0^a \int\limits_0^h \sum\limits_{k=1}^m \left| \frac{\partial u}{\partial x_k} \right|^2 dx_1 \ldots dx_{m-1}\, dx_m}{\int\limits_0^a \cdots \int\limits_0^a \int\limits_0^h | u |^2 dx_1 \ldots dx_{m-1}\, dx_m} \geqslant \lambda_h,$$

где через λ_h мы обозначили наименьшее собственное число задачи Дирихле для оператора $-\Delta$ в случае параллелепипеда $0 \leqslant x_k \leqslant a$, $k < m$; $0 \leqslant x_m \leqslant h$. Как хорошо известно,

$$\lambda_h = \pi^2 \left(\frac{m-1}{a^2} + \frac{1}{h^2} \right).$$

Отсюда

$$| u |_B^2 \geqslant \frac{(m-1)\pi^2}{a^2} \| u \|^2,$$

и оператор B — положительно определенный. Таким же будет и оператор A, причем

$$| u |_A^2 \geqslant \frac{(m-1)\mu_0\pi^2}{a^2} \| u \|^2.$$

Если Ω — область второго типа, а краевое условие имеет вид (41.3), то оператор A во всяком случае положителен; это легко усмотреть из формулы (41.8).

В случае бесконечной области первого или второго типа спектр оператора A не является дискретным. Для области первого типа это сразу вытекает из того факта, что оператор A только положителен; для области второго типа доказательство сложнее и мы на нем не будем останавливаться[1]).

В заключение отметим случай обыкновенного дифференциального оператора.

Рассмотрим оператор

$$-\frac{d}{dx}\left(p(x) \frac{du}{dx} \right) \tag{41.9}$$

на одном из двух бесконечных промежутков $(-\infty, \infty)$ или $(0, \infty)$; примем, что в соответствующем промежутке функция $p(x)$ непрерывно дифференцируема и положительно ограничена снизу и сверху. Оператор (41.9) будем рассматривать соответственно в пространстве $L_2(-\infty, \infty)$ или $L_2(0, \infty)$; в последнем случае мы будем подчинять функцию $u(x)$ одному из условий

$$u(0) = 0 \tag{41.10}$$

или

$$u'(0) - \alpha u(0) = 0, \qquad \alpha \geqslant 0. \tag{41.11}$$

В данном случае Ω есть бесконечная область первого типа: одномерный куб с ребром длины a есть отрезок длины a, и как луч $(0, \infty)$, так и прямая $(-\infty, \infty)$ содержат отрезок произвольной длины. Из сказанного в настоящем параграфе следует, что рассматриваемый нами одномерный оператор только положителен и, следовательно, спектр его недискретен.

---

[1]) Это вытекает из весьма общих результатов, приведенных, например в книге И. М. Глазмана [1], стр. 233.

§ 42. Условие дивергенции[1])

Пусть Ω — бесконечная область первого или второго типа[2]) и пусть в этой области требуется проинтегрировать уравнение

$$-\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right)=f(x) \tag{42.1}$$

при краевом условии (41.2) или (41.3); допустим, что коэффициенты $A_{jk}(x)$ удовлетворяют условиям § 41. Если $f(x)\in L_2(\Omega)$, а поставленная здесь задача имеет решение, также принадлежащее $L_2(\Omega)$, то это решение реализует минимум функционала

$$|u|_A^2-2\,\mathrm{Re}\int_\Omega u\bar{f}\,dx, \tag{42.2}$$

где $|u|_A^2$ определено формулой (41.7) или (41.8). В общем случае мы будем искать для уравнения (42.1) решение с конечной энергией; из общих соображений нам известно[3]), что такое решение существует тогда и только тогда, когда интеграл

$$\int_\Omega u\bar{f}\,dx, \qquad u\in D(A), \tag{42.3}$$

есть ограниченный в метрике H_A функционал. Выясним, при каких условиях это имеет место. Обозначим через M множество финитных функций, непрерывно дифференцируемых в Ω и обращающихся в нуль в пограничной полосе (своей для каждой функции) этой области.

Пусть почти всюду в Ω определен локально суммируемый[4]) вектор $G(x)$; допустим, что существует определенная почти всюду в Ω и также локально суммируемая скалярная функция $g(x)$, обладающая тем свойством, что для любой функции $\varphi(x)\in M$ справедливо тождество

$$\int_\Omega \varphi(x)\overline{g(x)}\,dx=-\int_\Omega \operatorname{grad}\varphi\cdot G(x)\,dx. \tag{42.4}$$

Будем тогда называть функцию $g(x)$ дивергенцией (обобщенной) вектора $G(x)$ и будем писать $g(x)=\operatorname{div}G(x)$.

Понятие обобщенной дивергенции, разумеется, можно отнести и к случаю конечной области Ω. В этом случае мы будем считать,

---

[1]) Основные результаты настоящего параграфа содержатся в статьях автора [8] и [14].

[2]) Некоторые из результатов настоящего параграфа справедливы и для областей более общего вида.

[3]) См. § 40.

[4]) То есть суммируемый в любой конечной подобласти Ω.

что функции $g(x)$ и $|G(x)|$ суммируемы в любой внутренней подобласти Ω, а под M будем понимать множество функций из $C^{(1)}(\Omega)$, каждая из которых обращается в нуль в некоторой пограничной полосе области Ω. Будем говорить, что $g(x) = \operatorname{div} G(x)$, если справедливо тождество

$$\int_\Omega \varphi(x)\,\overline{g}(x)\,dx = -\int_\Omega \operatorname{grad}\varphi \cdot G(x)\,dx.$$

Теорема 42.1. *Для того чтобы задача* (42.1), (41.2) *имела решение с конечной энергией, необходимо и достаточно, чтобы* $f(x) = \operatorname{div} F(x)$, *где* $|F| \in L_2(\Omega)$.

Условие теоремы 42.1 будем называть «условием дивергенции». Если задача (42.1), (41.2) имеет решение $u_0(x)$ с конечной энергией, то интеграл (42.3) можно представить в виде[1])

$$\int_\Omega \overline{f}(x)\,u(x)\,dx = [u,\ u_0]_A = \int_\Omega \sum_{j,\,k=1}^{m} A_{jk}(x)\,\frac{\partial \overline{u_0}}{\partial x_j}\,\frac{\partial u}{\partial x_k}\,dx = -\int_\Omega \operatorname{grad} u \cdot F(x)\,dx, \tag{42.5}$$

где $F(x)$ есть вектор, k-я составляющая которого равна

$$F_k(x) = -\sum_{j=1}^{m} \overline{A}_{jk}\,\frac{\partial u_0}{\partial x_j} = -\sum_{j=1}^{m} A_{kj}\,\frac{\partial u_0}{\partial x_j}. \tag{42.6}$$

Множество M, определенное выше, очевидно, входит в пространство H_A. Будем в равенстве (42.5) понимать под u произвольную функцию из M, тогда это равенство означает, что $f(x) = \operatorname{div} F(x)$. Далее, из неравенства (41.4) следует, что коэффициенты A_{jk} ограничены. Пусть $|A_{kj}| \leqslant N$, тогда из формулы (42.6) в силу неравенства Коши следует:

$$|F(x)|^2 = \sum_{j=1}^{m} |F_j(x)|^2 \leqslant m^2N^2\,|\operatorname{grad} u_0|^2;$$

так как $u_0 \in H_A \subset W_2^{(1)}(\Omega)$, то $|F| \in L_2(\Omega)$. Обратно, если $f(x)$ удовлетворяет условиям теоремы, то для любой функции $u \in M$

$$\int_\Omega \overline{f}(x)\,u(x)\,dx = \int_\Omega u\,\overline{\operatorname{div} F}\,dx = -\int_\Omega \operatorname{grad} u \cdot F\,dx.$$

Отсюда

$$\left|\int_\Omega \overline{f}(x)\,u(x)\,dx\right| \leqslant \left\{\int_\Omega |F|^2\,dx\right\}^{\frac{1}{2}} \left\{\int_\Omega |\operatorname{grad} u|^2\,dx\right\}^{\frac{1}{2}}$$

---

[1]) См. § 41.

и, по неравенству (41.4),

$$\left|\int_\Omega \bar{f}(x)\, u(x)\, dx\right| \leqslant \frac{1}{\mu_0}\left\{\int_\Omega |F|^2\, dx\right\}^{\frac{1}{2}} |u|_A. \tag{42.7}$$

Множество M плотно в H_A, и последнее неравенство показывает, что функционал (42.3) ограничен в H_A и, следовательно, задача (42.1), (41.2) имеет решение с конечной энергией.

Если $u_0(x)$ — это решение, то

$$\int_\Omega \bar{f}(x)\, u(x)\, dx = [u,\ u_0]_A.$$

В неравенстве (42.7), которое верно не только для элементов множества M, но и для всего пространства H_A, положим $u = u_0$. Мы получим тогда оценку энергетической норми решения:

$$|u_0|_A \leqslant \frac{1}{\mu_0}\left\{\int_\Omega |F(x)|^2\, dx\right\}^{\frac{1}{2}}. \tag{42.8}$$

Теорема 42.2. *Пусть Ω — область первого типа с конечной границей. Для того чтобы задача* (42.1), (41.3) *имела решение с конечной энергией, необходимо, чтобы $f(x) = \operatorname{div} F(x)$, $|F| \in L_2(\Omega)$, и достаточно, чтобы, кроме того, для любой финитной функции $u \in H_A$ была верна формула интегрирования по частям*

$$\int_\Omega u\bar{f}\, dx = -\int_\Omega (\operatorname{grad} u)\cdot F\, dx + \int_S u\bar{F}_\nu\, dS, \tag{42.9}$$

где F_ν — проекция вектора F на направление внешней нормали ν к поверхности S, и чтобы $F_\nu \in L_2(S)$.

Напомним, что в данном случае пространство H_A состоит из тех функций пространства $W_2^{(1)}(\Omega)$, квадраты которых на границе S суммируемы с весом σ; метрика в H_A определяется формулами (41.8). Необходимость доказывается, как в теореме 42.1, следует только предварительно заметить, что и в данном случае $M \subset H_A$.

Займемся доказательством достаточности. Пусть финитная функция $u \in H_A$. Из тождества (42.9) следует:

$$\left|\int_\Omega u\bar{f}\, dx\right|^2 \leqslant 2\int_\Omega |F|^2\, dx \int_\Omega |\operatorname{grad} u|^2\, dx + 2\int_S |F_\nu|^2\, dS \int_S |u|^2\, dS. \tag{42.10}$$

Допустим сначала, что $\sigma \not\equiv 0$. Тогда из теорем вложения С. Л. Соболева вытекает, что

$$\int_S |u|^2\, dS \leqslant C_1\left\{\int_\Omega (\operatorname{grad} u)^2\, dx + \int_S \sigma |u|^2\, dS\right\}, \qquad C_1 = \text{const}.$$

Отсюда, как легко видеть, следует, что

$$\left|\int_\Omega u\bar{f}\,dx\right|^2 \leqslant C_2\left\{\int_\Omega (\operatorname{grad} u)^2\,dx + \int_S \sigma|u|^2\,dS\right\} = C_2|u|_A^2, \qquad (42.11)$$

$$C_2 = \text{const}.$$

Функционал (42.3) оказывается ограниченным на плотном в H_A множестве финитных функций; в таком случае задача (42.1), (41.3) имеет решение с конечной энергией. Из неравенства (42.10) легко вытекает оценка $|u_0|_A \leqslant \sqrt{C_2}$, где u_0 — только что упомянутое решение.

Случай $\sigma \equiv 0$ требует особого рассмотрения. В этом случае решение задачи (42.1), (41.3), если оно существует, определено только с точностью до постоянного слагаемого. Мы устраним этот произвол, заменив пространство H_A его подпространством $\tilde{H}_A$, для элементов которого

$$\int_S u\,dS = 0, \qquad (42.12)$$

и потребовав, чтобы решение с конечной энергией принадлежало подпространству $\tilde{H}_A$. Докажем, что для элементов этого подпространства имеет место неравенство

$$\int_S |u|^2\,dS \leqslant C_1\int_\Omega |\operatorname{grad} u|^2\,dx \leqslant \frac{C_1}{\mu_0}|u|_A^2, \qquad C_1 = \text{const}. \qquad (42.13)$$

Построим сферу S_R с центром в начале координат и радиусом R, фиксированным, но настолько большим, чтобы вся граница S лежала внутри сферы. Обозначим через Ω_R часть области Ω, заключенную внутри S_R. В пространстве $W_2^{(1)}(\Omega_R)$ можно задать норму формулой

$$\|u\|^2_{W_2^{(1)}(\Omega_R)} = \int_{\Omega_R} |\operatorname{grad} u|^2\,dx + \left|\int_S u\,dS\right|^2,$$

а тогда из теорем вложения вытекает неравенство

$$\int_S |u|^2\,dS \leqslant C_1\left\{\int_{\Omega_R} |\operatorname{grad} u|^2\,dx + \left|\int_S u\,dS\right|^2\right\}, \qquad C_1 = \text{const}.$$

Если $u \in \tilde{H}_A$, то из этого неравенства получается:

$$\int_S |u|^2\,dS \leqslant C_1\int_{\Omega_R} |\operatorname{grad} u|^2\,dx \leqslant C_1\int_\Omega |\operatorname{grad} u|^2\,dx.$$

Из неравенств (42.10) и (42.13) следует теперь, что

$$\left|\int_\Omega u\bar{f}\,dx\right| \leqslant 2\left[\int_\Omega |F(x)|^2\,dx + \frac{C_1}{\mu_0}\int_S |F_\nu|^2\,dS\right]|u|_A^2;$$

функционал (42.3) ограничен в $\tilde{H}_A$, и наша задача имеет решение с конечной энергией. Пусть u_0 — это решение. Оно реализует минимум функционала

$$|u|_A^2 - 2\,\mathrm{Re}\int_\Omega u\bar{f}\,dx$$

в пространстве $\tilde{H}_A$. Тогда в точке u_0 вариация этого функционала равна нулю:

$$[u_0, \zeta]_A - \int_\Omega \zeta\bar{f}\,dx = 0, \qquad \zeta \in \tilde{H}_A,$$

или

$$[u_0, \zeta]_A + \int_\Omega \operatorname{grad}\zeta \cdot F\,dx = 0.$$

Полагая здесь $\zeta = u_0$ и оценивая интеграл по неравенству Буняковского и неравенству (41.4), получим:

$$|u_0|_A^2 \leqslant \left\{\int_\Omega |F|^2\,dx\right\}^{\frac{1}{2}} \left\{\int_\Omega |\operatorname{grad} u_0|^2\,dx\right\}^{\frac{1}{2}} \leqslant$$
$$\leqslant \frac{1}{\sqrt{\mu_0}}\left\{\int_\Omega |F(x)|^2\,dx\right\}^{\frac{1}{2}} |u_0|_A.$$

Это дает нам оценку решения с конечной энергией для задачи Неймана:

$$|u_0|_A \leqslant \frac{1}{\sqrt{\mu_0}}\left\{\int_\Omega |F|^2\,dx\right\}^{\frac{1}{2}}. \tag{42.14}$$

З а м е ч а н и е 1. Формула (42.9) и включение $F_\nu \in L_2(S)$ имеют место, если, например, $F_j \in W_p^{(1)}(\omega)$, где F_j — составляющие вектора F, ω — пересечение области Ω с некоторой окрестностью границы S и $\frac{(m-1)p}{m-p} > 2$ или $p > \frac{2m}{m+1}$.

З а м е ч а н и е 2. Соотношения $f(x) = \operatorname{div} F(x)$, $F \in L_2(\Omega)$ имеют место, например, в следующем случае[1]) Пусть x и y — точки области Ω, $\rho = |y|$, $r = |x - y|$ и

$$\alpha(\rho) = \begin{cases} 0, & \rho < 1, \\ 1, & \rho \geqslant 1. \end{cases}$$

[1]) См. статью автора [14].

Если существует точка, внешняя по отношению к Ω, то можно принять эту точку за начало координат и тогда просто положить $\alpha(\rho)=1$. Рассмотрим функцию

$$\psi(x)=\int_{\Omega}\left(\frac{1}{r^{m-2}}-\frac{\alpha(\rho)}{\rho^{m-2}}\right)f(y)\,dy; \tag{42.15}$$

мы полагаем при этом, что $m>2$; при $m=2$ мы положим

$$\psi(x)=\int_{\Omega}\left(\ln\frac{1}{r}-\alpha(\rho)\ln\frac{1}{\rho}\right)f(y)\,dy. \tag{42.15'}$$

Интеграл (42.15) существует при почти всех $x\in\Omega$, если, например, $f\in L_p(\Omega)$, $1<p<m$. Если $|\operatorname{grad}\psi|\in L_2(\Omega)$, то, как оказывается, можно положить $F(x)=C\operatorname{grad}\psi$, где C — подходящим образом выбранная постоянная.

§ 43. Другое условие разрешимости

Существование вектора $F(x)$ такого, что $f(x)=\operatorname{div}F(x)$ и $|F|\in L_2(\Omega)$, не всегда легко установить, поэтому мы укажем здесь другое, проще проверяемое достаточное условие разрешимости задачи (42.1) при краевом условии (41.2).

Для любой функции $u(x)$, определенной почти всюду в m-мерном евклидовом пространстве E_m, $m\geqslant 3$, имеющей обобщенные первые производные и сообщающей конечное значение интегралам

$$\int_{E_m}|u(x)|^2\,dx,\qquad \int_{E_m}|\operatorname{grad}u|^2\,dx, \tag{43.1}$$

справедливо неравенство[1])

$$\int_{E_m}\frac{|u(x)|^2}{|x|^2}\,dx\leqslant\frac{4}{(m-2)^2}\int_{E_m}|\operatorname{grad}u|^2\,dx. \tag{43.2}$$

Исходя из этого неравенства, нетрудно доказать, что задача (42.1), (41.2) имеет решение с конечной энергией, если размерность области Ω $m\geqslant 3$, граница S области Ω — конечная, и

$$\int_{\Omega}|x|^2\cdot|f(x)|^2\,dx<\infty. \tag{43.3}$$

Действительно,

$$\left|\int_{\Omega}u(x)\overline{f(x)}\,dx\right|\leqslant\left\{\int_{\Omega}|x|^2\cdot|f(x)|^2\,dx\right\}^{\frac{1}{2}}\left\{\int_{\Omega}\frac{|u(x)|^2}{|x|^2}\,dx\right\}^{\frac{1}{2}}. \tag{43.4}$$

[1]) См. Р. Курант и Д. Гильберт [1], гл. VI, § 5.

Так как $u \in H_A$, то существует последовательность $u_n(x)$ непрерывно дифференцируемых в $\overline{\Omega}$ финитных функций, равных нулю на S, таких, что

$$\int_{\Omega} |\operatorname{grad}(u_n - u)|^2 \, dx \xrightarrow[n\to\infty]{} 0. \tag{43.5}$$

Доопределим функции $u_n(x)$, положив их равными нулю вне Ω. Доопределенные таким образом, эти функции имеют обобщенные первые производные, а интегралы (43.1) для этих функций конечны; по неравенству (43.2)

$$\int_{E_m} \frac{|u_n(x)|^2}{|x|^2} \, dx \leqslant \frac{4}{(m-2)^2} \int_{E_m} |\operatorname{grad} u_n|^2 \, dx,$$

или, если отбросить равные нулю интегралы по множеству, дополнительному к Ω,

$$\int_{\Omega} \frac{|u_n(x)|^2}{|x|^2} \, dx \leqslant \frac{4}{(m-2)^2} \int_{\Omega} |\operatorname{grad} u_n|^2 \, dx. \tag{43.6}$$

Построим шар с центром в начале координат и радиусом R столь большим, чтобы граница S области Ω вся лежала внутри указанного шара. Часть этого шара, общую с областью Ω, обозначим через Ω_R. В силу формул (43.5) и (43.6), при любом R имеют место соотношения

$$\int_{\Omega_R} |\operatorname{grad}(u_n - u)|^2 \, dx \xrightarrow[n\to\infty]{} 0 \tag{43.7}$$

и

$$\int_{\Omega_R} \frac{|u_n(x)|^2}{|x|^2} \, dx \leqslant \frac{4}{(m-2)^2} \int_{\Omega} |\operatorname{grad} u_n|^2 \, dx. \tag{43.8}$$

Функции $u_n(x)$ и $u(x)$ обе обращаются в нуль на S, а тогда из (43.7) следует[1]), что $u_n \to u$ в метрике $L_2(\Omega_R)$. Переходя в (43.8) к пределу сначала по n, а затем по R, получим неравенство

$$\int_{\Omega} \frac{|u(x)|^2}{|x|^2} \, dx \leqslant \frac{4}{(m-2)^2} \int_{\Omega} |\operatorname{grad} u|^2 \, dx, \tag{43.9}$$

верное для любой функции $u \in H_A$. Отсюда и из неравенства (43.4) следует, что функционал (42.3) ограничен в H_A и, следовательно, задача (42.1), (41.2) имеет решение с конечной энергией.

---

[1]) Это легко вытекает из соболевских теорем вложения.

§ 44. Случай однородного дифференциального уравнения

Цель настоящего параграфа — отметить одно осложнение, которое может возникнуть при применении энергетического метода к краевым задачам для однородных дифференциальных уравнений в случае бесконечной области.

Ограничимся рассмотрением задачи Дирихле для уравнения Лапласа. Пусть Ω — бесконечная область с конечной границей и пусть требуется найти гармоническую в этой области функцию $v(x)$, удовлетворяющую краевому условию

$$v|_S = \psi|_S; \tag{44.1}$$

мы предполагаем при этом, что $\psi \in W_2^{(1)}(\Omega)$, так что

$$\int_\Omega |\operatorname{grad} \psi|^2 dx < \infty. \tag{44.2}$$

Поставим задачу об отыскании минимума интеграла

$$\int_\Omega |\operatorname{grad} v|^2 dx \tag{44.3}$$

на множестве функций, которые удовлетворяют условию (44.1) и для которых интеграл (44.3) конечен. Нетрудно видеть, что эта задача имеет решение. Чтобы доказать это, введем в рассмотрение гильбертово пространство, которое мы обозначим через H_0 и которое состоит из функций $u(x)$, удовлетворяющих условиям:

1) $\int_\Omega |\operatorname{grad} u|^2 dx < \infty$;

2) $u|_S = 0$;

3) существует последовательность $u_n(x)$ непрерывно дифференцируемых в $\overline{\Omega}$ финитных функций таких, что

$$\int_\Omega |\operatorname{grad}(u_n - u)|^2 dx \underset{n\to\infty}{\longrightarrow} 0. \tag{44.4}$$

Скалярное произведение в H_0 зададим формулой

$$[u_1, u_2]_0 = \int_\Omega \operatorname{grad} u_1 \cdot \operatorname{grad} u_2 \, dx. \tag{44.5}$$

Заметим, что H_0 совпадает с пространством H_A § 41, если $Au = -\Delta u$ и выполняется краевое условие (41.2). Положим в интеграле (44.3) $v = \psi - u$. Задача о минимуме интеграла (44.3) сводится тогда к задаче о минимуме в пространстве H_0 функционала

$$|u|_0^2 - 2\operatorname{Re} \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} \psi \, dx. \tag{44.6}$$

По неравенству Коши — Буняковского

$$\left| \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} \psi \, dx \right| \leqslant \left\{ \int_\Omega |\operatorname{grad} \psi|^2 \, dx \right\}^{\frac{1}{2}} |u|_0, \tag{44.7}$$

линейный функционал, входящий в выражение (44.6), ограничен в H_0, и задача о минимуме функционала (44.6) имеет некоторое решение $u_0 \in H_0$. Теперь функция $v(x) = \psi(x) - u_0(x)$ реализует минимум функционала (44.3).

Легко доказать (ср. ПМ, теорема 1 § 33), что функция $v_0(x)$ гармонична в любой конечной подобласти Ω. В таком случае в области вне достаточно большой сферы эта функция разлагается в ряд по сферическим функциям:

$$v_0(x) = C + \sum_{n=1}^{\infty} \sum_{k=1}^{k_{n,m}} a_{nk} |x|^n Y_{n,m}^{(k)}(\theta) + \sum_{n=0}^{\infty} \sum_{k=1}^{k_{n,m}} b_{nk} \frac{Y_{n,m}^{(k)}(\theta)}{|x|^{n+m-2}}.$$

Здесь C — постоянная, $\theta = \frac{x}{|x|}$, $Y_{n,m}^{(k)}(\theta)$ — линейно независимые сферические функции n-го порядка в пространстве E_m, $k_{n,m}$ — число таких линейно независимых функций. Интеграл

$$\int_\Omega |\operatorname{grad} v_0|^2 \, dx$$

должен быть конечным, поэтому необходимо $a_{nk} = 0$, и при достаточно больших x.

$$v_0(x) = C + \sum_{n=1}^{\infty} \sum_{k=1}^{k_n} b_{nk} \frac{Y_{n,m}^{(k)}(\theta)}{|x|^{n+m-2}}. \tag{44.8}$$

Если $m = 2$, то $v_0(x)$ решает поставленную в начале параграфа задачу. Если же $m > 2$, то в определение гармоничности функции $v_0(x)$ обычно включается требование $C = 0$ — без этого требования решение задачи Дирихле для бесконечной области с конечной границей неединственно.

Поместим начало координат внутри S.

Пусть $C = 0$. Равенство (44.8) показывает тогда, что интеграл

$$\int_\Omega \frac{v_0(x)}{|x|^m} \, dx \tag{44.9}$$

сходится. Очевидно и обратное: если интеграл (44.9) сходится, то в формуле (44.8) $C = 0$. Функция $u_0(x) \in H_A$; в силу неравенства

(43.9) интеграл

$$\int_{\Omega} \frac{u_0(x)}{|x|^m}\,dx$$

сходится. Действительно, пусть радиус R выбран, как в § 43. Тогда, если $R_1 > R$, то

$$\left| \int_{R<|x|<R_1} \frac{u_0(x)}{|x|^m}\,dx \right| \leqslant$$

$$\leqslant \left\{ \int_{R<|x|<R_1} \frac{|u_0(x)|^2}{|x|^2}\,dx \right\}^{\frac{1}{2}} \left\{ \int_{S_1} dS_1 \int_R^{R_1} \frac{d|x|}{|x|^{2m-2}} \right\}^{\frac{1}{2}} \leqslant$$

$$\leqslant \sqrt{\frac{\omega_m}{2m-3}\left(\frac{1}{R^{2m-3}} - \frac{1}{R_1^{2m-3}}\right)} \left\{ \int_{\Omega} \frac{|u_0(x)|^2}{|x|^2}\,dx \right\}^{\frac{1}{2}} \underset{R\to\infty}{\longrightarrow} 0,$$

где S_1 — единичная сфера в пространстве E_m, а ω_m — площадь поверхности этой сферы. Таким образом, чтобы построенная энергетическим методом функция $v_0(x)$ была гармонической в области Ω первого типа, размерность которой $m \geqslant 3$, необходимо и достаточно, чтобы сходился интеграл

$$\int_{\Omega} \frac{\psi(x)}{|x|^m}\,dx = \int_{\Omega} \frac{u_0(x)}{|x|^m}\,dx + \int_{\Omega} \frac{v_0(x)}{|x|^m}\,dx. \tag{44.10}$$

Важно отметить, что этого всегда можно добиться: если интеграл (44.10) расходится, то можно заменить функцию $\psi(x)$ произведением $\psi(x)\zeta(x)$, где $\zeta(x)$ — непрерывно дифференцируемая функция такая, что

$$\zeta(x) = \begin{cases} 1, & |x| \leqslant R, \\ 0, & |x| \geqslant R+1, \end{cases}$$

и число R достаточно велико; условия (44.1) и (44.2) при этом не нарушаются, а интеграл

$$\int_{\Omega} \frac{\psi(x)\zeta(x)}{|x|^m}\,dx$$

очевидно сходится.

§ 45. Вырождающиеся уравнения в конечных областях

Рассмотрим уравнение второго порядка

$$-\sum_{j,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}\frac{\partial u}{\partial x_k}\right) = f(x), \qquad A_{jk} = \overline{A}_{kj}, \tag{45.1}$$

эллиптическое в данной конечной области Ω, но вырождающееся на некоторой части S' границы S этой области. Допустим, что существует [1]) достаточно гладкое преобразование независимых переменных $x_1, x_2, \ldots, x_m$ в новые переменные $\xi_1, \xi_2, \ldots, \xi_m$, которое переводит S' в часть плоскости $\xi_m = 0$ и уничтожает коэффициенты $A_{jm} (j = 1, 2, \ldots, m-1)$. Имея это в виду, будем сразу считать, что $A_{jm} \equiv 0$ $(j < m)$ и что часть границы S', на которой уравнение (45.1) вырождается, лежит в плоскости $x_m = 0$. При этом мы допускаем, что S' является единственным в замкнутой области $\overline{\Omega}$ множеством вырождения уравнения (45.1), так что в тех точках $\overline{\Omega}$, в которых $x_m > 0$, форма

$$\sum_{j,k=1}^{m} A_{jk}\bar{t}_j t_k = \sum_{j,k=1}^{m-1} A_{jk}\bar{t}_j t_k + A_{mm} |t_m|^2$$

— положительно определенная, а при $x_m = 0$ она вырождается.

Примем следующее допущение: в промежутке $0 \leqslant x_m \leqslant h$, где h — максимальное значение x_m в Ω, существует непрерывно дифференцируемая неотрицательная функция $\varphi(x_m)$ такая, что отношение $\dfrac{A_{mm}(x)}{\varphi(x_m)}$ непрерывно и ограничено сверху и снизу положительными числами.

Очевидно, $\varphi(0) = 0$ и $\varphi(x_m) > 0$ при $x_m > 0$.

Обозначим через S'' часть границы S, дополнительную к S'. Дифференциальный оператор в левой части уравнения (45.1) будем рассматривать на множестве $N(\Omega)$ функций, удовлетворяющих следующим требованиям:

а) если $u(x) \in N(\Omega)$, то функции

$$u, \quad \frac{\partial u}{\partial x_j}, \quad \frac{\partial^2 u}{\partial x_j \partial x_k}, \quad j, k \leqslant m-1; \qquad \varphi(x_m)\frac{\partial u}{\partial x_m}, \quad \frac{\partial}{\partial x_m}\left(\varphi(x_m)\frac{\partial u}{\partial x_m}\right)$$

непрерывны в $\overline{\Omega}$;

б) $u|_{S''} = 0$; (45.2)

в) если интеграл

$$\int_0^h \frac{dt}{\varphi(t)} \tag{45.3}$$

сходится, то поставим еще условие

$$u|_{S'} = 0; \tag{45.4}$$

если же интеграл (44 3) расходится, то на S' не будем задавать никаких условий.

Оператор, определяемый на множестве $N(\Omega)$ дифференциальным выражением в левой части уравнения (45.1), оказывается в $L_2(\Omega)$ положительным и потому допускающим расширение до самосопряженного

[1]) Подробнее об этом см. статью автора [9], в которой изложены приводимые ниже сведения о вырождающихся уравнениях.

по Фридрихсу; это расширение обозначим через A. Оказывается, что оператор A — положительно определенный, если сходится интеграл

$$\int_0^h \frac{t\,dt}{\varphi(t)}; \tag{45.5}$$

при этом спектр этого оператора дискретен. Оператор A положительно определенный и в том случае, когда $at^2 \leqslant \varphi(t) \leqslant bt^2$, где a и b — положительные постоянные, но дискретность спектра в этом случае теряется. Наконец, если $C_1 t^\beta \leqslant \varphi(t) \leqslant C_2 t^\beta$, где C_1, C_2 — положительные и $\beta > 2$, и если существует непрерывная при $0 \leqslant t \leqslant h$ неотрицательная функция $\psi(t)$ такая, что $\psi(0) = 0$ и

$$\sum_{j,\,k=1}^{m-1} A_{jk}(x)\,\bar{t}_j t_k \leqslant \psi(x_m) \sum_{k=1}^{m-1} |t_k|^2, \tag{45.6}$$

то оператор A — только положительный, но не положительно определенный.

Можно ввести энергетическое пространство H_A. Норма в нем определяется формулой

$$|u|_A^2 = \int_\Omega \sum_{j,\,k=1}^{m} A_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k}\,dx; \tag{45.7}$$

оно состоит из функций, которые имеют в Ω обобщенные первые производные, сообщают конечное значение интегралу (45.7) и удовлетворяют условию (45.2).

Поставим «задачу D» об интегрировании уравнения (45.1) при условиях (45.2) и (45.4); последнее условие не ставится, если расходится интеграл (45.3). Задача D имеет решение с конечной энергией, если в пространстве H_A существует функция, реализующая в этом же пространстве минимум интеграла

$$\int_\Omega \left\{ \sum_{j,\,k=1}^{m} A_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k} - 2\,\mathrm{Re}\, u\bar{f} \right\} dx.$$

Теорема 45.1. *Пусть в уравнении* (45.1) *функция* $f(x)$ *суммируема в любой внутренней подобласти области* Ω.

Задача (45.1) — (45.2) *имеет решение с конечной энергией тогда и только тогда, когда* $f(x) = \mathrm{div}\, F(x)$, *где вектор* $F(x)$ *таков, что для любой функции* $u \in N(\Omega)$ *верна формула*

$$\int_\Omega u\,\mathrm{div}\,\bar{F}\,dx = -\int_\Omega \mathrm{grad}\, u \cdot F\,dx \tag{45.8}$$

и сходится интеграл

$$\int\limits_{\Omega} \sum_{k=1}^{m} F_k \overline{\omega}_k \, dx, \tag{45.9}$$

в котором функции ω_k *определены из системы уравнений*

$$\sum_{k=1}^{m} A_{jk}\omega_k = F_j. \tag{45.10}$$

Доказательство. Пусть $u \in N(\Omega)$. Рассмотрим функционал

$$lu = \int\limits_{\Omega} u\overline{f} \, dx = \int\limits_{\Omega} u \operatorname{div} \overline{F} \, dx, \qquad u \in H_A. \tag{45.11}$$

По формулам (45.8) и (45.10)

$$lu = -\int\limits_{\Omega} \sum_{j=1}^{m} \overline{F}_j \frac{\partial u}{\partial x_j} dx = -\int\limits_{\Omega} \sum_{j,\,k=1}^{m} \overline{A}_{jk} \frac{\partial u}{\partial x_j} \overline{\omega}_k \, dx =$$
$$= -\int\limits_{\Omega} \sum_{j,\,k=1}^{m} A_{jk}\overline{\omega}_j \frac{\partial u}{\partial x_k} dx.$$

Квадратичный функционал

$$\int\limits_{\Omega} \sum_{j,\,k=1}^{m} A_{jk}(x)\, \overline{\zeta_j(x)}\, \zeta_k(x) \, dx$$

неотрицателен и для него справедливо неравенство Коши — Буняковского. Отсюда

$$|lu| \leqslant \left\{ \int\limits_{\Omega} \sum_{j,\,k=1}^{m} A_{jk}\overline{\omega}_j \omega_k \, dx \right\}^{\frac{1}{2}} \left\{ \int\limits_{\Omega} \sum_{j,\,k=1}^{m} A_{jk} \frac{\partial \overline{u}}{\partial x_j} \frac{\partial u}{\partial x_k} dx \right\}^{\frac{1}{2}} =$$
$$= |u|_A \left\{ \int\limits_{\Omega} \sum_{j=1}^{m} F_j \overline{\omega}_j \, dx \right\}^{\frac{1}{2}}.$$

Последнее неравенство показывает, что функционал (45.9) ограничен в H_A и, следовательно, задача D имеет решение с конечной энергией.

Обратно, пусть задача D имеет решение $u_0(x)$ с конечной энергией. Тогда

$$lu = \int\limits_{\Omega} u\overline{f} \, dx = [u, u_0]_A = \int\limits_{\Omega} \sum_{j,\,k=1}^{m} A_{jk} \frac{\partial u}{\partial x_j} \frac{\partial \overline{u}_0}{\partial x_k} dx, \tag{45.12}$$
$$u \in N(\Omega),$$

и достаточно положить

$$F_j=\sum_{k=1}^{m} \bar{A}_{jk} \frac{\partial u_0}{\partial x_k}=\sum_{k=1}^{m} A_{kj} \frac{\partial u_0}{\partial x_k}.$$

Докажем это. Возьмем функцию $u(x) \in C^{(1)}(\Omega)$, равную нулю в некоторой пограничной полосе области Ω, и построим ее среднюю функцию [1]) $u_h(x)$. При достаточно малом радиусе усреднения h функция $u_h \in N(\Omega)$, и для этой функции справедливо тождество (45.12), которому можно придать вид

$$\int_\Omega u_h \bar{f}\, dx=-\int_\Omega \operatorname{grad} u_h \cdot F\, dx.$$

Пусть $h \to 0$. При этом u_h и $\operatorname{grad} u_h$ равномерно стремятся соответственно к u и $\operatorname{grad} u$, а функция f и вектор F суммируемы в любой внутренней подобласти Ω. В таком случае в последнем равенстве можно перейти к пределу:

$$\int_\Omega u \bar{f}\, dx=-\int_\Omega \operatorname{grad} u \cdot F\, dx.$$

Полученное равенство означает, что $f(x)=\operatorname{div} F(x)$. При принятом определении вектора F, очевидно, имеет место формула (45.8). Далее, как легко видеть, в этом случае $\omega_k=\frac{\partial u_0}{\partial x_k}$ и

$$\int_\Omega \sum_{k=1}^{m} F_k \bar{\omega}_k\, dx=\int_\Omega \sum_{j,\, k=1}^{m} A_{jk} \frac{\partial u_0}{\partial x_j} \frac{\partial \bar{u}_0}{\partial x_k}\, dx=|u|_A^2,$$

так что интеграл (45.9) сходится.

Рассмотрим для примера случай, когда функция $f(x)$ непрерывна в $\bar{\Omega}$; для простоты допустим, что область Ω выпуклая, а поверхность S'' — достаточно гладкая. Пусть прямая, параллельная оси x_1, встречает поверхность S'' в точках $\alpha(x_2, \ldots, x_m)$ и $\beta(x_2, \ldots, x_m)$, $\alpha<\beta$. Положим $F=(F_1, F_2, \ldots, F_m)$, где

$$F_1=\int_\alpha^{x_1} f(\xi, x_2, \ldots, x_m)\, d\xi, \quad F_j=0, \; j>1.$$

Тогда $f(x)=\operatorname{div} F(x)$. Если $u(x) \in N(\Omega)$, то

$$-\int_\Omega \operatorname{grad} u \cdot F(x)\, dx=-\int_\Omega \frac{\partial u}{\partial x_1} F_1(x)\, dx=$$
$$=\int_\Omega u f(x)\, dx-\int_{S''} u F_1(x) \cos(\nu, x_1)\, dS,$$

[1]) См. С. Л. Соболев [1], а также ПМ, § 16.

где ν — внешняя нормаль к S''; в силу условия (45.2) контурный интеграл исчезает, и тождество (45.8) справедливо. Задача D имеет решение с конечной энергией, если сходится интеграл

$$\int_{\Omega} F_1 \overline{\omega}_1 \, dx. \tag{45.9'}$$

Принимая во внимание, что $A_{jm} = 0$, $j < m$, легко найти из системы (45.10), что $\omega_1 = \frac{F_1}{\Delta} \Delta_1$, где

$$\Delta = \begin{vmatrix} A_{11} & A_{12} & \ldots & A_{1,\, m-1} \\ A_{21} & A_{22} & \ldots & A_{2,\, m-1} \\ \cdot & \cdot & \cdot & \cdot \\ A_{m-1,\, 1} & A_{m-1,\, 2} & \ldots & A_{m-1,\, m-1} \end{vmatrix},$$

а Δ_1 есть минор элемента A_{11} в определителе Δ. Допустим, что имеет место не только оценка (45.6), но и двусторонняя оценка:

$$a\psi(x_m) \sum_{k=1}^{m-1} |t_k|^2 \leqslant \sum_{j,\, k=1}^{m-1} A_{jk} \overline{t}_j t_k \leqslant \psi(x_m) \sum_{k=1}^{m-1} |t_k|^2, \tag{45.13}$$

где a — постоянная, $0 < a \leqslant 1$. Тогда все собственные числа матрицы коэффициентов A_{jk} $(j,\, k \leqslant m-1)$ имеют оценку $O(\psi(x_m))$. Отсюда, в свою очередь, следует, что при x_m, близком к нулю, величина Δ имеет в точности порядок малости $[\psi(x_m)]^{m-1}$. Положив в неравенстве (45.13) $t_1 = 0$, мы найдем, что

$$a\psi(x_m) \sum_{k=2}^{m-1} |t_k|^2 \leqslant \sum_{j,\, k=2}^{m-1} A_{jk} \overline{t}_j t_k \leqslant \psi(x_m) \sum_{k=2}^{m-1} |t_k|^2$$

и, следовательно, порядок малости (при малых x_m) величины Δ_1 в точности такой же, как и у величины $[\psi(x_m)]^{m-2}$. Подынтегральная функция в (45.12) имеет оценку

$$\frac{|F_1 \overline{\omega}_1|}{\Delta} \leqslant \frac{|F_1|^2 \Delta_1}{\Delta} \leqslant C \frac{|F_1|^2}{\psi(x_m)}, \qquad C = \mathrm{const}.$$

и решение с конечной энергией существует, если существует интеграл

$$\int_0^h \frac{dt}{\psi(t)}.$$

Далее,

$$|F_1| \leqslant (\beta - \alpha) \max_{\alpha \leqslant \xi \leqslant \beta} |f(\xi,\, x_2,\, \ldots,\, x_m)| \leqslant$$
$$\leqslant C_1 \max_{\alpha \leqslant x_1 \leqslant \beta} |f(x_1,\, x_2,\, \ldots,\, x_m)|, \qquad C_1 = \mathrm{const}.$$

Пусть $f(x) \xrightarrow[x_m \to 0]{} 0$, более определенно, пусть $|f(x)| \leqslant C_2 \chi(x_m)$, где $C_2 = \text{const}$, и $\chi(0) = 0$. Тогда решение с конечной энергией существует, если сходится интеграл

$$\int_0^h \frac{\chi^2(t)}{\psi(t)}\, dt.$$

§ 46. Координатные системы для одномерных задач в случае бесконечного промежутка[1])

1. В пространстве $L_2(0, \infty)$ рассмотрим оператор A, действующий по формуле

$$Au = -\frac{d^2u}{dx^2} + u \tag{46.1}$$

и определенный на множестве $D(A)$ дважды непрерывно дифференцируемых финитных функций, удовлетворяющих краевому условию

$$u(0) = 0. \tag{46.2}$$

Оператор A — положительно определенный, так как

$$(Au, u) = \int_0^\infty (-u''\bar{u} + |u|^2)\, dx = -u'\bar{u}\Big|_0^\infty + \int_0^\infty (|u'|^2 + |u|^2)\, dx =$$

$$= \int_0^\infty (|u'|^2 + |u|^2)\, dx \geqslant \int_0^\infty |u|^2\, dx = \|u\|^2.$$

Пространство H_A представляет собой замыкание множества $D(A)$ в метрике

$$[u, v]_A = \int_0^\infty (u'\bar{v}' + u\bar{v})\, dx.$$

Можно доказать, что H_A состоит из тех функций из $L_2(0, \infty)$, которые абсолютно непрерывны на любом конечном промежутке, для которых $u(0) = 0$ и

$$|u|_A^2 = \int_0^\infty (|u'|^2 + |u|^2)\, dx < \infty. \tag{46.3}$$

Рассмотрим систему функций:

$$\varphi_k = \frac{2}{\sqrt{\pi}} \frac{\sin(2k \operatorname{arc\,tg} x)}{\sqrt{1+x^2}} \qquad (k = 1, 2, \ldots). \tag{46.4}$$

---

[1]) В этом параграфе излагаются результаты статьи Ю. С. Вержбинской [1].

Докажем, что она удовлетворяет требованиям 1) — 3) § 6. Выполнение требований 1) и 2) очевидно.

Убедимся, что система $\{\varphi_k\}$ полна в H_A. Для этого нам достаточно доказать, что из равенства:

$$\int_0^\infty (f'\varphi_k' + f\varphi_k)\,dx = 0 \qquad (k = 1, 2, \ldots), \tag{46.5}$$

где $f \in H_A$, вытекает, что $f \equiv 0$.

Имеем:

$$\int_0^\infty \frac{2kf'}{(1+x^2)^{\frac{3}{2}}} \cos(2k \operatorname{arctg} x)\,dx + \\ + \int_0^\infty \left[\frac{-xf'}{(1+x^2)^{\frac{3}{2}}} + \frac{f}{\sqrt{1+x^2}}\right] \sin(2k \operatorname{arctg} x)\,dx = 0.$$

Беря второй интеграл по частям, получим:

$$\int_0^\infty \frac{2kf'}{(1+x^2)^{\frac{3}{2}}} \cos(2k \operatorname{arctg} x)\,dx + \\ + \left[\sin(2k \operatorname{arctg} x) \cdot \int_0^x \left[\frac{-tf'}{(1+t^2)^{\frac{3}{2}}} + \frac{f}{(1+t^2)^{\frac{1}{2}}}\right] dt\right]\Bigg|_0^\infty - \\ - \int_0^\infty \left[\int_0^x \left[-\frac{tf'}{(1+t^2)^{\frac{3}{2}}} + \frac{f}{(1+t^2)^{\frac{1}{2}}}\right] dt\right] 2k \frac{\cos(2k \operatorname{arctg} x)}{(1+x^2)}\,dx = 0 \\ (k = 1, 2, \ldots).$$

Легко доказать, что в последней формуле внеинтегральный член пропадает, и наше равенство принимает вид:

$$2k \int_0^\infty \left\{\frac{f'}{(1+x^2)^{\frac{1}{2}}} - \int_0^x \left(-\frac{tf'}{(1+t^2)^{\frac{3}{2}}} + \frac{f}{(1+t^2)^{\frac{1}{2}}}\right) dt\right\} \times \\ \times \frac{\cos(2k \operatorname{arctg} x)}{(1+x^2)}\,dx = 0 \qquad (k = 1, 2, \ldots).$$

Выражение в фигурных скобках обозначим через $\Phi(x)$. Положив $x = \operatorname{tg}\frac{t}{2}$, получим

$$\int_0^\pi \Phi\left(\operatorname{tg}\frac{t}{2}\right) \cos kt\,dt = 0 \qquad (k = 1, 2, \ldots).$$

Система функций $\{\cos kt\}$ полна в $L_2(0, \pi)$ и потому $\Phi(x) \equiv \text{const}$. Отсюда

$$\frac{f'}{(1+x^2)^{\frac{1}{2}}} = \int_0^x \left[-\frac{tf'}{(1+t^2)^{\frac{3}{2}}} + \frac{f}{(1+t^2)^{\frac{1}{2}}} \right] dt + \text{const}. \tag{46.6}$$

Следовательно, функция f, удовлетворяющая равенству (46.5) и принадлежащая пространству H_A, имеет вторую производную. Дифференцируя уравнение (46.6), находим, что $-f'' + f = 0$. Отсюда $f = c_1 e^{-x} + c_2 e^{x}$. Но $c_2 = 0$, так как функция $f \in L_2(0, \infty)$, а $c_1 = 0$, так как $f(0) = 0$. Окончательно, $f \equiv 0$, что и требовалось доказать.

Система (46.4) сильно минимальна в H_A. Действительно, пространство H_A вкладывается в $L_2(\Omega)$, так как оператор A положительно определенный. Далее, система функций (46.4) ортонормирована в метрике $L_2(0, \infty)$; в силу следствия 4.2 она сильно минимальна в пространстве H_A.

Теперь рассмотрим оператор более общего вида:

$$\left.\begin{aligned} A_1 u &= -\frac{d}{dx}\left[p(x)\frac{du}{dx}\right] + q(x)u, \\ u(0) &= 0, \qquad x \in [0, \infty]. \end{aligned}\right\} \tag{46.7}$$

Наложим на $p(x)$ и $q(x)$ следующие требования:

а) $p(x)$ — положительно ограниченная сверху и снизу и имеющая непрерывную и ограниченную первую производную функция; пусть

$$0 < p_0 \leqslant p(x) \leqslant p_1; \tag{46.8}$$

б) $q(x)$ — измеримая, положительно ограниченная сверху и снизу функция; пусть

$$0 < q_0 \leqslant q(x) \leqslant q_1. \tag{46.9}$$

В неравенствах (46.8) и (46.9) величины p_0, q_0, p_1, q_1 — постоянные.

Метрика в пространстве H_A определяется так:

$$[u, u]_{A_1} = \int_0^\infty \left(p(x)u'^2 + q(x)u^2\right) dx.$$

Из неравенств (46.8) и (46.9) вытекает, что метрика пространства H_{A_1} эквивалентна метрике пространства H_A и, следовательно, система функций (46.4) удовлетворяет условиям 1) — 3) § 6 и сильно минимальна в пространстве H_{A_1}.

2. Рассмотрим теперь оператор (46.1) при краевом условии

$$u'(0) = 0, \qquad x \in [0, \infty]. \tag{46.10}$$

За область определения этого оператора, который будем, по-прежнему, обозначать буквой A, примем множество финитных дважды непрерывно дифференцируемых функций, удовлетворяющих условию (46.10). Рассматриваемый нами оператор положительно определенный; как нетрудно доказать, его энергетическое пространство H_A состоит из функций, абсолютно непрерывных на любом конечном промежутке, для которых

$$|u|_A^2 = \int_0^\infty (u'^2 + u^2)\,dx < \infty.$$

Краевое условие (46.10) — естественное, и функции из H_A не обязательно ему удовлетворяют.

Рассмотрим систему функций:

$$\psi_k(x) = \frac{2}{\sqrt{\pi}} \frac{\cos(2k \operatorname{arctg} x)}{\sqrt{1+x^2}} \qquad (k = 0,\ 1,\ 2,\ \ldots). \tag{46.11}$$

Нетрудно доказать, что функции $\psi_k(x)$ удовлетворяют требованиям 1) и 2) § 6.

Докажем, что система функций $\psi_k(x)$ полна в пространстве H_A. Для этого достаточно из равенств

$$\int_0^\infty (f'\psi_k' + f\psi_k)\,dx = 0 \qquad (k = 0,\ 1,\ 2,\ \ldots), \tag{46.12}$$

в которых $f \in H_A$, извлечь, что $f \equiv 0$.

Имеем:

$$\int_0^\infty \left[-\frac{f'}{(1+x^2)^{\frac{3}{2}}}\, 2k \sin(2k \operatorname{arctg} x) \right] dx + \\ + \int_0^\infty \left[-\frac{xf'}{(1+x^2)^{\frac{3}{2}}} + \frac{f}{(1+x^2)^{\frac{1}{2}}} \right] \cos(2k \operatorname{arctg} x)\,dx = 0. \tag{46.13}$$

Беря второй интеграл по частям, получим:

$$\int_0^\infty \left[-\frac{f'}{(1+x^2)^{\frac{3}{2}}}\, 2k \sin(2k \operatorname{arctg} x) \right] dx + \\ + \left[\int_0^x \left(-\frac{tf'}{(1+t^2)^{\frac{3}{2}}} + \frac{f}{(1+t^2)^{\frac{1}{2}}} \right) dt \cos(2k \operatorname{arctg} x) \right]_0^\infty + \\ + 2k \int_0^\infty \left[\int_0^x \left(-\frac{tf'}{(1+t^2)^{\frac{3}{2}}} + \frac{f}{(1+t^2)^{\frac{1}{2}}} \right) dt \right] \sin(2k \operatorname{arctg} x) \frac{1}{1+x^2}\,dx = 0.$$

Докажем, что в последнем равенстве внеинтегральный член исчезает. Действительно, при $x=0$ это очевидно, а при $x=\infty$ внеинтегральный член исчезает, так как из равенства (46.12) при $k=0$ имеем:

$$\int_0^\infty \left[-\frac{tf'}{(1+t^2)^{\frac{3}{2}}}+\frac{f}{(1+t^2)^{\frac{1}{2}}}\right]dt=0.$$

Равенство (46.13) теперь принимает вид

$$2k\int_0^\infty \left\{-\frac{f'}{(1+x^2)^{\frac{1}{2}}}+\int_0^x \left[-\frac{tf'}{(1+t^2)^{\frac{3}{2}}}+\frac{f}{(1+t^2)^{\frac{1}{2}}}\right]dt\right\}\times$$
$$\times\frac{\sin(2k\operatorname{arctg}x)}{1+x^2}\,dx=0.$$

Обозначая выражение, взятое в фигурные скобки, через $\Phi(x)$, имеем:

$$\int_0^\infty \Phi(x)\frac{\sin(2k\operatorname{arctg}x)}{1+x^2}\,dx=0.$$

Рассуждая так же, как и в предыдущем параграфе, приходим к выводу, что последнее равенство влечет за собой тождество $f\equiv 0$.

Система (46.11) сильно минимальна в пространстве H_A. Действительно, эта система ортонормирована в $H=L_2(0,\infty)$, так как

$$(\psi_k,\psi_m)=\frac{2}{\pi}\int_0^\infty \cos(2k\operatorname{arctg}x)\cos(2m\operatorname{arctg}x)\frac{2}{1+x^2}\,dx=$$
$$=\frac{2}{\pi}\int_0^\pi \cos kt\cos mt\,dt=\delta_{km}.$$

Далее, оператор (46.1), (46.10) положительно определен в $L_2(0,\infty)$, и сильная минимальность системы (46.11) в пространстве H_A вытекает из следствия 4.2.

Рассмотрим оператор более общего вида

$$A_2u=-\frac{d}{dx}\left[p(x)\frac{du}{dx}\right]+q(x)u,\qquad u'(0)=0,\qquad x\in[0,\infty],$$

где p и q подчинены требованиям п. 1.

Очевидно, что пространство H_{A_2} с метрикой

$$|u|_{A_2}^2=\int_0^\infty \left[p(x)u'^2+q(x)u^2\right]dx$$

эквивалентно пространству H_A. Следовательно, функции (46.11) можно брать в качестве координатных для операторов A_2 и при этом система $\{\psi_k(x)\}$ сильно минимальна в пространстве H_{A_2}.

3. Рассмотрим оператор

$$\left.\begin{aligned} &-\frac{d}{dx}\left[p(x)\frac{du}{dx}\right]+q(x)\,u, \\ &u'(0)-\alpha u(0)=0, \quad x\in[0,\ \infty], \quad \alpha>0. \end{aligned}\right\} \tag{46.14}$$

Этот оператор обозначим через A_3. Подчиним $p(x)$ и $q(x)$ требованиям п. 1.

Краевые условия (46.10) и (46.14) являются естественными для дифференциальных операторов второго порядка.

Как легко видеть, отсюда вытекает, что метрики пространств H_{A_2} и H_{A_3} эквивалентны.

В таком случае систему функций (46.11) можно использовать в качестве координатной системы для оператора A_3. При этом система функций $\{\psi_k(x)\}$ сильно минимальна в пространстве H_{A_3}.

4. Рассмотрим дифференциальный оператор

$$-\frac{d}{dx}\left[p(x)\frac{du}{dx}\right]+q(x)\,u$$

в пространстве $L_2(-\infty,\ +\infty)$; обозначим этот оператор через A_4. На функции $p(x)$ и $q(x)$ наложим условия п. 1.

Рассуждая так же, как и в случае первой и второй краевой задачи, можно доказать, что система функций

$$\frac{1}{\sqrt{1+x^2}},\quad \frac{\sin(2\operatorname{arctg}x)}{\sqrt{1+x^2}},\quad \frac{\cos(2\operatorname{arctg}x)}{\sqrt{1+x^2}},\ \ldots$$
$$\ldots,\ \frac{\sin(2k\operatorname{arctg}x)}{\sqrt{1+x^2}},\quad \frac{\cos(2k\operatorname{arctg}x)}{\sqrt{1+x^2}},\ \ldots$$

может служить координатной системой для оператора A_4 и сильно минимальна в пространстве H_{A_4}.

5. Рассмотрим оператор

$$-\frac{d}{dx}\left[p(x)\frac{du}{dx}\right]+q(x)\,u$$

в пространстве $L_2(0,\ \infty)$ при краевом условии (46.2). Этот оператор обозначим A_5.

На функции $p(x)$ и $q(x)$ наложим следующие требования:

а) $p(x)$ — ограниченная сверху и снизу и имеющая непрерывную первую производную функция, причем

$$0<p_0\leqslant p(x)\leqslant p_1, \qquad p_0=\text{const}, \qquad p_1=\text{const}.$$

б) $q(x)$ — измеримая неотрицательная ограниченная сверху функция такая, что при больших значениях x

$$q(x) = O(x^{-\alpha}), \qquad \alpha \geqslant 2.$$

Пространство H_{A_5} состоит из тех функций из $L_2(0, \infty)$, которые абсолютно непрерывны на любом конечном отрезке, для которых $u(0) = 0$ и

$$|u|_A^2 = \int_0^\infty \left[p(x)|u'|^2 + q(x)|u|^2\right] dx < \infty. \tag{46.15}$$

Докажем, что система функций

$$\eta_k(x) = \int_0^x \frac{e^{-\frac{t}{2}} L_k(t)}{k!} dt \qquad (k = 0, 1, 2, \ldots), \tag{46.16}$$

где $L_k(t)$ — полиномы Лагерра, удовлетворяет условиям 1) — 3) § 6. Действительно, как известно,

$$L_k(t) = e^t \frac{d^k}{dt^k}\left(t^k e^{-t}\right).$$

Интегрируя по частям, легко найдем:

$$\eta_k(x) = \int_0^x \frac{e^{\frac{t}{2}}}{k!} \frac{d^k}{dt^k}\left(t^k e^{-t}\right) dt = (-1)^k 2 + \sum_{m=0}^{k} c_m x^m e^{-\frac{x}{2}}, \tag{46.17}$$

где c_m - некоторые постоянные.

Из формулы (46.17) и равенства $\eta_k(0) = 0$, очевидно, следует, что $\eta_k \in H_{A_5}$. Очевидно также, что функции $\eta_0(x), \eta_1(x), \ldots, \eta_n(x)$ линейно независимы при любом n. Докажем, что в пространстве H_{A_5} эта система также и полна.

Будем исходить из неравенства Харди [1])

$$\int_0^\infty \frac{|u(x)|}{x^2} dx \leqslant 4 \int_0^\infty |u'(x)|^2 dx, \tag{46.18}$$

которое справедливо для функций, абсолютно непрерывных на любом отрезке вида $0 \leqslant x \leqslant N$ и обращающихся в нуль при $x = 0$.

Из условия $q(x) = O(x^{-\alpha})$, $\alpha \geqslant 2$, следует, что

$$q(x) \leqslant \frac{C}{x^2}, \qquad C = \text{const}.$$

[1]) См. Г. Г. Харди, Дж. Е. Литтльвуд и Г. Полиа, [1], стр. 289.

А тогда неравенство (46.18) дает оценку

$$\int_0^\infty q(x)\,|u|^2\,dx \leqslant 4C \int_0^\infty |u'(x)|^2\,dx, \qquad u \in H_A, \tag{46.19}$$

из которой в свою очередь следует, что

$$|u|_A^2 \leqslant (1+C)\int_0^\infty |u'|^2\,dx. \tag{46.20}$$

Как хорошо известно, система функций

$$\eta_k'(x) = \frac{e^{-\frac{x}{2}}\,L_k(x)}{k!}$$

ортонормирована и полна в $L_2(0, \infty)$, поэтому, если $u \in H_A$ и ε — произвольное положительное число, то при n, достаточно большом,

$$\int_0^\infty \left| u'(x) - \sum_{k=1}^n a_k \eta_k'(x) \right|^2 dx < \varepsilon^2, \qquad a_k = \int_0^\infty u'(x)\,\eta_k'(x)\,dx.$$

В силу неравенства (46.20)

$$\left| u - \sum_{k=1}^n a_k \eta_k \right|_{A_5}^2 \leqslant (1+C)\,\varepsilon^2;$$

так как ε — произвольное положительное число, то последнее неравенство означает, что система $\{\eta_k\}$ полна в H_{A_5}.

Система $\{\eta_k\}$ почти ортонормирована в H_{A_5}. Действительно, система $\{\eta_k'\}$ ортонормирована в $L_2(0, \infty)$; это равносильно тому, что система $\{\eta_k\}$ ортонормирована в пространстве H_0 функций $u(x)$ таких, что $u(0) = 0$, с нормой

$$\| u \|_0^2 = \int_0^\infty |u'(x)|^2\,dx.$$

Формулы (46.15) и (46.20) показывают, что нормы $\|u\|_0$ и $|u|_{A_5}$ эквивалентны; по теореме 4.3, система $\{\eta_k\}$ почти ортонормирована в H_{A_5}.

§ 47. Координатные системы для многомерных задач в случае бесконечной области с конечной границей

Будем считать, что область Ω расположена вне конечного числа кусочно гладких замкнутых поверхностей (линий, если Ω — область на плоскости). Совокупность этих поверхностей — границу области Ω — обозначим, как всегда, через S; начало координат поместим внутри S. Рассмотрим первую краевую задачу

$$-\sum_{j,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk} \frac{\partial u}{\partial x_k}\right) = f(x), \tag{47.1}$$

$$u|_S = 0. \tag{47.2}$$

Коэффициенты A_{jk} подчиним требованиям § 41, а функцию $f(x)$ — условию дивергенции. Наша цель — построить систему функций $\{\varphi_k\}$, сильно минимальную или, еще лучше, почти ортонормированную в метрике

$$|u|_A = \int_\Omega \sum_{j,k=1}^{m} A_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k} dx, \tag{47.3}$$

где

$$u|_S = 0. \tag{47.4}$$

В силу неравенств (41.4), достаточно, чтобы искомая система обладала указанным свойством в метрике

$$|u|_0^2 = \int_\Omega \sum_{k=1}^{m} \left| \frac{\partial u}{\partial x_k} \right|^2 dx. \tag{47.5}$$

Отметим, что метрика (47.5) есть энергетическая метрика для оператора $-\Delta$ при краевом условии $u|_S = 0$; здесь Δ — оператор Лапласа.

Введем сферическую систему координат с центром в точке $x = 0$. Тогда

$$|u|_0^2 = \int_\Omega \left\{ \left| \frac{\partial u}{\partial \rho} \right|^2 + \frac{1}{\rho^2} \sum_{k=1}^{m-1} \frac{1}{q_k} \left| \frac{\partial u}{\partial \vartheta_j} \right|^2 \right\} dx. \tag{47.6}$$

В формуле (47.6) ρ, ϑ_1, ϑ_2, ..., ϑ_{m-1} — сферические координаты точки x; далее,

$$q_1 = 1, \qquad q_k = \prod_{j=1}^{k-1} \sin^2 \vartheta_j, \qquad k > 1.$$

Положим $\rho = \frac{1}{\rho'}$. Тогда область Ω преобразуется в некоторую область Ω'; если x и x' — точки областей Ω и Ω', соответствующие друг другу при сделанном нами преобразовании, то $dx = |J|\, dx'$, где J — якобиан преобразования $\rho = \frac{1}{\rho'}$; нетрудно видеть, что $|J| = \rho^{2m}$, и потому

$$dx = \rho^{2m}\, dx'.$$

Положим еще в интеграле (47.6) $u=\frac{u'}{\rho^{m-2}}$. Простые вычисления дают

$$|u|_0^2=\int_{\Omega'}\left\{\left|\frac{\partial u'}{\partial\rho'}-\frac{m-2}{\rho'}u'\right|^2+\frac{1}{\rho'^2}\sum_{k=1}^{m-1}\frac{1}{q_k}\left|\frac{\partial u'}{\partial\vartheta_k}\right|^2\right\}dx'. \qquad (47.7)$$

Рассмотрим сначала случай, когда $m=2$. Тогда, обозначая полярный угол через ϑ, имеем:

$$|u|_0^2=\int_{\Omega'}\left\{\left|\frac{\partial u'}{\partial\rho'}\right|^2+\frac{1}{\rho'^2}\left|\frac{\partial u'}{\partial\vartheta}\right|^2\right\}dx'. \qquad (47.8)$$

Правая часть формулы (47.8) есть квадрат энергетической нормы элемента u' для оператора Лапласа (взятого со знаком минус) при краевом условии $u'|_{S'}=0$, где S' — граница области Ω'. Обозначая энергетическое пространство указанного оператора через H_1 и норму в нем через $|\ |_1$, имеем $|u|_0=|u'|_1$. Таким образом, в случае $m=2$ преобразование $\rho=\frac{1}{\rho'}$, $u=u'$ изометрично отображает пространство H_0 на пространство H_1.

Возьмем теперь в H_1 произвольную почти ортонормированную систему $\{\varphi_n'(\rho',\vartheta)\}$, тогда система $\{\varphi_n\}$, где $\varphi_n(\rho,\vartheta)=\varphi_n'\left(\frac{1}{\rho},\vartheta\right)$, почти ортонормирована в H_0, а следовательно, и в H_A. Если область Ω' можно отобразить на одну из областей, названных в § 33, то почти ортонормированными в H_A будут указанные в этом параграфе координатные системы. В частности, если Ω есть внешность круга радиуса 1, то можно положить

$$\varphi_{kn}(\rho,\vartheta)=c_nJ_n\left(\frac{\gamma_{nk}}{\rho}\right)\begin{matrix}\cos\\ \sin\end{matrix}n\vartheta \qquad (n=0,1,2,\ldots), \qquad (47.9)$$

где c_n определяется, например, из условия $|\varphi_{kn}|_A=1$.

Вернемся к формуле (47.7) и будем считать теперь, что $m>2$. В известном неравенстве

$$|a+b|^2\leqslant 2|a|^2+2|b|^2 \qquad (47.10)$$

заменим a на $a-b$ и затем b на $-b$. Мы получим тогда

$$|a+b|^2\geqslant\frac{|a|^2}{2}-|b|^2. \qquad (47.11)$$

Отсюда

$$\left|\frac{\partial u'}{\partial\rho'}-\frac{m-2}{\rho'}u'\right|^2\geqslant\frac{1}{2}\left|\frac{\partial u'}{\partial\rho'}\right|^2-\frac{(m-2)^2}{\rho'^2}|u'|^2$$

и, следовательно,

$$\| u \|_0^2 \geqslant \frac{1}{2} \int_{\Omega'} \left\{ \left| \frac{\partial u'}{\partial \rho'} \right|^2 + \frac{1}{\rho'^2} \sum_{k=1}^{m-1} \frac{1}{q_k} \left| \frac{\partial u'}{\partial \vartheta_k} \right|^2 \right\} dx' - \\ - (m-2)^2 \int_{\Omega'} \frac{|u'|^2}{\rho'^2} dx'. \qquad (47.12)$$

Как и выше, введем пространство H_1. Тогда первый интеграл в (47.12) равен $\| u' \|_1^2$. Второй интеграл оценивается по неравенству (43.9)

$$\int_{\Omega'} \frac{|u'|^2}{\rho'^2} dx' = \int_{\Omega} \frac{u^2}{\rho^2} dx \leqslant \frac{4}{(m-2)^2} \int_{\Omega} |\operatorname{grad} u|^2 dx = \frac{4}{(m-2)^2} \| u \|_0^2.$$

Подставив это в неравенство (47.12), найдем:

$$\| u \|_0^2 \geqslant \frac{1}{10} \| u' \|_1^2. \qquad (47.13)$$

Из теоремы 4.2 следует теперь, что всякая система, сильно минимальная (в частности, ортонормированная) в H_1, переводится преобразованием $\rho = \frac{1}{\rho'}$, $u = \frac{u'}{\rho^{m-2}}$ в систему, сильно минимальную в H_0, а следовательно, и в H_1. Если, например, Ω есть внешность трехмерной единичной сферы, то сильно минимальной в H_A будет система функций

$$\varphi_{klm} = \frac{c_{klm}}{\sqrt{\rho'}} J_{k+\frac{1}{2}} \left(\frac{\gamma_{k+\frac{1}{2}, l}}{\rho'} \right) P_k^{(m)}(\cos\theta) \begin{matrix} \cos \\ \\ \sin \end{matrix} m\varphi \qquad (47.14)$$

$$(k = 0, 1, 2, \ldots; \quad l = 1, 2, \ldots; \quad m = 0, 1, \ldots, k),$$

где коэффициент c_{klm} можно определить из условия $|\varphi_{klm}|_A = 1$ или какого-либо другого условия, но так, чтобы при любых значениях индексов k, l, m нормы $|\varphi_{klm}|_A$ были ограничены сверху и снизу положительными постоянными, не зависящими от упомянутых индексов.

Обратимся ко второй краевой задаче и будем рассматривать уравнение (47.1) при краевом условии

$$\left[\sum_{j, k=1}^{m} A_{jk} \frac{\partial u}{\partial x_j} \cos(\nu, x_k) + \sigma u \right]_S = 0; \qquad (47.15)$$

как обычно, будем считать заданную на S функцию σ измеримой, неотрицательной и ограниченной. Энергетические произведения и норма в данном случае определяются формулами (41.8). Координатная система будет сильно минимальной, соответственно, почти ортонормированной в H_A, если она будет обладать тем же свойством в пространстве H_0, которое состоит из тех же функций, что и пространство H_A (см. § 41), и в котором норма определена формулой

$$
\begin{aligned}
\| u \|_0^2 &= \int_\Omega \sum_{k=1}^{m} \left| \frac{\partial u}{\partial x_k} \right|^2 dx + \int_S \sigma |u|^2 \, dS = \\
&= \int_\Omega \left\{ \left| \frac{\partial u}{\partial \rho} \right|^2 + \frac{1}{\rho^2} \sum_{k=1}^{m-1} \frac{1}{q_k} \left| \frac{\partial u}{\partial \vartheta_j} \right|^2 \right\} dx + \int_S \sigma |u|^2 \, dS. \qquad (47\ 16)
\end{aligned}
$$

В интеграле (47.16) выполним преобразование $\rho = \frac{1}{\rho'}$, $u = \rho^{m-2} u'$. Это приведет нас к равенству

$$
\begin{aligned}
\| u \|_0^2 = \int_{\Omega'} \left\{ \left| \frac{\partial u'}{\partial \rho'} - \frac{m-2}{\rho'} u' \right|^2 + \frac{1}{\rho'^2} \sum_{k=1}^{m-1} \frac{1}{q_k} \left| \frac{\partial u'}{\partial \vartheta_k} \right|^2 \right\} dx' + \\
+ \int_{S'} \sigma K |u'|^2 \, dS'; \qquad (47\ 17)
\end{aligned}
$$

здесь S' — граница области Ω' и K — некоторая ограниченная функция.

В случае $m = 2$ мы, как и выше, убедимся, что в качестве координатной можно взять, например, систему собственных функций задачи Неймана для оператора Лапласа и области Ω', подвергнутую преобразованию $\rho = \frac{1}{\rho'}$; такая система будет почти ортонормированной в H_A. В частности, если Ω — внешность единичного круга, то можно положить

$$
\varphi_{kn}(\rho, \vartheta) = c_{kn} J_n \left(\frac{\tilde{\gamma}_{n,k}}{\rho} \right) \begin{matrix} \cos \\ \sin \end{matrix} n\vartheta \qquad (n = 0, 1, 2, \ldots; \ k = 1, 2, \ldots), \qquad (47.18)
$$

где c_{kn} таково, что $|\varphi_{kn}|_A = 1$.

Если $m > 2$, то, как и выше, мы придем к неравенству

$$
\begin{aligned}
\| u \|_0^2 \geqslant \frac{1}{2} \int_{\Omega'} \left\{ \left| \frac{\partial u'}{\partial \rho'} \right|^2 + \frac{1}{\rho'^2} \sum_{k=1}^{m-1} \frac{1}{q_k} \left| \frac{\partial u'}{\partial \vartheta_k} \right|^2 \right\} dx' + \\
+ \int_{S} \sigma K |u'|^2 \, dS' - (m-2)^2 \int_\Omega \frac{|u|^2}{\rho^2} \, dx. \qquad (47.19)
\end{aligned}
$$

Функцию u можно продолжить[1]) на все пространство E_m так, чтобы продолженная функция имела обобщенные первые производные и чтобы

$$
\int_\Omega |\operatorname{grad} u|^2 \, dx \geqslant c \int_{E_m} |\operatorname{grad} u|^2 \, dx,
$$

[1]) См. В. М. Бабич [1].

где c зависит только от области Ω. В таком случае соотношение (41.2) дает следующую цепочку неравенств:

$$\int\limits_{\Omega} \frac{|u|^2}{\rho^2}\,dx \leqslant \int\limits_{E_m} \frac{|u|^2}{\rho^2}\,dx \leqslant \frac{4}{(m-2)^2} \int\limits_{E_m} |\operatorname{grad} u|^2\,dx \leqslant$$
$$\leqslant \frac{4}{c\,(m-2)^2} \int\limits_{\Omega} |\operatorname{grad} u|^2\,dx \leqslant \frac{4}{c\,(m-2)^2}\,|u|_0^2.$$

Теперь из формулы (47.19) вытекает оценка

$$|u|_0^2 \geqslant C_1 \left[\int\limits_{\Omega} |\operatorname{grad}' u'|^2\,dx' + \int\limits_{S'} \sigma K\,|u'|^2\,dS'\right]; \qquad (47.20)$$

здесь $C_1 = \mathrm{const}$, а grad' означает дифференцирование по координатам точки x'. Выражение в скобках в формуле (47.20) определяет некоторую норму в соболевском пространстве $W_2^{(1)}(\Omega')$, эквивалентную основной норме

$$\|u\|^2_{W_2^{(1)}(\Omega')} = \int\limits_{\Omega'} [\,|\operatorname{grad}' u'|^2 + |u'|^2]\,dx'. \qquad (47.21)$$

Теперь можно дословно повторить все сказанное выше о случае первой краевой задачи: любая система, сильно минимальная в метрике (47.21), переводится преобразованием $\rho = \frac{1}{\rho'}$, $u = \rho^{m-2}u'$ в систему, сильно минимальную и в энергетическом пространстве второй краевой задачи. В частности, можно использовать систему собственных функций задачи Неймана для оператора Лапласа и области Ω'; если, например, Ω есть внешность трехмерной единичной сферы, мы таким путем получим координатную систему

$$\varphi_{klm} = \frac{c_{klm}}{\sqrt{\rho}}\, J_{k+\frac{1}{2}}\left(\frac{\gamma^*_{k+\frac{1}{2},\,l}}{\rho}\right) P_k^{(m)}(\cos\theta)\, {\cos \atop \sin}\, m\varphi; \qquad (47.22)$$

коэффициент c_{klm} можно определить из условия $|\varphi_{klm}|_A = 1$.

§ 48. Координатные системы для областей с бесконечной границей

1. В настоящем параграфе мы будем рассматривать уравнение

$$-\sum_{j,\,k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk}\,\frac{\partial u}{\partial x_j}\right) = f(x) \qquad (48.1)$$

при краевых условиях первой или второй краевой задачи; эти условия имеют вид (41.2) или (41.3) соответственно. Операторы этих задач, расширенные до самосопряженных, будем обозначать через A_1 и A_2. Будем считать, что существует достаточно гладкое преобразование области Ω в цилиндрическую область, бесконечную или полубесконечную. Раз такое преобразование возможно, будем считать, что оно уже выполнено и что область Ω представляет собой цилиндр, определяемый соотношениями

$$\xi \in \omega, \qquad 0 \leqslant x_m < \infty, \tag{48.2'}$$

или

$$\xi \in \omega, \qquad -\infty < x_m < \infty. \tag{48.2''}$$

Здесь мы через ξ обозначили точку $(m-1)$-мерного евклидова пространства координат $x_1, x_2, \ldots, x_{m-1}$, через ω — область этого пространства. Ограничимся случаем, когда граница s этой области конечная; сама же область ω может быть бесконечной или конечной; в соответствии с этим Ω будет областью первого или второго типа.

Будем считать, что коэффициенты A_{jk} подчинены условиям § 41. О коэффициенте σ в краевом условии (41.3) допустим, что он измерим и что существует заданная на s функция σ_0 неотрицательная, ограниченная и измеримая, такая, что на цилиндрической части S_0 границы S выполняется соотношение

$$c_1\sigma_0 \leqslant \sigma \leqslant c_2\sigma_0, \qquad c_1,\ c_2 = \mathrm{const} > 0. \tag{48.3}$$

Координатная система будет сильно минимальной и полной в пространстве H_{A_i} $(i=1, 2)$, если она будет обладать тем же свойством в метрике $|\cdot|_i$ $(i=1, 2)$, где

$$|u|_1^2 = \int_\Omega |\operatorname{grad} u|^2\,dx, \qquad u|_S = 0 \tag{48.4}$$

и

$$|u|_2^2 = \int_\Omega |\operatorname{grad} u|^2\,dx + \int_{S_0} \sigma_0 |u|^2\,dS. \tag{48.5}$$

2. Рассмотрим сначала задачу (48.1), (41.2) и связанный с ней оператор A_1, причем пока допустим, что ω — конечная область и что Ω — полуцилиндр (48.2′). Обозначим через $\tau_k(\xi)$ собственные функции оператора

$$-\sum_{j=1}^{m-1} \frac{\partial^2\tau}{\partial x_j^2}, \qquad \tau|_s = 0. \tag{48.6}$$

Эти функции ортогональны и образуют полную систему как в метрике $L_2(\omega)$, так и в метрике, порождаемой скалярным произведением

$$[\tau, \chi]_0 = \int_\omega \sum_{j=1}^{m-1} \frac{\partial\tau}{\partial x_j}\,\frac{\partial\bar{\chi}}{\partial x_j}\,d\xi; \tag{48.7}$$

будем считать, что они нормированы в метрике $L_2(\omega)$, так что

$$\int\limits_{\omega} \tau_k(\xi)\,\tau_l(\xi)\,d\xi = \delta_{kl}. \tag{48.8}$$

Для задачи (48.1), (41.2) в качестве координатной возьмем систему функций

$$\varphi_{kn}(x) = \frac{\sqrt{2}}{\pi}\,\tau_k(\xi)\,\frac{\sin(2n\,\mathrm{arctg}\,x_m)}{\sqrt{1+x_m^2}} \qquad (k,\ n = 1,\ 2,\ \ldots). \tag{48.9}$$

Очевидно, $\varphi_{kn} \in H_{A_1}$. Докажем, что система (48.9) полна в метрике H_{A_1}; для этого достаточно доказать полноту этой системы в метрике (48.4).

Пусть $v \in H_{A_1}$ и $[\varphi_{kn},\ v]_1 = 0$ $(k,\ n = 1,\ 2,\ \ldots)$. Обозначим для краткости

$$\frac{\sin(2n\,\mathrm{arctg}\,x_m)}{\sqrt{1+x_m^2}} = \varphi_n(x_m).$$

Тогда

$$\int\limits_0^\infty dx_m \int\limits_\omega \left[\sum_{j=1}^{m-1} \frac{\partial\tau_k}{\partial x_j}\,\frac{\partial\bar{v}}{\partial x_j}\,\varphi_n(x_m) + \tau_k\,\frac{\partial\bar{v}}{\partial x_m}\,\frac{\partial\varphi_n}{\partial x_m}\right] d\xi = 0$$

$$(k,\ n = 1,\ 2,\ \ldots).$$

Первый внутренний интеграл возьмем по частям. Приняв во внимание тождество

$$-\sum_{j=1}^{m-1} \frac{\partial^2\tau_k}{\partial x_j^2} = \mu_k\tau_k,$$

где μ_k есть k-е собственное число оператора (48.6), получим:

$$\int\limits_\omega \tau_k d\xi \int\limits_0^\infty \left[\mu_k\varphi_n(x_m)\,\bar{v} + \frac{d\varphi_n}{dx_m}\,\frac{\partial\bar{v}}{\partial x_m}\right] dx_m = 0.$$

Система функций $\{\tau_k(\xi)\}$ полна в $L_2(\omega)$, поэтому при любых натуральных n и k и почти при всех ξ справедливо тождество

$$\int\limits_0^\infty \left[\mu_k\varphi_n(x_m)\,\bar{v} + \frac{d\varphi_n}{dx_m}\,\frac{\partial\bar{v}}{\partial x_m}\right] dx_m \equiv 0.$$

Написав последнее тождество для двух различных значений μ_k и вычитая, получим:

$$\int\limits_0^\infty \varphi_n(x_m)\,\bar{v}\,dx_m \equiv 0,$$

и так как система $\{\varphi_n\}$ полна в $L_2(0, \infty)$ (см. § 46), то $v \equiv 0$, что и требовалось доказать.

Метрика (48.4) есть частный случай метрики пространства H_{A_1}, когда $A_{jk} = \delta_{jk}$. Оператор A_1 — положительно определенный (см. § 41), и по следствию 4.2 система (48.9), ортонормированная в метрике $L_2(\Omega)$, сильно минимальна в H_{A_1}.

Заметим, что система (48.9) остается полной и сильно минимальной в метрике H_{A_1} и в том случае, когда уравнение (48.1) заменяется уравнением

$$-\sum_{j, k=1}^{m} \frac{\partial}{\partial x_j}\left(A_{jk} \frac{\partial u}{\partial x_k}\right) + Cu = f(x), \tag{48.10}$$

где C — неотрицательная[1]), измеримая и ограниченная в Ω функция. Это легко вытекает из того факта, что энергетические метрики операторов (48.1), (41.2) и (48.10), (41.2) в этом случае эквивалентны.

Множитель $\frac{2}{\sqrt{\pi}}$ в (48.9) можно опустить.

Если Ω — бесконечный цилиндр (48.2″), то к системе (48.9) следует добавить функции

$$\tau_k(x) \frac{\cos(2n \operatorname{arctg} x_m)}{\sqrt{1 + x_m^2}} \qquad (n = 0, 1, 2, \ldots). \tag{48.11}$$

3. Перейдем теперь к задаче (48.1) — (41.3) и соответствующему оператору A_2; как и в п. 2, допустим, что область ω конечна. В случае полубесконечного цилиндра (48.2′) координатную систему возьмем в виде

$$\varphi_{kn}(x) = \frac{\sqrt{2}}{\pi} \chi_k(\xi) \frac{\cos(2n \operatorname{arctg} x_m)}{\sqrt{1 + x_m^2}} \tag{48.12}$$
$$(k = 1, 2, \ldots; \ n = 0, 1, 2, \ldots),$$

где $\chi_k(\xi)$ суть собственные функции оператора

$$-\sum_{j=1}^{m-1} \frac{\partial^2 \chi}{\partial x_j^2}, \qquad \left.\frac{\partial \chi}{\partial \nu}\right|_s = 0, \tag{48.13}$$

и ν есть нормаль к s; в случае бесконечного цилиндра (48.2″) к функциям (48.12) следует присоединить еще и функции

$$\chi_k(\xi) \frac{\sin(2n \operatorname{arctg} x_m)}{\sqrt{1 + x_m^2}} \qquad (k, n = 1, 2, \ldots). \tag{48.14}$$

---

[1]) Это требование можно ослабить.

Дословно повторив рассуждения п. 2, мы убедимся, что система (48.12) (система (48.12), (48.14) в случае (48.2″)) полна и сильно минимальна в пространстве H_{A_2}.

4. Если область ω бесконечна, то для построения сильно минимальной координатной системы следует в формулах (48.9), (48.11), (48.12) и (48.14) заменить функции $\tau_k(\xi)$ и $\chi_k(\xi)$ функциями $\tilde{\tau}_k(\xi)$ и $\tilde{\chi}_k(\xi)$, которые строятся следующим образом.

Будем считать, что $m \geqslant 3$; обозначим $|\xi| = \rho$ и выполним преобразование $\rho' = \frac{1}{\rho}$; мы предполагаем при этом, что в пространстве координат $x_1, x_2, \ldots, x_{m-1}$ начало координат лежит внутри s. Наше преобразование переведет область ω и ее границу s в конечную область ω' и соответствующую границу s'. Пусть $\tau'_k(\xi')$ и $\chi'_k(\xi')$ суть собственные функции задач Дирихле и Неймана для оператора Лапласа и области ω', и пусть при преобразовании $\rho' = \frac{1}{\rho}$ точке $\xi' \in \omega'$ соответствует точка $\xi \in \omega$. Мы положим тогда

$$\tilde{\tau}_k(\xi) = \frac{1}{\rho^{m-3}} \tau'_k(\xi'), \qquad \tilde{\chi}_k(\xi) = \frac{1}{\rho^{m-3}} \chi'_k(\xi'). \tag{48.15}$$

Построенные таким образом координатные системы полны и сильно минимальны в соответствующих энергетических пространствах.

§ 49. Примеры

Во всех примерах настоящего параграфа приводимые нами координатные системы сильно минимальны в соответствующих энергетических пространствах.

1. Полоса на двумерной плоскости

$$0 < x < 1, \qquad -\infty < y < \infty; \tag{49.1}$$

первая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+y^2}} \sin k\pi x \begin{matrix}\sin\\ \cos\end{matrix} (2n \operatorname{arctg} y) \qquad (k = 1, 2, \ldots; \; n = 0, 1, 2, \ldots).$$

2. Та же полоса; вторая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+y^2}} \cos k\pi x \begin{matrix}\sin\\ \cos\end{matrix} (2n \operatorname{arctg} y) \qquad (k, n = 0, 1, 2, \ldots).$$

3. Полуполоса на двумерной плоскости

$$0 < x < 1, \qquad 0 \leqslant y < \infty; \tag{49.2}$$

первая краевая задача; координатные функции

$$\frac{1}{\sqrt{1+y^2}} \sin k\pi x \sin(2n \operatorname{arctg} y) \qquad (k, n = 1, 2, \ldots).$$

4. Та же полуполоса; вторая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+y^2}}\cos k\pi x\cos(2n\operatorname{arctg}y)\qquad (k,\ n=0,\ 1,\ 2,\ \ldots).$$

Укажем еще сильно минимальные координатные системы для некоторых задач со смешанными краевыми условиями.

5. Полоса (49.1); на прямой $x=0$ задано условие первой краевой задачи, на прямой $x=1$ — условие второй задачи. Координатные функции:

$$\frac{1}{\sqrt{1+y^2}}\sin\frac{2k+1}{2}\pi x\ {\sin \atop \cos}(2n\operatorname{arctg}y)\qquad (k,\ n=0,\ 1,\ 2,\ \ldots).$$

6. Полуполоса (49.2); на прямых $x=0$, $x=1$ заданы условия первой задачи, на отрезке $y=0$, $0\leqslant x\leqslant 1$ — условие второй задачи. Координатные функции:

$$\frac{1}{\sqrt{1+y^2}}\sin k\pi x\cos(2n\operatorname{arctg}y)\qquad (k=1,\ 2,\ \ldots;\ n=0,\ 1,\ 2,\ \ldots).$$

Число примеров подобного рода нетрудно увеличить.

7. Внутренность бесконечной трехмерной прямоугольной призмы

$$0<x<a,\qquad 0<y<b,\qquad -\infty<z<\infty;\tag{49.3}$$

первая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+z^2}}\sin\frac{k\pi x}{a}\sin\frac{l\pi y}{b}\ {\sin \atop \cos}(2n\operatorname{arctg}z)$$
$$(k,\ l=1,\ 2,\ \ldots;\ n=0,\ 1,\ 2,\ \ldots).$$

8. Та же призма; вторая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+z^2}}\cos\frac{k\pi x}{a}\cos\frac{l\pi y}{b}\ {\sin \atop \cos}(2n\operatorname{arctg}z)\qquad (k,\ l,\ n=0,\ 1,\ 2,\ \ldots).$$

Нетрудно указать координатные функции и для некоторых смешанных задач; приведем один пример.

9. Призма (49.3). На гранях $x=0$, $y=0$ заданы условия первой краевой задачи, на гранях $x=a$, $y=b$ — условия второй задачи. Координатные функции:

$$\frac{1}{\sqrt{1+z^2}}\sin\frac{2k+1}{2a}\pi x\sin\frac{2l+1}{2b}\pi y\ {\sin \atop \cos}(2n\operatorname{arctg}z)$$
$$(k,\ l,\ n=0,\ 1,\ 2,\ \ldots).$$

10. Внутренность полубесконечной трехмерной прямоугольной призмы

$$0<x<a,\qquad 0<y<b,\qquad 0<z<\infty;\tag{49.4}$$

первая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+z^2}} \sin\frac{k\pi x}{a} \sin\frac{l\pi y}{b} \sin(2n \operatorname{arctg} z) \qquad (k,\ l,\ n=1,\ 2,\ \ldots).$$

11. Та же призма, вторая краевая задача; координатные функции:

$$\frac{1}{\sqrt{1+z^2}} \cos\frac{k\pi x}{a} \cos\frac{k\pi y}{b} \cos(2n \operatorname{arctg} z) \qquad (k,\ l,\ n=0,\ 1,\ 2,\ \ldots).$$

12. Бесконечный трехмерный круглый цилиндр

$$0<\rho<1, \qquad 0\leqslant\vartheta\leqslant 2\pi, \qquad -\infty<z<\infty; \tag{49.5}$$

первая краевая задача; координатные функции:

$$c_{kln}\frac{1}{\sqrt{1+z^2}} J_k(\gamma_{k,l}\rho) \begin{matrix}\cos\\ \sin\end{matrix} k\vartheta \begin{matrix}\cos\\ \sin\end{matrix}(2n \operatorname{arctg} z)$$
$$(k,\ n=0,\ 1,\ 2,\ \ldots;\ l=1,\ 2,\ \ldots);$$

$\gamma_{k,l}$ означает l-й положительный корень функции $J_k(t)$.

13. Тот же цилиндр, вторая краевая задача. Под знаком функции Бесселя следует заменить $\gamma_{k,l}$ на $\tilde{\gamma}_{k,l}$, где $\tilde{\gamma}_{k,l}$ есть l-й положительный корень функции $J_k'(t)$.

14. Полубесконечный цилиндр

$$0<\rho<1, \qquad 0\leqslant\vartheta\leqslant 2\pi, \qquad 0<z<\infty; \tag{49.6}$$

первая краевая задача; координатные функции:

$$c_{kln}\frac{1}{\sqrt{1+z^2}} J_k(\gamma_{k,l}\rho) \begin{matrix}\cos\\ \sin\end{matrix} k\vartheta \sin(2n \operatorname{arctg} z)$$
$$(k=0,\ 1,\ 2,\ \ldots;\ l,\ n=1,\ 2,\ \ldots).$$

15. Тот же цилиндр, вторая краевая задача; координатные функции:

$$c_{kln}\frac{1}{\sqrt{1+z^2}} J_k(\tilde{\gamma}_{k,l}\rho) \begin{matrix}\cos\\ \sin\end{matrix} k\vartheta \cos(2n \operatorname{arctg} z)$$
$$(k,\ n=0,\ 1,\ 2,\ \ldots;\ l=1,\ 2,\ \ldots).$$

16. Внешность бесконечного трехмерного круглого цилиндра

$$1<\rho<\infty, \qquad 0\leqslant\vartheta\leqslant 2\pi, \qquad -\infty<z<\infty; \tag{49.7}$$

первая краевая задача; координатные функции:

$$c_{kln}\frac{1}{\sqrt{1+z^2}} J_k\left(\frac{\gamma_{k,l}}{\rho}\right) \begin{matrix}\cos\\ \sin\end{matrix} k\vartheta \begin{matrix}\cos\\ \sin\end{matrix}(2n \operatorname{arctg} z)$$
$$(k,\ n=0,\ 1,\ 2,\ \ldots;\ l=1,\ 2,\ \ldots).$$

17. Та же область; вторая краевая задача; координатные функции получаются из предшествующих заменой $\gamma_{k,l}$ на $\tilde{\gamma}_{k,l}$.

18. Полубесконечная область

$$1 < \rho < \infty, \qquad 0 \leqslant \vartheta \leqslant 2\pi, \qquad 0 < z < \infty; \tag{49.8}$$

первая краевая задача; координатные функции:

$$c_{kln} \frac{1}{\sqrt{1+z^2}} J_k\left(\frac{\gamma_{k,l}}{\rho}\right) \begin{matrix}\cos \\ \sin\end{matrix} k\vartheta \sin(2n \operatorname{arctg} z)$$

$$(k = 0,\ 1,\ 2,\ \ldots;\ n,\ l = 1,\ 2,\ \ldots).$$

19. Область (49.8); вторая краевая задача; координатные функции:

$$c_{kln} \frac{1}{\sqrt{1+z^2}} J_k\left(\frac{\tilde{\gamma}_{k,l}}{\rho}\right) \begin{matrix}\cos \\ \sin\end{matrix} k\vartheta \cos(2n \operatorname{arctg} z)$$

$$(k,\ n = 0,\ 1,\ 2,\ \ldots;\ l = 1,\ 2,\ \ldots).$$

В примерах 13—19 нормирующий коэффициент можно определить из условия, чтобы энергетическая норма каждой координатной функции равнялась единице.

Для областей (49.6)—(49.8) легко построить сильно минимальные координатные системы и в некоторых случаях смешанных краевых условий; столь же легко, используя результаты § 33 и § 34, построить сильно минимальные координатные системы для бесконечного или полубесконечного цилиндра, сечение которого представляет собой не круг, а круговой сектор (в частности, полукруг). Очевидно также построение формул, аналогичных формулам настоящего параграфа, в случае областей с числом измерений $m > 3$.

§ 50. Координатные системы для вырождающихся уравнений в конечной области

Мы ограничимся здесь случаем, который описывается следующими допущениями:

а) Ω — цилиндрическая область m-мерного пространства, описываемая соотношениями

$$\xi = (x_1,\ x_2,\ \ldots,\ x_{m-1}) \in \omega, \qquad 0 < x_m < 1, \tag{50.1}$$

где ω — конечная область $(m-1)$-мерного пространства с кусочно гладкой границей s.

б) Ставится задача об интегрировании уравнения (45.1) при краевом условии (45.2) и, если нужно, (45.4).

в) Имеет место двусторонняя оценка

$$\mu_0\left[\psi(x_m)\sum_{k=1}^{m-1}|t_k|^2+\varphi(x_m)|t_m|^2\right]\leqslant$$
$$\leqslant\sum_{j,k=1}^{m}A_{jk}\bar{t}_j t_k\leqslant M_0\left[\psi(x_m)\sum_{k=1}^{m-1}|t_k|^2+\varphi(x_m)|t_m|^2\right]; \qquad (50.2)$$

здесь μ_0 и M_0 — положительные постоянные, $\varphi(t)$ и $\psi(t)$ — функции, непрерывные при $0\leqslant t\leqslant 1$ и положительные при $0<x\leqslant 1$, причем хотя бы одна из величин $\varphi(0)$ и $\psi(0)$ равна нулю.

Чтобы координатная система была полна и сильно минимальна в энергетической метрике нашей задачи, достаточно, чтобы эта система была полна и сильно минимальна в метрике

$$|u|_0^2=\int\limits_\Omega\left\{\psi(x_m)\sum_{k=1}^{m-1}\left|\frac{\partial u}{\partial x_k}\right|^2+\varphi(x_m)\left|\frac{\partial u}{\partial x_m}\right|^2\right\}dx \qquad (50.3)$$

и чтобы функции, входящие в названную систему, удовлетворяли краевым условиям задачи.

Отдельно рассмотрим случаи расходимости и сходимости интеграла (45.3). Если он расходится, то перечисленным выше условиям удовлетворяет, например, система функций

$$\varphi_{kn}(x)=\tau_k(\xi)\int\limits_{x_m}^{1}p_n(t)\,dt, \qquad (50.4)$$

где $\tau_k(\xi)$ суть ортонормированные в $L_2(\omega)$ собственные функции задачи

$$\sum_{j=1}^{m-1}\frac{\partial^2\tau}{\partial x_j^2}+\mu\tau=0, \qquad \tau|_s=0, \qquad (50.5)$$

а $p_n(t)$ — полиномы n-й степени, ортонормированные с весом $\varphi(t)$ на отрезке $0\leqslant t\leqslant 1$. В формуле (50.4) индекс k пробегает значения $k=1, 2, 3, \ldots$, а индекс n — значения $n=0, 1, 2, \ldots$

Докажем это. Прежде всего, функции (50.4), очевидно, удовлетворяют краевым условиям задачи. Далее, как мы сейчас убедимся, система (50.4) полна в метрике (50.3). Действительно, пусть $v(x)$ удовлетворяет краевым условиям задачи, $|v|_0<\infty$ и $[\varphi_{kn}, v]_0=0$ для всех допустимых значений k и n. Полагая

$$\int\limits_x^1 p_n(t)\,dt=q_n(x),$$

имеем:

$$\int_0^1 dx_m \int_\omega \left\{ \psi(x_m) \sum_{j=1}^{m-1} \frac{\partial \tau_k}{\partial x_j} \frac{\partial \bar{v}}{\partial x_j} q_n + \varphi(x_m) \tau_k \bar{v} p_n \right\} d\xi = 0,$$

или

$$\int_\omega \tau_k \, d\xi \int_0^1 \{\mu_k \psi(x_m) q_n(x_m) + \varphi(x_m) p_n\} \bar{v} \, dx_m = 0.$$

Поэтому при всех допустимых k и n

$$\int_0^1 \{\mu_k \psi(x_m) q_n(x_m) + \varphi(x_m) p_n\} \bar{v} \, dx_m = 0$$

и, следовательно,

$$\int_0^1 \varphi(x_m) p_n(x_m) \bar{v} \, dx_m \equiv 0. \tag{50.6}$$

Последнее равенство имеет место при всех значениях n; в силу полноты системы полиномов p_n имеем $v \equiv 0$.

Система (50.4) ортонормирована и потому сильно минимальна в метрике

$$|u|_1^2 = \int_\Omega \varphi(x_m) \left| \frac{\partial u}{\partial x_m} \right|^2 dx.$$

Но $|u|_0 \geqslant |u|_1$; по теореме 4.2 упомянутая система сильно минимальна в метрике (50.3), что и требовалось доказать.

Пусть теперь интеграл (45.3) сходится. В этом случае систему, полную и сильно минимальную в H_A, можно задать той же формулой (50.4), если по-иному определить полиномы $p_n(t)$: это, по-прежнему, будут полиномы степени n, но подчиненные условиям

$$\int_0^1 p_n(t) \, dt = 0 \qquad (n = 1, 2, 3, \ldots), \tag{50.7}$$

$$\int_0^1 \varphi(t) p_r(t) p_s(t) \, dt = \delta_{rs} \qquad (r, s = 1, 2, 3, \ldots); \tag{50.8}$$

полином $p_0(t)$ мы в рассмотрение не вводим. Нетрудно убедиться, что последовательность полиномов, удовлетворяющих условиям (50.7) и (50.8), построить можно.

Будем считать, что в системе (50.4) теперь $k, n = 1, 2, 3, \ldots$

В силу соотношения (50.7) функции (50.4) удовлетворяют не только условию (45.2), но и условию (45.4), — короче, они удовлетворяют всем краевым условиям задачи. Система (50.4) сильно минимальна в метрике (50.3) — это доказывается так же, как и выше; докажем, что упомянутая система в метрике (50.3) полна. Пусть функция $v(x)$ удовлетворяет всем краевым условиям задачи, и пусть еще

$$|v|_0 < \infty, \qquad [\varphi_{kn}, v]_0 = 0 \qquad (k, n = 1, 2, 3, \ldots).$$

Предшествующие рассуждения приведут нас к тождеству (50.6), верному для $n = 1, 2, 3, \ldots$ Это тождество можно понимать так: для почти всех $\xi \in \omega$ функция $\varphi(x_m) v(x)$, рассматриваемая как функция только от x_m, ортогональна в метрике $L_2(0, 1)$ к любому полиному ненулевой степени. Но тогда функция $\varphi(x_m) v(x)$ не зависит от x_m и функция $v(x)$ имеет вид

$$v(x) = \frac{c(\xi)}{\varphi(x_m)}.$$

Но

$$v(x)\big|_{x_m=1} = \frac{c(\xi)}{\varphi(1)} = 0,$$

поэтому $c(\xi) \equiv 0$ и $v(x) \equiv 0$, что и требовалось доказать.

Если $\varphi(0) > 0$, то можно принять $\varphi(x_m) \equiv 1$; в этом случае

$$p_n(t) = \sqrt{2n+1}\, P_n(2t-1),$$

где P_n — полиномы Лежандра.

ГЛАВА VII

УСТОЙЧИВОСТЬ ПРОЦЕССА РИТЦА В ЗАДАЧАХ О СПЕКТРЕ

В этой небольшой главе будут выяснены условия устойчивости процесса Ритца, примененного для приближенного вычисления собственных чисел и собственных элементов (точнее, собственных подпространств) положительно определенного оператора. Будет дано и определение соответствующих понятий устойчивости; это надо сделать уже потому, что задачи о спектре (даже линейного оператора) суть задачи нелинейные, и определения устойчивости, данные в предшествующих главах применительно к линейным задачам, могут оказаться непригодными.

Устойчивость процесса Ритца в задачах о спектре была исследована в работе Л. Н. Довбыш [2], результаты которой мы здесь и излагаем с некоторыми небольшими изменениями.

§ 51. Общая теорема

Лемма 51.1 [1]). *Пусть S и T — самосопряженные неотрицательные вполне непрерывные операторы, действующие в некотором гильбертовом пространстве $\mathfrak{H}$, и пусть σ_k^S и σ_k^T — их собственные числа, занумерованные в порядке убывания. Тогда*

$$|\sigma_k^S - \sigma_k^T| \leqslant \|S - T\|. \tag{51.1}$$

Положим

$$\left.\begin{aligned} \sigma^S(v_1, v_2, \ldots, v_{k-1}) &= \max(Su, u), \\ \sigma^T(v_1, v_2, \ldots, v_{k-1}) &= \max(Tu, u), \end{aligned}\right\} \tag{51.2}$$

оба раза максимум берется по множеству элементов $u \in \mathfrak{H}$ таких, что $\|u\| = 1$ и $(u, v_j) = 0$ $(j = 1, 2, \ldots, k-1)$. В силу минимаксимального принципа,

$$\sigma_k^S = \min \sigma^S(v_1, v_2, \ldots, v_{k-1}), \tag{51.3}$$

$$\sigma_k^T = \min \sigma^T(v_1, v_2, \ldots, v_{k-1}), \tag{51.4}$$

---

[1]) Лемма 51.1 является частным случаем известной леммы Г. Вейля; см., например, Ф. Рисс и Б. Секефальви-Надь [1].

где минимум берется по всевозможным наборам элементов v_1, $v_2, \ldots, v_{k-1}$, принадлежащих пространству $\mathfrak{H}$.

Для любых вещественных переменных x и y справедливо неравенство $\sup(x+y) \leqslant \sup x + \sup y$. Положив $S = T + Q$, найдем отсюда:

$$\sigma^S(v_1, v_2, \ldots, v_{k-1}) - \sigma^T(v_1, v_2, \ldots, v_{k-1}) \leqslant$$
$$\leqslant \sup_{\substack{\|u\|=1 \\ (u, v_j)=0}} (Qu, u) \leqslant \sup_{\|u\|=1} |(Qu, u)| = \|Q\| = \|S-T\|.$$

Поменяв T и S местами, найдем аналогично:

$$\sigma^T(v_1, v_2, \ldots, v_{k-1}) - \sigma^S(v_1, v_2, \ldots, v_{k-1}) \leqslant \|S-T\|$$

и, окончательно,

$$|\sigma^S(v_1, v_2, \ldots, v_{k-1}) - \sigma^T(v_1, v_2, \ldots, v_{k-1})| \leqslant \|S-T\|. \qquad (51.5)$$

Пусть, например, $\sigma_k^T \leqslant \sigma_k^S$. В формуле (51.5) положим $v_j = v_j^0$ $(j = 1, 2, \ldots, k-1)$, где v_j^0 — те элементы, на которых достигается минимум в формуле (51.4). Формула (51.5) принимает вид

$$\sigma^S\left(v_1^0, v_2^0, \ldots, v_{k-1}^0\right) - \sigma_k^T \leqslant \|S-T\|.$$

Но, в силу формулы (51.3), $\sigma_k^S \leqslant \sigma^S\left(v_1^0, v_2^0, \ldots, v_{k-1}^0\right)$ и, следовательно, $\sigma_k^S - \sigma_k^T \leqslant \|S-T\|$. Это совпадает с неравенством (51.1), так как $\sigma_k^S \geqslant \sigma_k^T$.

Рассмотрим теперь последовательность гильбертовых пространств $\mathfrak{H}_n$ $(n = 1, 2, 3, \ldots)$. В пространстве $\mathfrak{H}_n$ зададим два самосопряженных положительно определенных оператора A_n и B_n таких, что операторы A_n^{-1} и $\bar{B}_n = A_n^{-\frac{1}{2}} B_n A_n^{-\frac{1}{2}}$ вполне непрерывны в $\mathfrak{H}_n$; так как

$$A_n^{-\frac{1}{2}} B_n A_n^{-\frac{1}{2}} = A_n^{-\frac{1}{2}} B_n^{\frac{1}{2}} B_n^{\frac{1}{2}} A_n^{-\frac{1}{2}} = \left(B_n^{\frac{1}{2}} A_n^{-\frac{1}{2}}\right)^* \left(B_n^{\frac{1}{2}} A_n^{-\frac{1}{2}}\right),$$

то оператор $A_n^{-\frac{1}{2}} B_n A_n^{-\frac{1}{2}}$ вполне непрерывен, если вполне непрерывно произведение $B_n^{\frac{1}{2}} A_n^{-\frac{1}{2}}$. Легко доказать, что для этого, в свою очередь, необходимо и достаточно, чтобы всякое множество, ограниченное в энергетической метрике оператора A_n, было компактно в энергетической метрике оператора B_n. Обозначим через $\sigma_k^{(n)}$ k-е (в порядке возрастания) собственное число уравнения

$$(A_n - \sigma B_n) u_n = 0. \qquad (51.6)$$

Наряду с уравнением (51.6) рассмотрим уравнение

$$[(A_n+\Gamma_n)-\mu(B_n+\Delta_n)]\,v_n=0, \tag{51.7}$$

где Γ_n и Δ_n — ограниченные самосопряженные операторы, действующие в пространстве $\mathfrak{H}_n$.

Будем говорить, что процесс вычисления k-го собственного числа уравнения (51.6) (короче, «процесс (51.6)») устойчив, если существуют такие независящие от n постоянные (они могут зависеть от k) p, q, r, что при $\|\Gamma_n\|\leqslant r$ и при любом Δ_n таком, что сумма $B_n+\Delta_n$ остается положительно определенной, выполнены следующие требования: 1) операторы $A_n+\Gamma_n$ — положительно определенные; 2) операторы

$$T_n=(I_n+\bar{\Gamma}_n)^{-\frac{1}{2}}(\bar{B}_n+\bar{\Delta}_n)(I_n+\bar{\Gamma}_n)^{-\frac{1}{2}}$$

вполне непрерывны в соответствующих пространствах $\mathfrak{H}_n$; здесь I_n — тождественный оператор в $\mathfrak{H}_n$,

$$\bar{\Gamma}_n=A_n^{-\frac{1}{2}}\Gamma_nA_n^{-\frac{1}{2}},\qquad \bar{\Delta}_n=A_n^{-\frac{1}{2}}\Delta_nA_n^{-\frac{1}{2}};$$

3) если $\mu_k^{(n)}$ означает k-е, в порядке возрастания, собственное число уравнения (51.7), то

$$\left|\frac{\sigma_k^{(n)}}{\mu_k^{(n)}}-1\right|\leqslant p\|\Gamma_n\|+q\|\Delta_n\|. \tag{51.8}$$

Теорема 51.1. *Для того чтобы процесс* (51.6) *был устойчив, необходимо, чтобы*

$$\|A_n^{-1}\|\leqslant C_1, \tag{51.9}$$

и достаточно, чтобы были выполнены условия (51.9) *и*

$$\sigma_k^{(n)}\leqslant C_2,\qquad \frac{(B_nu,\,u)}{(A_nu,\,u)}\leqslant C_3,\qquad u\in\mathfrak{H}_n, \tag{51.10}$$

где постоянные C_1, C_2, C_3 не зависят от n[1]).

Необходимость условия (51.9). Если процесс (51.6) устойчив, то при $\|\Gamma_n\|\leqslant r$ оператор $A_n+\Gamma_n$ положительно определен. Положим $\Gamma_n=-rI_n$. Тогда во всяком случае

$$((A_n-rI_n)\,u,\;u)\geqslant 0,\qquad u\in\mathfrak{H}_n.$$

Отсюда $(A_nu,\;u)\geqslant r\|u\|^2$. Следовательно (ПМ, стр. 21, формула (5)), $\|A_n^{-1}\|\leqslant\frac{1}{r}$.

[1]) Постоянная C_2 может зависеть от k.

Достаточность условий (51.9) и (51.10). Выберем постоянную β, $0<\beta<1$, положим $r=\frac{\beta}{C_1}$ и потребуем, чтобы $\|\Gamma_n\|\leqslant r$. По только что упомянутой формуле ((5), стр. 21, ПМ) нижняя грань γ_n^2 положительно определенного оператора A_n равна $\frac{1}{\|A_n^{-1}\|}$; по неравенству (51.9) $\gamma_n^2\geqslant\frac{1}{C_1}$ и, следовательно,

$$((A_n+\Gamma_n)u,\ u)=(A_nu,\ u)+(\Gamma_nu,\ u)\geqslant$$
$$\geqslant(\gamma_n^2-\|\Gamma_n\|)\|u\|^2\geqslant\frac{1-\beta}{C_1}\|u\|^2.$$

Таким образом, оператор $A_n+\Gamma_n$ — положительно определенный.

Теперь очевидно, что введенный выше оператор T_n неотрицателен в пространстве $\mathfrak{H}_n$. Докажем, что в этом пространстве оператор T_n также и вполне непрерывен. Из спектрального разложения легко вытекает формула

$$\left\|A_n^{-\frac{1}{2}}\right\|=\|A_n^{-1}\|^{\frac{1}{2}}=\frac{1}{\gamma_n},$$

отсюда $\left\|A_n^{-\frac{1}{2}}\right\|\leqslant\sqrt{C_1}$. Теперь $\|\overline{\Gamma}_n\|\leqslant\|\Gamma_n\|\left\|A_n^{-\frac{1}{2}}\right\|^2\leqslant\beta<1$, оператор $I_n+\overline{\Gamma}_n$ положительно определенный, а оператор $(I_n+\overline{\Gamma}_n)^{-\frac{1}{2}}$ существует и ограничен; при этом

$$\left\|(I_n+\overline{\Gamma}_n)^{-\frac{1}{2}}\right\|=\left\|\sum_{m=0}^{\infty}(-1)^m\binom{m}{\frac{1}{2}}\overline{\Gamma}_n^m\right\|\leqslant$$
$$\leqslant\sum_{m=0}^{\infty}\binom{m}{\frac{1}{2}}\|\overline{\Gamma}_n\|^m\leqslant\sum_{m=0}^{\infty}\binom{m}{\frac{1}{2}}\beta^m=\frac{1}{\sqrt{1-\beta}}.\qquad(51.11)$$

Оператор $\overline{B}_n$ вполне непрерывен по условию, а оператор $\overline{\Delta}_n$ — как произведение из двух вполне непрерывных и одного ограниченного. Полная непрерывность оператора T_n теперь очевидна.

Займемся теперь установлением оценки (51.8). Полагая $A_n^{\frac{1}{2}}u=v$, имеем:

$$\frac{(B_nu,\ u)}{(A_nu,\ u)}=\frac{\left(B_nA_n^{-\frac{1}{2}}v,\ A_n^{-\frac{1}{2}}v\right)}{\|v\|^2}=\frac{(\overline{B}_nv,\ v)}{\|v\|^2}.$$

Оператор $\overline{B}_n$ самосопряженный, поэтому

$$\|\overline{B}_n\|=\sup\frac{(\overline{B}_nv,\ v)}{\|v\|^2},$$

и второе неравенство (51.10) показывает, что

$$\|\bar{B}_n\| \leqslant C_3. \tag{51.12}$$

Замена

$$A_n^{\frac{1}{2}} u_n = \tilde{u}_n, \qquad (I_n + \bar{\Gamma}_n)^{\frac{1}{2}} A_n^{\frac{1}{2}} v_n = \tilde{v}_n$$

приводит уравнения (51.6) и (51.7) к виду

$$(I_n - \lambda \bar{B}_n)\tilde{u}_n = 0, \qquad (I_n - \mu T_n)\tilde{v}_n = 0;$$

отсюда видно, что $\frac{1}{\sigma_k^{(n)}}$ и $\frac{1}{\mu_k^{(n)}}$ суть собственные числа самосопряженных вполне непрерывных операторов $\bar{B}_n$ и T_n соответственно. По лемме 51.1

$$\left|\frac{1}{\mu_k^{(n)}} - \frac{1}{\sigma_k^{(n)}}\right| \leqslant \|T_n - \bar{B}_n\|.$$

Далее,

$$\begin{aligned} \| T_n - \bar{B}_n\| &= \left\|(I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} \bar{B}_n (I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} - \bar{B}_n + \right. \\ &\qquad \left. + (I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} \bar{\Delta}_n (I_n + \bar{\Gamma}_n)^{-\frac{1}{2}}\right\| \leqslant \\ &\leqslant \left\|\left[(I_n + \Gamma_n)^{-\frac{1}{2}} - I_n\right] \bar{B}_n (I_n + \Gamma_n)^{-\frac{1}{2}}\right\| + \\ &+ \left\|\bar{B}_n\left[(I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} - I_n\right]\right\| + \left\|(I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} \bar{\Delta}_n (I_n + \bar{\Gamma}_n)^{-\frac{1}{2}}\right\| \leqslant \\ &\leqslant \frac{C_3(1 + \sqrt{1-\beta})}{\sqrt{1-\beta}} \left\|(I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} - I_n\right\| + \frac{C_1\|\Delta_n\|}{1-\beta}. \end{aligned}$$

Но

$$\begin{aligned} \left\|(I_n + \bar{\Gamma}_n)^{-\frac{1}{2}} - I_n\right\| &= \left\|\sum_{m=1}^{\infty} \binom{m}{\frac{1}{2}} \bar{\Gamma}_n^m\right\| \leqslant \|\bar{\Gamma}_n\| \left\|\sum_{m=1}^{\infty} \binom{m}{\frac{1}{2}} \bar{\Gamma}_n^{m-1}\right\| \leqslant \\ &\leqslant C_1\|\Gamma_n\| \sum_{m=1}^{\infty} \binom{m}{\frac{1}{2}} \beta^{m-1} = \frac{C_1\|\Gamma_n\|}{\sqrt{1-\beta}(1+\sqrt{1-\beta})}, \end{aligned}$$

и мы приходим к неравенству

$$\left|\frac{1}{\mu_k^{(n)}} - \frac{1}{\sigma_k^{(n)}}\right| \leqslant \frac{C_1 C_3}{1-\beta} \|\Gamma_n\| + \frac{C_1}{1-\beta} \|\Delta_n\|.$$

Умножая это на $\sigma_k^{(n)}$ и вспоминая, что $\sigma_k^{(n)} \leqslant C_2$, получаем окончательно:

$$\left|\frac{\sigma_k^{(n)}}{\mu_k^{(n)}} - 1\right| \leqslant \frac{C_1 C_2}{1-\beta} (C_3\|\Gamma_n\| + \|\Delta_n\|). \tag{51.13}$$

Теорема доказана. Как видно из формулы (51.13), можно положить

$$p=\frac{C_1C_2C_3}{1-\beta}, \qquad q=\frac{C_1C_2}{1-\beta}.$$

§ 52. Об устойчивости процесса Ритца в задаче о собственных числах

Пусть A — положительно определенный самосопряженный оператор с дискретным спектром, действующий в сепарабельном гильбертовом пространстве H, и пусть λ_k $(k=1, 2, 3, \ldots)$ — собственные числа оператора A, занумерованные в порядке возрастания. Для приближенного вычисления чисел λ_k можно использовать процесс Ритца, сущность которого сводится (см. ВМ, § 33) к следующему: выбираем координатную систему $\{\varphi_n\}$, удовлетворяющую условиям 1)—3) § 6, составляем матрицы

$$R_n=\begin{Vmatrix} [\varphi_1, \varphi_1] & [\varphi_2, \varphi_1] & \ldots & [\varphi_n, \varphi_1] \\ [\varphi_1, \varphi_2] & [\varphi_2, \varphi_2] & \ldots & [\varphi_n, \varphi_2] \\ \cdot & \cdot & \cdot & \cdot \\ [\varphi_1, \varphi_n] & [\varphi_2, \varphi_n] & \ldots & [\varphi_n, \varphi_n] \end{Vmatrix} \tag{52.1}$$

и

$$r_n=\begin{Vmatrix} (\varphi_1, \varphi_1) & (\varphi_2, \varphi_1) & \ldots & (\varphi_n, \varphi_1) \\ (\varphi_1, \varphi_2) & (\varphi_2, \varphi_2) & \ldots & (\varphi_n, \varphi_2) \\ \cdot & \cdot & \cdot & \cdot \\ (\varphi_1, \varphi_n) & (\varphi_2, \varphi_n) & \ldots & (\varphi_n, \varphi_n) \end{Vmatrix}. \tag{52.2}$$

Приближенный собственный элемент

$$u_n=\sum_{k=1}^{n} a_k^{(n)}\varphi_n$$

и приближенное собственное число $\lambda_k^{(n)}$ определяются из уравнения

$$(R_n-\sigma r_n)\,a^{(n)}=0, \tag{52.3}$$

где

$$a^{(n)}=\left(a_1^{(n)},\ a_2^{(n)},\ \ldots,\ a_n^{(n)}\right).$$

Мы оказались в условиях общей схемы § 51: в нашем случае пространство $\mathfrak{H}_n$ есть n-мерное унитарное пространство E_n, $A_n=R_n$, $B_n=r_n$, где на этот раз R_n и r_n суть операторы, порожденные в пространстве E_n матрицами (52.1) и (52.2) соответственно.

Будучи порождены матрицами Грама линейно независимых элементов, эти операторы — положительно определенные. Вопрос о полной непрерывности здесь не возникает, так как в конечномерном пространстве E_n любой линейный оператор вполне непрерывен.

Теорема 52.1. *Для того чтобы процесс вычисления по Ритцу первого собственного числа самосопряженного положительно определенного оператора был устойчив, необходимо и достаточно, чтобы координатная система была сильно минимальна в соответствующей энергетической метрике*[1]).

Если процесс устойчив, то, по теореме 51.1, $\|R_n^{-1}\| \leqslant C_1$. Но $\|R_n^{-1}\| = \frac{1}{\lambda_1^{(n)}}$, где $\lambda_1^{(n)}$ — наименьшее собственное число матрицы R_n, поэтому $\lambda_1^{(n)} \geqslant \frac{1}{C_1}$ и координатная система сильно минимальна в энергетической метрике оператора A.

Доказательство достаточности сводится к проверке условий (51.9) и (51.10) для $k=1$.

Если координатная система сильно минимальна в H_A, то существует такая постоянная $\lambda_0 > 0$, что $\lambda_1^{(n)} \geqslant \lambda_0$. Но тогда $\|R_n^{-1}\| = \frac{1}{\lambda_1^{(n)}} \leqslant \frac{1}{\lambda_0}$ и можно положить $C_1 = \frac{1}{\lambda_0}$. Далее, в рассматриваемом случае $\sigma_1^{(n)}$ есть собственное число уравнения (52.3). При возрастающем n величина $\sigma_1^{(n)}$ не возрастает (ВМ, § 32), поэтому $\sigma_1^{(n)} \leqslant \sigma_1^{(k)}$, и можно положить $C_3 = \sigma_1^{(k)}$.

Проверим, наконец, второе из условий (51.9). Оператор A — положительно определенный, поэтому существует такая положительная постоянная γ, что $|u|_A \geqslant \gamma \|u\|$, $u \in H_A$. Пусть $a^{(n)} = \left(a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}\right)$ — произвольный вектор из E_n. Положим

$$u_n = \sum_{k=1}^{n} a_k^{(n)} \varphi_k;$$

очевидно, $u_n \in H_A$. Имеем:

$$\frac{(r_n a^{(n)}, a^{(n)})}{(R_n a^{(n)}, a^{(n)})} = \frac{\|u_n\|^2}{|u_n|_A^2} \leqslant \frac{1}{\gamma^2},$$

и можно положить $C_2 = \frac{1}{\gamma^2}$. Все условия теоремы 51.1 выполнены, и теорема 52.1 доказана.

Пользуясь формулой (51.13), легко дать оценку погрешности «неточного» приближения по Ритцу. Именно, пусть Γ_n и Δ_n суть ошибки матриц R_n и r_n соответственно и пусть $\|\Gamma_n\| \leqslant \frac{\beta}{C_1} = \beta\lambda_0$.

---

[1]) Эту теорему доказала в статье [2] Л. Н. Довбыш, исходившая из других соображений.

Тогда по формуле (51.13)

$$\left|\frac{\sigma_k^{(n)}}{\mu_k^{(n)}}-1\right| \leqslant \frac{\sigma_k^{(k)}}{\lambda_0(1-\beta)}\left(\frac{\|\Gamma_n\|}{\gamma^2}+\|\Delta_n\|\right); \tag{52.4}$$

здесь $\sigma_k^{(n)}$ — точное, а $\mu_k^{(n)}$ — неточное n-е приближение по Ритцу к k-му собственному числу оператора A.

Для $k > 1$ Л. Н. Довбыш [2] доказала следующую теорему.

Теорема 52.2. *Для устойчивости процесса вычисления по Ритцу k-го ($k > 1$) собственного числа положительно определенного оператора достаточно, а в классе минимальных систем и необходимо, чтобы координатная система была сильно минимальна в соответствующей энергетической метрике.*

Достаточность можно доказать, как в теореме 52.1; необходимость мы доказывать не будем. Оговорка о классе минимальных систем необходима: примеры показывают, что при $k > 1$ процесс Ритца может оказаться устойчивым и для координатной системы, неминимальной в энергетической метрике.

§ 53. Об устойчивости процесса Ритца в задаче о собственных подпространствах

Пусть A — положительно определенный оператор с дискретным спектром и пусть $\sigma_k = \sigma_{k+1} = \ldots = \sigma_{k+l-1}$ — собственное число кратности l и $H^{(k)}$, соответствующее этому числу, — собственное подпространство оператора A. Пусть $\sigma_{k+j}^{(n)}$ ($j = 0, 1, \ldots, l-1$) — собственные числа уравнения (52.3), а $\mu_{k+j}^{(n)}$ — собственные числа уравнения

$$[(R_n+\Gamma_n)-\mu(r_n+\Delta_n)]\, b^{(n)}=0, \tag{53.1}$$

где Γ_n и Δ_n — ошибки матриц R_n и r_n. Обозначим через $E_{k,l}^{(n)}$ ортогональную сумму собственных подпространств уравнения (52.3), отвечающих собственным числам $\sigma_k^{(n)}, \sigma_{k+1}^{(n)}, \ldots, \sigma_{k+l-1}^{(n)}$, и через $F_{k,l}^{(n)}$ — аналогичную сумму для уравнения (53.1) и собственных чисел $\mu_k^{(n)}, \mu_{k+1}^{(n)}, \ldots, \mu_{k+l-1}^{(n)}$. В пространстве $E^{(n)}$ введем новую метрику

$$\left[\alpha^{(n)}, \beta^{(n)}\right]=\left(R_n\alpha^{(n)}, \beta^{(n)}\right), \qquad \left|\alpha^{(n)}\right|=\sqrt{\left[\alpha^{(n)}, \alpha^{(n)}\right]}, \tag{53.2}$$

и в этой метрике определим угол раствора θ_n подпространств $E_{k,l}^{(n)}$ и $F_{k,l}^{(n)}$

$$\cos\theta_n=\sup\frac{\left|\left[\alpha^{(n)}, \beta^{(n)}\right]\right|}{\left|\alpha^{(n)}\right|\left|\beta^{(n)}\right|}; \tag{53.3}$$

точная верхняя граница берется по всевозможным отличным от нуля векторам $\alpha^{(n)} \in \widetilde{E}_{k,l}^{(n)}$ и $\beta^{(n)} \in \widetilde{F}_{k,l}^{(n)}$, где $\widetilde{E}_{k,l}^{(n)}$ и $\widetilde{F}_{k,l}^{(n)}$ суть ортогональные дополнения подпространств $E_{k,l}^{(n)}$ и $F_{k,l}^{(n)}$ к их общему подпространству. Будем говорить, что процесс Ритца определения подпространства $E_{k,l}^{(n)}$ устойчив, если при $\|\Gamma_n\| \to 0$ и $\|\Delta_n\| \to 0$ угол θ_n стремится к нулю равномерно относительно n, иначе говоря, если по данному числу $\varepsilon > 0$ можно найти такие числа $r > 0$ и $s > 0$, что из неравенств $\|\Gamma_n\| \leqslant r$, $\|\Delta_n\| \leqslant s$ вытекает неравенство $\theta < \varepsilon$.

В статье [2] Л. Н. Довбыш доказала следующую теорему: *Процесс определения подпространства $E_{k,l}^{(n)}$ устойчив, если координатная система сильно минимальна в H_A.*

Доказательство мы проведем здесь только для простейшего случая, когда $H^{(n)} = E_{1,1}^{(n)}$ есть собственное подпространство оператора A, соответствующее его наименьшему собственному числу λ_1, причем это собственное число — простое.

Пусть $\|\Gamma_n\| \leqslant r$, $\|\Delta_n\| \leqslant s$; числа r и s будем считать столь малыми, чтобы процесс определения по Ритцу собственного числа λ_1 был устойчивым. При n достаточно большом наименьшие собственные числа $\sigma_1^{(n)}$ и $\mu_1^{(n)}$ обеих приближенных задач — «точной» и «неточной» — будут простыми. Пусть им соответствуют (одномерные) собственные подпространства $\Xi_1^{(n)}$ и $M_1^{(n)}$.

Натянем на $\Xi_1^{(n)}$ и $M_1^{(n)}$ двумерное пространство P и обозначим через ψ_n угол (в смысле метрики, порождаемой матрицей R_n) между направлением $\Xi_1^{(n)}$ и произвольным вектором $x \in P$:

$$\cos\psi_n = \frac{(R_n x, \xi_n)}{\sqrt{(R_n x, x)\,(R_n \xi_n, \xi_n)}},$$

где $\xi_n \in \Xi_1^{(n)}$ — собственный вектор уравнения $R_n\xi - \sigma_1^{(n)} r_n \xi = 0$. Вектор ξ_n можно считать единичным: $(R_n\xi_n, \xi_n) = 1$, и тогда

$$\cos\psi_n = \frac{(R_n x, \xi_n)}{\sqrt{(R_n x, x)}}. \tag{53.4}$$

Докажем, что для любого вектора $x \in P$ соотношение

$$\frac{(r_n x, x)}{(R_n x, x)} = a_n \sin^2\psi_n + b_n \cos^2\psi_n, \tag{53.5}$$

где величины a_n и b_n не зависят от вектора x. В плоскости P возьмем вектор η_n, единичный и ортогональный (в метрике матрицы R_n) к вектору ξ_n: $(R_n\eta_n, \eta_n) = 1$, $(R_n\xi_n, \eta_n) = 0$.

Тогда $x = (\xi_n \cos\psi_n + \eta_n \sin\psi_n)\,|x|$;

$$(r_n x, x) = [(r_n\xi_n, \xi_n)\cos^2\psi_n + (r_n\eta_n, \eta_n)\sin^2\psi_n + 2\,\mathrm{Re}\,(r_n\xi_n, \eta_n)\cos\psi_n \sin\psi_n]\,|x|^2.$$

Но

$$(r_n\xi_n, \eta_n) = \frac{1}{\sigma_1^{(n)}}(R_n\xi_n, \eta_n) = 0,$$

и мы приходим к формуле (53.5), в которой

$$a_n = (r_n\eta_n, \eta_n), \qquad b_n = (r_n\xi_n, \xi_n). \tag{53.6}$$

Из формул (53.6) и из определения векторов ξ_n и η_n сразу следует, что

$$b_n = \frac{1}{\sigma_1^{(n)}}, \qquad a_n \leqslant \frac{1}{\sigma_2^{(n)}}. \tag{53.7}$$

Далее,

$$b_n - a_n \geqslant \frac{1}{\sigma_1^{(n)}} - \frac{1}{\sigma_2^{(n)}} \underset{n\to\infty}{\longrightarrow} \frac{1}{\lambda_1} - \frac{1}{\lambda_2},$$

где λ_2 — второе собственное число оператора A. Так как собственное число λ_1 — простое, то $\frac{1}{\lambda_1} - \frac{1}{\lambda_2} > 0$ и, следовательно, существует такая постоянная $\beta > 0$, что при n достаточно большом

$$b_n - a_n \geqslant \beta. \tag{53.8}$$

На плоскости P рассмотрим отношение

$$\frac{((r_n + \Delta_n)x, x)}{((R_n + \Gamma_n)x, x)} = \frac{(a_n \sin^2\psi_n + b_n\cos^2\psi_n) + \dfrac{(\Delta_n x, x)}{(R_n x, x)}}{1 + \dfrac{(\Gamma_n x, x)}{(R_n x, x)}}. \tag{53.9}$$

Отношения в числителе и знаменателе не зависят от $|x|$. Обозначим

$$\frac{(\Gamma_n x, x)}{(R_n x, x)} = \tau_1^{(n)}(\psi_n), \qquad \frac{(\Delta_n x, x)}{(R_n x, x)} = \tau_2^{(n)}(\psi_n),$$

так что

$$\frac{((r_n + \Delta_n)x, x)}{((R_n + \Gamma_n)x, x)} = \frac{a_n + (b_n - a_n)\cos^2\psi_n + \tau_2^{(n)}(\psi_n)}{1 + \tau_1^{(n)}(\psi_n)}. \tag{53.9$'$}$$

Абсолютный максимум отношения (53.9) достигается на собственном векторе «неточной» задачи, соответствующем собственному числу этой задачи $\mu_1^{(n)}$.

Заметим, что $x \in E_n$, так как пространству E_n принадлежат оба подпространства $\Xi_1^{(n)}$ и $M_1^{(n)}$. Отсюда

$$|(\Gamma_n x, x)| \leqslant \|\Gamma_n\| \|x\|^2, \qquad |(R_n x, x)| \geqslant \tilde{\lambda} \|x\|^2,$$

где $\|x\| = \|x\|_{E_n}$, а $\tilde{\lambda}$ — положительная нижняя граница собственных чисел матриц R_n. Теперь

$$|\tau_1^{(n)}(\psi_n)| \leqslant \frac{\|\Gamma_n\|}{\tilde{\lambda}} \leqslant \frac{r}{\tilde{\lambda}}. \tag{53.10}$$

Аналогично

$$|\tau_2^{(n)}(\psi_n)| \leqslant \frac{\|\Delta_n\|}{\tilde{\lambda}} \leqslant \frac{s}{\tilde{\lambda}}. \tag{53.11}$$

Правую часть формулы (53.9) обозначим через $\Omega(\psi_n)$, так что

$$\Omega(\psi)=\frac{a_n+(b_n-a_n)\cos^2\psi+\tau_2^{(n)}(\psi)}{1+\tau_1^{(n)}(\psi)}.$$

Если $\tau_1^{(n)}(\psi)=\tau_2^{(n)}(\psi)\equiv 0$, т. е. если речь идет о «точной» задаче, то в силу соотношения (53.8) функция $\Omega(\psi)$ достигает максимума при $\psi=0$. Докажем теперь, что при малых $\tau_1^{(n)}$ и $\tau_2^{(n)}$ функция $\Omega(\psi)$ достигает абсолютного максимума в окрестности точки $\psi=0$. Более определенно докажем следующее: каково бы ни было число $\varepsilon>0$, можно найти такие числа r и s, что если $\|\Gamma_n\|\leqslant r$ и $\|\Delta_n\|\leqslant s$, то функция $\Omega(\psi)$ достигает абсолютного максимума в окрестности значения $\psi=0$, определяемой неравенством $\sin^2\psi<\varepsilon$.

Из формул (53.10) и (53.11) вытекает, что

$$\Omega(\psi)\geqslant\frac{a_n+(b_n-a_n)\cos^2\psi-\dfrac{s}{\tilde\lambda}}{1+\dfrac{r}{\tilde\lambda}}.$$

Отсюда, обозначая символом max абсолютный максимум, имеем:

$$\max\Omega(\psi)\geqslant\Omega(0)\geqslant\frac{b_n-\dfrac{s}{\tilde\lambda}}{1+\dfrac{r}{\tilde\lambda}}.\tag{53.12}$$

Пусть $\max\Omega(\psi)=\Omega(\psi_0)$, где $\sin^2\psi_0\geqslant\varepsilon$. Тогда $\cos^2\psi_0\leqslant 1-\varepsilon$ и, как легко видеть,

$$\Omega(\psi_0)<\frac{a_n+(b_n-a_n)\cos^2\psi_0+\dfrac{s}{\tilde\lambda}}{1-\dfrac{r}{\tilde\lambda}}\leqslant\frac{a_n+(b_n-a_n)(1-\varepsilon)+\dfrac{s}{\tilde\lambda}}{1-\dfrac{r}{\tilde\lambda}}=$$

$$=\frac{b_n-(b_n-a_n)\varepsilon+\dfrac{s}{\tilde\lambda}}{1-\dfrac{r}{\tilde\lambda}}\leqslant\frac{b_n-\beta\varepsilon+\dfrac{s}{\tilde\lambda}}{1-\dfrac{r}{\tilde\lambda}}.$$

При фиксированном ε и при достаточно малых r и s последняя дробь меньше правой части неравенства (53.12), и приходим к противоречию.

Итак, абсолютный максимум отношения (53.9) реализуется на векторе, для которого угол $\psi_n<\arcsin\varepsilon$. Но этот вектор, очевидно, есть собственный вектор, соответствующий собственному числу $\mu_1^{(n)}$ «неточной» задачи. Теорема доказана.

В классе минимальных систем условие сильной минимальности и необходимо (см. Л. Н. Довбыш [2]).

ГЛАВА VIII

ЭФФЕКТ ПОГРЕШНОСТИ В УРАВНЕНИИ

В настоящей главе мы рассматриваем следующий вопрос. Пусть в целях упрощения вычислений оператор данной задачи заменен другим, в некотором смысле близким к данному; как велика обусловленная этой заменой погрешность решения? С таким упрощением оператора мы встречаемся, например, в случаях, когда коэффициенты дифференциального уравнения или область интегрирования заменяются близкими, но более простыми по структуре; примеры такого рода будут рассмотрены ниже в настоящей главе.

Поставленный здесь вопрос близок к вопросу об устойчивости вычислительного процесса, но не тождествен с ним. Исследование устойчивости вычислительного процесса привело нас к необходимости сравнивать решения двух последовательностей уравнений

$$A_n x^{(n)} = y^{(n)} \qquad (n = 1, 2, \ldots)$$

и

$$(A_n + \Gamma_n) z^{(n)} = y^{(n)} + \delta^{(n)} \qquad (n = 1, 2, \ldots).$$

При этом на погрешность Γ_n оператора A_n накладывались довольно жесткие условия – во всяком случае, мы требовали, чтобы произведение $A_n^{-1}\Gamma_n$ было ограничено, — но зато оценки мы получали равномерные относительно n. В настоящей же главе мы будем рассматривать не две последовательности уравнений, а только два уравнения, которые можно было бы записать в виде

$$Ax = y, \qquad (A + \Gamma) z = y$$

(ниже мы используем иные обозначения). Задача тем самым упрощена, потому что исчезают параметр n и необходимость равномерных по этому параметру оценок, но оценки мы будем строить в новых предположениях относительно операторов A и Γ; в частности, мы не будем предполагать ограниченности произведения $A^{-1}\Gamma$.

Мы не будем рассматривать в этой главе погрешности, вызванной изменением свободного члена, так как этот вопрос решается достаточно просто.

Основные результаты настоящей главы содержатся в статьях автора [6], [7] и [10]; весьма сжатая формулировка этих результатов и один иллюстрирующий пример приведены в ВМ, § 63.

§ 54. Постановка задачи и оценка погрешности решения

Рассмотрим уравнение

$$Au = f, \tag{54.1}$$

в котором A — самосопряженный положительный оператор, действующий в гильбертовом пространстве H; разумеется, мы не исключаем случая, когда этот оператор на самом деле положительно определенный. Относительно правой части f примем, что символ (u, f) может быть определен на некотором множестве M, плотном в энергетическом пространстве H_A, и представляет собой функционал, линейный и ограниченный в этом пространстве. Тогда уравнение (54.1) имеет решение с конечной энергией; это решение обозначим через u_0.

Наряду с уравнением (54.1) рассмотрим уравнение

$$Bu = f, \tag{54.2}$$

где B — также самосопряженный положительный оператор, *полусходный с оператором* A. Энергетические пространства H_A и H_B состоят из одних и тех же элементов и метрики их эквивалентны; существуют, следовательно, такие положительные постоянные α и β, что

$$\alpha |u|_B^2 \leqslant |u|_A^2 \leqslant \beta |u|_B^2, \qquad u \in H_A. \tag{54.3}$$

Нетрудно видеть, что вместе с уравнением (54.1) имеет решение с конечной энергией и уравнение (54.2). Действительно, из неравенств (54.3) легко вытекает, что множество M, плотное в H_A, плотно и в H_B. Далее, функционал (u, f) ограничен в H_A:

$$|(u, f)| \leqslant c |u|_A, \qquad c = \text{const}.$$

Но тогда

$$|(u, f)| \leqslant c \sqrt{\beta}\, |u|_B,$$

функционал (u, f) ограничен в H_B и уравнение (54.2) имеет решение с конечной энергией. Это решение обозначим через u_1. Нашей ближайшей задачей является оценка величины $|u_0 - u_1|_B$.

Любой элемент пространства $u_1 \in H_B$ можно рассматривать как решение с конечной энергией для уравнения (54.2) при подходящем выборе функционала (u, f); достаточно положить $(u, f) = [u, u_1]_B$. В таком случае оператор, который решению (с конечной энергией) u_1 уравнения (54.2) приводит в соответствие такое же решение u_0 уравнения (54.1), определен на всем пространстве H_A Обозначим этот оператор через T, так что $u_0 = Tu_1$.

Имеем:

$$(u, f) = [u, u_0]_A, \qquad (u, f) = [u, u_1]_B, \quad u \in H_A.$$

Отсюда

$$[u, Tu_1]_A = [u, u_1]_B, \qquad u \in H_B.$$

Полагая здесь $u = u_1$, находим:

$$[Tu_1, u_1]_A = |u_1|_B^2 > 0, \qquad u_1 \neq 0; \tag{54.4}$$

оператор T, следовательно, самосопряженный в пространстве H_A. Из неравенства (54.3) получаем:

$$\frac{1}{\beta}|u_1|_A^2 \leqslant [Tu_1, u_1]_A \leqslant \frac{1}{\alpha}|u_1|_A^2.$$

Это неравенство означает, что спектр оператора T, рассматриваемого как оператор в пространстве H_A, заключен в сегменте $\left[\frac{1}{\beta}, \frac{1}{\alpha}\right]$. Но тогда спектр оператора $T-I$ заключен в сегменте $\left[\frac{1}{\beta}-1, \frac{1}{\alpha}-1\right]$ и

$$|T - I|_A \leqslant \max\left(\frac{|\alpha - 1|}{\alpha}, \frac{|\beta - 1|}{\beta}\right) = \eta. \tag{54.5}$$

Отсюда

$$|u_1 - u_0|_A = |(T - I)u_1|_A \leqslant \eta |u_1|_A. \tag{54.6}$$

Это и есть искомая оценка.

Укажем еще некоторые оценки, вытекающие из неравенства (54.6). Прежде всего, неравенства (54.3) позволяют написать такие формулы

$$|u_0 - u_1|_A \leqslant \eta \sqrt{\beta}\, |u_1|_B, \tag{54.7}$$

$$|u_0 - u_1|_B \leqslant \eta \sqrt{\frac{\beta}{\alpha}}\, |u_1|_B. \tag{54.8}$$

Одну полезную формулу можно получить из следующих соображений. Заменим оператор B через kB, где k — положительная постоянная, и постараемся подобрать k так, чтобы величина η стала наименьшей. При такой замене u_1 заменится через $\frac{1}{k}u_1$, а α и β — через $\frac{\alpha}{k}$ и $\frac{\beta}{k}$. Дело сводится к выбору такого $k > 0$, при котором

$$\max\left(\frac{|\alpha - k|}{\alpha}, \frac{|\beta - k|}{\beta}\right) = \min.$$

Рассуждая, как в § 12, найдем, что минимальное значение коэффициента η равно $\frac{\beta - \alpha}{\beta + \alpha}$, при этом $\frac{1}{k} = \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)$, и мы приходим к формуле, из иных соображений полученной М. Г. Слободянским [1]:

$$\left|\frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)u_1 - u_0\right|_A \leqslant \frac{\beta - \alpha}{2\alpha\beta}|u_1|_A. \tag{54.9}$$

Несколько иные оценки можно получить, если допустить, что в уравнениях (54.1) и (54.2) $f \in H$, а решения с конечной энергией u_0 и u_1 суть одновременно решения в обычном смысле слова, так что $u_0 \in D(A)$, $u_1 \in D(B)$. Одновременно $u_0 \in D\left(A^{\frac{1}{2}}\right)$ и $u_1 \in D\left(B^{\frac{1}{2}}\right)$, что легко вывести из соответствующих спектральных разложений. Построим самосопряженный оператор $Y = B^{-\frac{1}{2}} A^{\frac{1}{2}} A^{\frac{1}{2}} B^{-\frac{1}{2}}$ и найдем границы его спектра. Для этого в неравенстве (54.3) положим $u = B^{-\frac{1}{2}} v$, где v — любой элемент из области $D\left(B^{\frac{1}{2}}\right) = D\left(A^{\frac{1}{2}}\right)$. Напомним, что для таких элементов

$$|u|_B^2 = \left\| B^{\frac{1}{2}} u \right\|^2 = \| v \|^2.$$

Далее,

$$|u|_A^2 = \left\| A^{\frac{1}{2}} u \right\|^2 = \left\| A^{\frac{1}{2}} B^{-\frac{1}{2}} v \right\|^2 = \left(A^{\frac{1}{2}} B^{-\frac{1}{2}} v,\ A^{\frac{1}{2}} B^{-\frac{1}{2}} v\right) = (Yv,\ v),$$

и неравенство (54.3) принимает вид

$$\alpha \| v \|^2 \leqslant (Yv,\ v) \leqslant \beta \| v \|^2. \tag{54.10}$$

Неравенство (54.10) пока доказано для элементов $v \in D\left(B^{-\frac{1}{2}}\right)$. Но область определения самосопряженного оператора $D\left(B^{-\frac{1}{2}}\right)$ плотна в H, а оператор Y ограничен, потому что операторы A и B — полусходные. Теперь неравенство (54.10) можно распространить на все пространство H, и ясно, что спектр оператора Y лежит на сегменте $[\alpha,\ \beta]$.

Положим $v_0 = B^{\frac{1}{2}} u_0$, $v_1 = B^{\frac{1}{2}} u_1$. Умножив уравнение $Au_0 = f$ слева на $B^{-\frac{1}{2}}$ и заменив u_0 на $B^{-\frac{1}{2}} v_0$, получим $Yv_0 = B^{-\frac{1}{2}} f = v_1$. Отсюда $v_0 - v_1 = \left(Y^{-1} - I\right) v$ и, следовательно, $\| v_0 - v_1 \| \leqslant \leqslant \| Y^{-1} - 1 \| \| v_1 \|$. Спектр оператора $Y^{-1} - I$ заключен в сегменте $\left[\frac{1}{\beta} - 1,\ \frac{1}{\alpha} - 1\right]$, поэтому $\| Y^{-1} - I \| \leqslant \eta$, где η определяется формулой (54.5). Теперь

$$\| v_0 - v_1 \| \leqslant \eta \| v_1 \|.$$

Но

$$\| v_0 - v_1 \| = \left\| B^{\frac{1}{2}} (u_0 - u_1) \right\| = | u_0 - u_1 |_B,$$

$$\| v_1 \| = \left\| B^{\frac{1}{2}} u_1 \right\| = | u_1 |_B,$$

и окончательно:

$$| u_0 - u_1 |_B \leqslant \eta | u_1 |_B. \tag{54.11}$$

От оценки (54.6) эта оценка отличается тем, что в ней нормы берутся в метрике H_B. Справедлива и формула, аналогичная формуле (54.9):

$$\left|\frac{1}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)u_1-u_0\right|_B\leqslant\frac{\beta-\alpha}{2\alpha\beta}|u_1|_B. \tag{54.12}$$

§ 55. Применение к уравнениям второго порядка

1. Первая краевая задача. Пусть Ω — конечная область пространства координат $x_1, x_2, \ldots, x_m$, ограниченная кусочно гладкой поверхностью S. Поставим для этой области задачу об интегрировании невырождающегося эллиптического уравнения

$$-\sum_{j,k=1}^{m}\frac{\partial}{\partial x_j}\left(A_{jk}(x)\frac{\partial u}{\partial x_k}\right)+A_0(x)u=f(x),\qquad A_{jk}=\overline{A}_{kj} \tag{55.1}$$

при краевом условии

$$u|_S=0. \tag{55.2}$$

Будем считать, что коэффициенты $A_{jk}(x)$ и $A_0(x)$ ограничены и измеримы в области Ω. Для простоты допустим еще, что $A_0(x)\geqslant 0$ $f\in L_2(\Omega)$. Самосопряженное расширение по Фридрихсу оператора (55.1) — (55.2) обозначим через A, решение уравнения $Au=f(x)$ — через $u_0(x)$. Далее, через $u_1(x)$ обозначим решение уравнения $Bu=f$, где B есть расширение по Фридрихсу оператора, определяемого краевым условием (55.2) и невырождающимся эллиптическим дифференциальным выражением

$$-\sum_{j,k=1}^{m}\frac{\partial}{\partial x_j}\left(B_{jk}(x)\frac{\partial u}{\partial x_k}\right)+B_0(x)u,\qquad B_{jk}=\overline{B}_{kj}. \tag{55.3}$$

Коэффициенты B_{jk} и B_0 подчиним тем же условиям, что и коэффициенты A_{jk} и A_0. Оценим отношение

$$\frac{|u_0-u_1|_B}{|u_1|_B}; \tag{55.4}$$

как это вытекает из результатов предшествующего параграфа, достаточно будет определить постоянные α и β в неравенстве (54.3).

Имеем:

$$|u|_A^2=\int_\Omega\left\{\sum_{j,k=1}^{m}A_{jk}\frac{\partial\overline{u}}{\partial x_j}\frac{\partial u}{\partial x_k}+A_0|u|^2\right\}dx,$$

$$|u|_B^2=\int_\Omega\left\{\sum_{j,k=1}^{m}B_{jk}\frac{\partial\overline{u}}{\partial x_j}\frac{\partial u}{\partial x_k}+B_0|u|^2\right\}dx.$$

Составим уравнение

$$\mathrm{Det}\,|A_{jk}(x)-\varkappa B_{jk}(x)|=0.$$

Обозначим через $\varkappa_1(x)$ наименьший корень этого уравнения, а через $\varkappa_m(x)$ — его наибольший корень. Положим, далее,

$$\delta_1=\inf_{x\in\Omega}\varkappa_1(x),\qquad \delta_2=\sup_{x\in\Omega}\varkappa_m(x).$$

Тогда

$$(\delta_1-1)\int_\Omega\sum_{j,\,k=1}^m B_{jk}\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx\leqslant\int_\Omega\sum_{j,\,k=1}^m(A_{jk}-B_{jk})\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx\leqslant(\delta_2-1)\int_\Omega\sum_{j,\,k=1}^m B_{jk}\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx. \tag{55.5}$$

Далее, через λ_0 обозначим наименьшее собственное число оператора

$$-\sum_{j,\,k=1}^m\frac{\partial}{\partial x_j}\left(B_{jk}(x)\frac{\partial u}{\partial x_k}\right) \tag{55.6}$$

при краевом условии (55.2). Положим еще

$$\delta_0=\sup_{x\in\Omega}|A_0(x)-B_0(x)|.$$

Тогда, как легко видеть,

$$-\frac{\delta_0}{\lambda_0}\int_\Omega\sum_{j,\,k=1}^m B_{jk}\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx\leqslant\int_\Omega(A_0-B_0)|u|^2dx\leqslant\frac{\delta_0}{\lambda_0}\int_\Omega\sum_{j,\,k=1}^m B_{jk}\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx. \tag{55.7}$$

Неравенства (55.5) и (55.7) сложим и ко всем членам полученного нового неравенства прибавим величину $|u|_B^2$. Мы получим тогда:

$$\left(\delta_1-\frac{\delta_0}{\lambda_0}\right)\int_\Omega\sum_{j,\,k=1}^m B_{jk}\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx+\int_\Omega B_0|u|^2dx\leqslant|u|_A^2\leqslant\left(\delta_2+\frac{\delta_0}{\lambda_0}\right)\int_\Omega\sum_{j,\,k=1}^m B_{jk}\frac{\partial\bar u}{\partial x_j}\frac{\partial u}{\partial x_k}dx+\int_\Omega B_0|u|^2dx. \tag{55.8}$$

Теперь ясно, что неравенство (54.3) будет выполнено, если положим

$$\alpha=\min\left(\delta_1-\frac{\delta_0}{\lambda_0},\ 1\right),\qquad\beta=\max\left(\delta_2+\frac{\delta_0}{\lambda_0},\ 1\right); \tag{55.9}$$

мы предполагаем при этом, что $\delta_1 - \frac{\delta_0}{\lambda_0} > 0$. В частном случае, когда $A_0(x) = B_0(x) \equiv 0$, можно положить

$$\alpha = \delta_1, \qquad \beta = \delta_2. \tag{55.10}$$

Если $A_{jk}(x) \to B_{jk}(x)$ и $A_0(x) \to B_0(x)$ равномерно в Ω, то $\alpha \to 1$ и $\beta \to 1$. Из оценок § 51 и из факта положительной определенности оператора B вытекает тогда, что

$$\| u_0 - u_1 \|_{W_2^{(1)}(\Omega)} \to 0. \tag{55.11}$$

Результаты настоящего параграфа без изменений переносятся на случай, когда Ω — бесконечная область второго типа (§ 41), только под λ_0 следует понимать не наименьшее собственное число оператора (55.6), (55.2), а точную нижнюю границу его спектра. На случай бесконечной области первого типа эти результаты можно распространить, если, например, $A_0(x) = B_0(x)$; при этом, если $\inf A_0(x) = 0$, то формулу (55.11) следует заменить такой:

$$\| \operatorname{grad} u_0 - \operatorname{grad} u_1 \|_{L_2(\Omega)} \to 0. \tag{55.12}$$

2. Вторая краевая задача. Допустим теперь, что уравнение (55.1) решается при краевом условии

$$\left[\sum_{j,\, k=1}^{m} A_{jk}(x) \frac{\partial u}{\partial x_j} \cos(\nu,\ x_k) + \sigma_0(x) u \right]_S = 0, \tag{55.13}$$

а уравнение (55.3) — при условии

$$\left[\sum_{j,\, k=1}^{m} B_{jk}(x) \frac{\partial u}{\partial x_j} \cos(\nu,\ x_k) + \sigma_1(x) u \right]_S = 0. \tag{55.14}$$

Коэффициенты A_0, B_0, A_{jk}, B_{jk} подчиняем прежним ограничениям; по-прежнему, область Ω считаем конечной, ее границу S — кусочно гладкой. Заданные при $x \in S$ функции $\sigma_0(x)$ и $\sigma_1(x)$ будем считать измеримыми, ограниченными и, для простоты, строго положительными, так что

$$\inf \sigma_0 > 0, \qquad \inf \sigma_1 > 0. \tag{55.15}$$

В данном случае A будет означать у нас расширение по Фридрихсу оператора (55.1), (55.13), B — такое же расширение оператора (55.3), (55.14). Эти операторы — положительно определенные и полуходные, поэтому можно дать оценку величины (55.4), а для этого следует предварительно найти значения постоянных α и β, входящих в неравенство (54.3).

В случае условий (55.13) и (55.14) имеем:

$$|u|_A^2 = \int\limits_\Omega \left\{ \sum_{j,\,k=1}^m A_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k} + A_0 |u|^2 \right\} dx + \int\limits_S \sigma_0 |u|^2 dS,$$

$$|u|_B^2 = \int\limits_\Omega \left\{ \sum_{j,\,k=1}^m B_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k} + B_0 |u|^2 \right\} dx + \int\limits_S \sigma_1 |u|^2 dS.$$

Сохраним введенные выше обозначения δ_1, δ_2, δ_0, а через λ_0 обозначим наименьшее собственное число оператора (55.6) при краевом условии (55.14). Тогда сохраняют силу неравенства (55.5) и (55.7).

Положим

$$\bar{\delta} = \sup_{x \in S} \frac{|\sigma_0(x) - \sigma_1(x)|}{\sigma_1(x)}; \qquad (55.16)$$

как это следует из (55.4), величина $\bar{\delta}$ конечна, и $\bar{\delta} \to 0$, если $\sigma_0(x) \to \sigma_1(x)$ равномерно на S. Имеем теперь

$$-\bar{\delta} \int\limits_S \sigma_1 |u|^2 dS \leqslant \int\limits_S (\sigma_0 - \sigma_1) |u|^2 dS \leqslant \delta \int\limits_S \sigma_1 |u|^2 dS. \qquad (55.17)$$

Сложив неравенства (55.5), (55.8) и (55.17) и прибавив величину $|u|_B^2$ ко всем членам полученного нового неравенства, мы придем к соотношению

$$\left(\delta_1 - \frac{\delta_0}{\lambda_0}\right) \int\limits_\Omega \sum_{j,\,k=1}^m B_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k} dx + \int\limits_\Omega B_0 |u|^2 dx +$$

$$+ (1 - \bar{\delta}) \int\limits_S \sigma_1 |u|^2 dS \leqslant |u|_A^2 \leqslant \left(\delta_2 + \frac{\delta_0}{\lambda_0}\right) \int\limits_\Omega \sum_{j,\,k=1}^m B_{jk} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial u}{\partial x_k} dx +$$

$$+ \int\limits_\Omega B_0 |u|^2 dx + (1 + \bar{\delta}) \int\limits_S \sigma_1 |u|^2 dS.$$

Неравенство (54.3) будет выполнено, если положим

$$\alpha = \min\left(\delta_1 - \frac{\delta_0}{\lambda_0},\ 1 - \bar{\delta}\right), \qquad \beta = \max\left(\delta_2 + \frac{\delta_0}{\lambda_0},\ 1 + \bar{\delta}\right); \qquad (55.18)$$

мы предполагаем при этом, что последняя формула приводит к положительному значению α. Более простой, но отличающийся некоторыми особенностями случай $A_0 = B_0 = \sigma_0 \equiv 0$ изложен в ВМ, § 63.

§ 56. Применение к линейной теории оболочек. Постановка задачи

В этом и в следующих параграфах настоящей главы мы будем заниматься оценкой погрешности расчета упругой оболочки как плоской пластины. Эта задача возникла в связи с необходимостью исследовать распределение напряжений в лопасти водяной турбины. Это исследование наталкивается на значительные трудности, связанные как с довольно сложной формой самой лопасти, так и с выходящими за привычные стандарты условиями ее закрепления. Лопасть водяной турбины представляет собой оболочку переменной толщины, которую можно рассматривать как слабо изогнутую пластинку, иначе говоря, как пологую оболочку; часть ее края жестко закреплена, другая часть края свободна. Естественным образом возникла мысль [1]) приближенно рассчитать эту оболочку как плоскую пластинку; в связи с этим оказалось полезным оценить погрешность такого приближенного расчета. Общий прием для такой оценки указан в § 51; ниже мы применим его к только что сформулированной задаче.

Как известно, существует несколько вариантов уравнений теории тонких оболочек; эти варианты различаются в основном тем, какие члены считаются малыми и потому могут быть отброшены. В статье [10] автор предложил свой вариант уравнений теории оболочек; как обычно, они получаются из принципа минимума потенциальной энергии, но интеграл энергии предварительно преобразуется путем отбрасывания ряда слагаемых, которые автор считает малыми. Расхождение между уравнениями автора и уравнениями, приведенными, например, в книгах В. З. Власова [1] и В. В. Новожилова [1], не очень значительны.

Оценки, о которых сказано выше, проводились для предложенных автором уравнений теории оболочек. Результаты оказались следующими. В общем случае погрешность от замены пологой оболочки плоской пластинкой может оказаться сколь угодно большой, если только толщина оболочки достаточно мала. Благоприятное исключение представляет случай чисто моментного напряженного состояния. Экспериментальное исследование показало [2]), что напряженное состояние лопасти можно с большой степенью точности считать чисто моментным. Наши оценки показали, что относительная энергетическая погрешность расчета лопасти как плоской пластины для различных типов исследованных нами лопастей колеблется от 12,2% до 96,4%.

В общем случае, анализируя причину появления большой погрешности при замене тонкой оболочки плоской пластинкой, удалось сформулировать простые, более общие и, возможно, более точные, чем известные до того, уравнения равновесия пологих оболочек.

---

[1]) См. Л. М. Качанов [2].

[2]) См. С. П. Шихобалов, В. М. Краснов, Т. Д. Максутова, В. В. Цейтц, Е. И. Эдельштейн [1].

§ 57. Потенциальная энергия деформации оболочки

Обозначим через Ω срединную поверхность оболочки, о которой примем для простоты, что она имеет постоянную толщину, равную $2h$. Введем ортогональные криволинейные координаты α_1, α_2, α_3, где линии $\alpha_1=\mathrm{const}$ и $\alpha_2=\mathrm{const}$ суть линии кривизны срединной поверхности, а α_3 есть расстояние по нормали от данной точки до поверхности Ω, взятое с соответствующим знаком. Пусть в координатах α_1, α_2 длина элементарной дуги на срединной поверхности определяется формулой

$$ds^2=A_1^2(\alpha_1, \alpha_2)\,d\alpha_1^2+A_2^2(\alpha_1, \alpha_2)\,d\alpha_2^2.$$

Тогда, как известно, для элемента длины в пространстве имеем:

$$ds^2=(1-2k_1\alpha_3)\,A_1^2\,d\alpha_1^2+(1-2k_2\alpha_3)\,A_2^2\,d\alpha_2^2+d\alpha_3^2,$$

где k_1 и k_2 — главные кривизны срединной поверхности. Во всем последующем примем, что координатная сетка — правильная, так что коэффициенты A_1 и A_2 непрерывны, непрерывно дифференцируемы и нигде не обращаются в нуль в Ω.

Обозначим через u_1, u_2, u_3 составляющие вектора смещений $\boldsymbol{u}$ в осях α_1, α_2, α_3 и через $e_{ik}=e_{ik}(\boldsymbol{u})$ составляющие тензора деформаций; в местной системе координат x_1, x_2, x_3, образованной касательными к координатным линиям, эти составляющие имеют вид

$$e_{ik}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_k}+\frac{\partial u_k}{\partial x_i}\right). \tag{57.1}$$

Примем обычные гипотезы теории оболочек, по которым при вычислении потенциальной энергии деформации можно положить

$$e_{13}=e_{23}=0, \quad e_{33}=-\frac{\sigma}{1-\sigma}(e_{11}+e_{22}), \tag{57.2}$$

$$e_{ik}=\varepsilon_{ik}(\alpha_1, \alpha_2)+\alpha_3\varkappa_{ik}(\alpha_1, \alpha_2) \qquad (i,\ k=1,\ 2), \tag{57.3}$$

здесь σ — постоянная Пуассона.

Величины ε_{ik} будем далее называть деформациями срединной поверхности, а величины $\varkappa_{ik}$ — ее искривлениями. Эти величины можно рассматривать как составляющие двумерных тензоров E и K; в координатной системе α_1, α_2, где линии $\alpha_1=\mathrm{const}$ и $\alpha_2=\mathrm{const}$ суть линии кривизны поверхности Ω, деформации и искривления выражаются следующими формулами:

$$\left.\begin{aligned}
\varepsilon_{11}&=\frac{1}{A_1}\frac{\partial w_1}{\partial\alpha_1}+\frac{1}{A_1A_2}\frac{\partial A_1}{\partial\alpha_2}w_2-k_1w_3,\\
\varepsilon_{12}&=\frac{1}{2}\left\{\frac{A_1}{A_2}\frac{\partial}{\partial\alpha_2}\left(\frac{w_1}{A_1}\right)+\frac{A_2}{A_1}\frac{\partial}{\partial\alpha_1}\left(\frac{w_2}{A_2}\right)\right\},\\
\varepsilon_{22}&=\frac{1}{A_1A_2}\frac{\partial A_2}{\partial\alpha_1}w_1+\frac{1}{A_2}\frac{\partial w_2}{\partial\alpha_2}-k_2w_3,
\end{aligned}\right\} \tag{57.4}$$

$$\left.\begin{aligned}
\varkappa_{11} &= -\frac{1}{A_1}\frac{\partial k_1}{\partial \alpha_1} w_1 - \frac{1}{A_2}\frac{\partial k_1}{\partial \alpha_2} w_2 - k_1^2 w_3 - \\
&\qquad - \frac{1}{A_1}\frac{\partial}{\partial \alpha_1}\left(\frac{1}{A_1}\frac{\partial w_3}{\partial \alpha_1}\right) - \frac{1}{A_1 A_2^2}\frac{\partial A_1}{\partial \alpha_2}\frac{\partial w_3}{\partial \alpha_2}, \\
\varkappa_{12} &= \frac{1}{2}\left\{(k_2 - k_1)\left[\frac{A_1}{A_2}\frac{\partial}{\partial \alpha_2}\left(\frac{w_1}{A_1}\right) - \frac{A_2}{A_1}\frac{\partial}{\partial \alpha_2}\left(\frac{w_2}{A_2}\right)\right] - \right. \\
&\qquad \left. - \frac{A_1}{A_2}\frac{\partial}{\partial \alpha_2}\left(\frac{1}{A_1^2}\frac{\partial w_3}{\partial \alpha_1}\right) - \frac{A_2}{A_1}\frac{\partial}{\partial \alpha_1}\left(\frac{1}{A_2^2}\frac{\partial w_3}{\partial \alpha_2}\right)\right\}, \\
\varkappa_{22} &= -\frac{1}{A_1}\frac{\partial k_2}{\partial \alpha_1} w_1 - \frac{1}{A_2}\frac{\partial k_2}{\partial \alpha_2} w_2 - k_2^2 w_3 - \\
&\qquad - \frac{1}{A_1^2 A_2}\frac{\partial A_2}{\partial \alpha_1}\frac{\partial w_3}{\partial \alpha_1} - \frac{1}{A_2}\frac{\partial}{\partial \alpha_2}\left(\frac{1}{A_2}\frac{\partial w_3}{\partial \alpha_2}\right).
\end{aligned}\right\} \tag{57.5}$$

Здесь $w = (w_1, w_2, w_3)$ есть вектор смещения точек серединной поверхности.

Если пренебречь некоторыми малыми членами, то потенциальную энергию деформации оболочки можно привести к виду

$$Э(w) = \frac{E}{2(1-\sigma^2)} \int\limits_{\Omega}\int \left\{h\left[\varepsilon_{11}^2 + 2\sigma\varepsilon_{11}\varepsilon_{12} + \varepsilon_{22}^2 + 2(1-\sigma)\varepsilon_{12}^2\right] + \right.$$
$$\left. + \frac{h^3}{12}\left[\varkappa_{11}^2 + 2\sigma\varkappa_{11}\varkappa_{12} + \varkappa_{22}^2 + 2(1-\sigma)\varkappa_{12}^2\right]\right\} dS, \tag{57.6}$$

где E — модуль упругости материала оболочки и $dS = A_1 A_2\, d\alpha_1\, d\alpha_2$ — элемент площади поверхности Ω.

Внимательное рассмотрение подынтегрального выражения в (57.6) приводит к выводу [1]), что в некоторых случаях можно в выражениях искривлений (57.5) отбросить члены, зависящие от кривизн и их производных. Это можно сделать, в частности, в следующих двух случаях: 1) напряженное состояние оболочки не чисто моментное, так что тензор E отличен от тождественного нуля; 2) оболочка достаточно пологая. Последнее условие сформулируем точнее. Пусть параметры α_1 и α_2 — безразмерные, и координатная сетка — правильная, так что A_1 и A_2 нигде в Ω не обращаются в нуль. Наше условие состоит в том, что произведения

$$A_i k_j, \quad A_i \frac{\partial k_j}{\partial \alpha_m} \qquad (i,\ j,\ m = 1,\ 2), \tag{57.7}$$

должны быть достаточно малыми.

В последующем мы будем предполагать, что то или другое (или оба вместе) из наших условий выполнено. В таком случае вместо

---

[1]) Подробнее об этом см. статьи автора [7] и [10].

формул (57.5) можно написать следующие, значительно более простые:

$$\left.\begin{aligned}
\varkappa_{11} &= -\frac{1}{A_1}\frac{\partial}{\partial\alpha_1}\left(\frac{1}{A_1}\frac{\partial w_3}{\partial\alpha_1}\right) - \frac{1}{A_1A_2^2}\frac{\partial A_1}{\partial\alpha_2}\frac{\partial w_3}{\partial\alpha_2},\\
\varkappa_{12} &= -\frac{1}{2}\left\{\frac{A_1}{A_2}\frac{\partial}{\partial\alpha_2}\left(\frac{1}{A_1^2}\frac{\partial w_3}{\partial\alpha_1}\right) + \frac{A_2}{A_1}\frac{\partial}{\partial\alpha_1}\left(\frac{1}{A_2^2}\frac{\partial w_3}{\partial\alpha_2}\right)\right\},\\
\varkappa_{22} &= -\frac{1}{A_1^2A_2}\frac{\partial A_2}{\partial\alpha_1}\frac{\partial w_3}{\partial\alpha_1} - \frac{1}{A_2}\frac{\partial}{\partial\alpha_2}\left(\frac{1}{A_2}\frac{\partial w_3}{\partial\alpha_2}\right).
\end{aligned}\right\} \qquad (57.8)$$

Для дальнейшего важно, что формулы (57.8) остаются справедливыми в любой ортогональной системе координат α_1, α_2. Что касается формул (57.4), то при переходе к произвольным ортогональным координатам они заменяются следующими:

$$\left.\begin{aligned}
\varepsilon_{11} &= \frac{1}{A_1}\frac{\partial w_1}{\partial\alpha_1} + \frac{1}{A_1A_2}\frac{\partial A_1}{\partial\alpha_2}w_2 - \frac{L}{A_1^2}w_3,\\
\varepsilon_{12} &= \frac{1}{2}\left\{\frac{A_1}{A_2}\frac{\partial}{\partial\alpha_2}\left(\frac{w_1}{A_1}\right) + \frac{A_2}{A_1}\frac{\partial}{\partial\alpha_1}\left(\frac{w_2}{A_2}\right)\right\} - \frac{M}{A_1A_2}w_3,\\
\varepsilon_{22} &= \frac{1}{A_1A_2}\frac{\partial A_2}{\partial\alpha_1}w_1 + \frac{1}{A_2}\frac{\partial w_2}{\partial\alpha_2} - \frac{N}{A_2^2}w_3,
\end{aligned}\right\} \qquad (57.9)$$

где L, M, N — коэффициенты второй квадратичной формы поверхности Ω.

Заметим еще, что в любой ортогональной системе координат верна формула (57.6).

В дальнейшем мы будем рассматривать только такие оболочки, у которых хотя бы часть контура срединной поверхности жестко закреплена. Докажем, что если деформации и искривления этой поверхности тождественно равны нулю, то смещения ее точек также равны нулю.

Воспользуемся какой-либо изометрической системой координат, в которой $A_1 = A_2 = A$. Приравняв нулю тензор K, получим из уравнений (57.8):

$$\left.\begin{aligned}
\frac{\partial}{\partial\alpha_1}\left(\frac{1}{A}\frac{\partial w_3}{\partial\alpha_1}\right) + \frac{1}{A^2}\frac{\partial A}{\partial\alpha_2}\frac{\partial w_3}{\partial\alpha_2} = 0,\\
\frac{\partial}{\partial\alpha_2}\left(\frac{1}{A}\frac{\partial w_3}{\partial\alpha_2}\right) + \frac{1}{A^2}\frac{\partial A}{\partial\alpha_1}\frac{\partial w_3}{\partial\alpha_1} = 0,
\end{aligned}\right\} \qquad (57.10)$$

$$\frac{\partial}{\partial\alpha_2}\left(\frac{1}{A^2}\frac{\partial w_3}{\partial\alpha_1}\right) + \frac{\partial}{\partial\alpha_1}\left(\frac{1}{A^2}\frac{\partial w_3}{\partial\alpha_2}\right) = 0. \qquad (57.11)$$

Из уравнения (57.11) вытекает существование такой функции $\psi(\alpha_1, \alpha_2)$, что

$$\frac{\partial w_3}{\partial\alpha_1} = A^2\frac{\partial\psi}{\partial\alpha_1}, \qquad \frac{\partial w_3}{\partial\alpha_2} = -A^2\frac{\partial\psi}{\partial\alpha_2}. \qquad (57.12)$$

Подставив это в уравнения (57.10) и сложив полученные равенства, найдем:

$$\frac{\partial^2\psi}{\partial\alpha_1^2}+\frac{\partial^2\psi}{\partial\alpha_2^2}=0.$$

Таким образом, $\psi(\alpha_1, \alpha_2)$ есть гармоническая функция переменных α_1 и α_2 в области их изменения. В силу условий закрепления на некоторой части границы этой области

$$w_3=\frac{\partial w_3}{\partial\nu}=0$$

(ν — нормаль к границе поверхности Ω). Отсюда легко следует, что на указанной части границы

$$\frac{\partial w_1}{\partial\alpha_1}=\frac{\partial w_2}{\partial\alpha_2}=0$$

и, в силу уравнений (57.12),

$$\frac{\partial\psi}{\partial\alpha_1}=\frac{\partial\psi}{\partial\alpha_2}=0.$$

Но тогда во всей области $\psi=\text{const}$. Те же уравнения (57.12) дают теперь $w_3=\text{const}$, и так как на некоторой части границы $w_3=0$, то $w_3\equiv 0$.

По предположению, тензор $E\equiv 0$. Приняв во внимание, что $w_3\equiv 0$, из соотношений (57.9) получим:

$$\frac{\partial w_1}{\partial\alpha_1}+\frac{1}{A}\frac{\partial A}{\partial\alpha_2}w_2=0, \qquad \frac{1}{A}\frac{\partial A}{\partial\alpha_1}w_1+\frac{\partial w_2}{\partial\alpha_2}=0, \tag{57.13}$$

$$\frac{\partial}{\partial\alpha_2}\left(\frac{w_1}{A}\right)+\frac{\partial}{\partial\alpha_1}\left(\frac{w_2}{A}\right)=0. \tag{57.14}$$

Из равенства (57.14) вытекает существование такой функции $\varphi(\alpha_1, \alpha_2)$, что

$$w_1=A\frac{\partial\varphi}{\partial\alpha_1}, \qquad w_2=-A\frac{\partial\varphi}{\partial\alpha_2}.$$

Подставив это в (57.13) и сложив полученные равенства, получим:

$$\frac{\partial^2\varphi}{\partial\alpha_1^2}+\frac{\partial^2\varphi}{\partial\alpha_2^2}=0.$$

Таким образом, φ — гармоническая функция переменных α_1 и α_2 в области изменения этих переменных. На закрепленной части границы $w_1=w_2=0$ или $\frac{\partial\varphi}{\partial\alpha_1}=\frac{\partial\varphi}{\partial\alpha_2}=0$, поэтому $\varphi\equiv\text{const}$ и, следовательно, $w_1=w_2\equiv 0$.

§ 58. Оператор теории оболочек

Будем рассматривать оболочку, край которой жестко закреплен либо полностью, либо частично. В последнем случае допустим, что незакрепленная часть края — свободная от действия внешних сил.

Дифференциальные уравнения равновесия оболочки и естественные граничные условия на незакрепленной части ее края можно получить обычным путем из условия минимума потенциальной энергии оболочки

$$Э(w) - \iint_{\Omega} qw\,dS, \tag{58.1}$$

где $Э(w)$ определяется формулой (57.6), а q есть вектор внешней нагрузки, действующей на оболочку. Оператор, порождаемый упомянутыми дифференциальными уравнениями и краевыми условиями на всей границе срединной поверхности Ω, обозначим через P, так что совокупность уравнений равновесия и краевых условий можно записать в виде

$$Pw = q. \tag{58.2}$$

Мы не будем выписывать явно оператор P, нам это не понадобится.

Оператор P назовем оператором теории оболочек.

Важно отметить, что при $L = M = N = 0$ оператор P переходит в оператор, соответствующий задаче о равновесии плоской пластинки при той же внешней нагрузке и тех же краевых условиях.

Введем в рассмотрение вещественное гильбертово пространство H, элементы которого суть векторы, определенные в точках поверхности Ω; скалярное произведение и норму в H зададим формулами

$$(W', W'') = \iint_{\Omega} W'W''\,dS = \iint_{\Omega} \left(W_1'W_1'' + W_2'W_2'' + W_3'W_3''\right) dS,$$

$$\| W \|^2 = \iint_{\Omega} | W |^2\,dS = \iint_{\Omega} \left(W_1^2 + W_2^2 + W_3^2\right) dS.$$

Легко видеть, что оператор P симметричен и удовлетворяет тождеству

$$(Pw, w) = 2Э(w). \tag{58.3}$$

Из равенства (58.3) видно, что оператор P положителен. Действительно, $(Pw, w) \geqslant 0$. Если $(Pw, w) = 0$, то $Э(W) = 0$, а тогда $\varepsilon_{ik} \equiv 0$ и $\varkappa_{ik} \equiv 0$; по доказанному в § 54, в этом случае $w \equiv 0$. Можно доказать, что *оператор P положительно определенный*; здесь мы докажем это для того частного случая, когда оболочка однозначно проектируется на некоторую плоскость. Примем ее за плоскость (x_1, x_2). Пусть координаты x_1, x_2, x_3 определяются уравнениями

$$x_i = x_i(\alpha_1, \alpha_2, \alpha_3) \qquad (i = 1, 2, 3).$$

Тогда уравнения

$$x_i = x_i(\alpha_1, \alpha_2, 0) \qquad (i = 1, 2)$$

определяют проекцию Ω' серединной поверхности оболочки на плоскость x_1, x_2; можно тогда α_1 и α_2 рассматривать как криволинейные ортогональные координаты в Ω', при этом

$$ds^2 = dx_1^2 + dx_2^2 = A_1^2\, d\alpha_1^2 + A_2^2\, d\alpha_2^2.$$

Допустим, что оператор P — не положительно определенный. Тогда существует такая последовательность векторов $w_n \in D(P)$, что

$$Э(w_n) \xrightarrow[n\to\infty]{} 0, \qquad \|w_n\| = 1. \tag{58.4}$$

В таком случае, как показывает формула (54.6), в метрике $L_2(\Omega)$ стремятся к нулю все составляющие деформации $\varepsilon_{jk}^{(n)}$ и искривления $\varkappa_{jk}^{(n)}$, соответствующие векторам w_n.

Обозначим

$$w_n = \left(w_1^{(n)},\ w_2^{(n)},\ w_3^{(n)}\right).$$

В метрике $L_2(\Omega)$ выполняются предельные соотношения, вытекающие из формул (57.8):

$$\left.\begin{aligned}
&\frac{\partial^2 w_3^{(n)}}{\partial\alpha_1^2} - \frac{1}{A_1}\frac{\partial A_1}{\partial\alpha_1}\frac{\partial w_3^{(n)}}{\partial\alpha_1} + \frac{A_1}{A_2^2}\frac{\partial A_1}{\partial\alpha_2}\frac{\partial w_3^{(n)}}{\partial\alpha_2} \to 0,\\
&\frac{\partial^2 w_3^{(n)}}{\partial\alpha_1\,\partial\alpha_2} - \frac{1}{A_1}\frac{\partial A_1}{\partial\alpha_2}\frac{\partial w_3^{(n)}}{\partial\alpha_1} - \frac{1}{A_2}\frac{\partial A_2}{\partial\alpha_1}\frac{\partial w_3^{(n)}}{\partial\alpha_2} \to 0,\\
&\frac{\partial^2 w_3^{(n)}}{\partial\alpha_2^2} + \frac{A_2}{A_1^2}\frac{\partial A_2}{\partial\alpha_1}\frac{\partial w_3^{(n)}}{\partial\alpha_1} - \frac{1}{A_2}\frac{\partial A_2}{\partial\alpha_2}\frac{\partial w_3^{(n)}}{\partial\alpha_2} \to 0.
\end{aligned}\right\} \tag{58.5}$$

Докажем, что вторые производные от $w_3^{(n)}$ ограничены в метрике $L_2(\Omega')$. Допустим противное. Тогда существует такая подпоследовательность (обозначим ее, по-прежнему, через $w_3^{(n)}$), что

$$\sum_{j,\,k=1}^{2}\left\|\frac{\partial^2 w_3^{(n)}}{\partial\alpha_j\,\partial\alpha_k}\right\| \to \infty.$$

Левую часть последней формулы обозначим через Q_n. Для функции $\tilde w_3^{(n)} = Q_n^{-1} w_3^{(n)}$ соотношения (58.5) остаются верными. При этом нормы вторых производных от этих функций ограничены (они не превосходят единицы), а первые производные и сами функции обращаются в нуль на жестко закрепленной части контура. В силу теорем вложения множество первых производных $\dfrac{\partial\tilde w_3^{(n)}}{\partial\alpha_1}$, $\dfrac{\partial\tilde w_3^{(n)}}{\partial\alpha_2}$ и множество самих функций $\tilde{w}_3^{(n)}$ компактны в $L_2(\Omega)$; можно выделить подпоследовательность $\tilde{w}_3^{(n_k)}$, которая в метрике $W_2^{(1)}(\Omega)$ сходится к некоторой функции $\tilde{w}_3$, удовлетворяющей краевым условиям на жестко закрепленной части контура. Формулы (58.5) показывают, что вторые производные от $\tilde{w}_3^{(n_k)}$ также стремятся к некоторым пределам и, следовательно, $\tilde{w}_3 \in W_2^{(2)}(\Omega)$. Те же формулы (58.5) показывают, что искривления, соответствующие функции $\tilde{w}_3$, тождественно равны нулю. По доказанному в § 57, отсюда следует, что $\tilde{w}_3 \equiv 0$. Но это невозможно, потому что

$$\sum_{j,\, k=1}^{2} \left\| \frac{\partial^2 \tilde{w}_3}{\partial \alpha_j\, \partial \alpha_k} \right\| = \lim_{n \to \infty} \sum_{j,\, k=1}^{2} \left\| \frac{\partial^2 \tilde{w}_3^{(n)}}{\partial \alpha_j\, \partial \alpha_k} \right\| = 1.$$

Итак, нормы вторых производных функций $w_3^{(n)}$ ограничены. Повторив предшествующие рассуждения, мы докажем, что

$$\| w_3^{(n)} \|_{W_2^{(2)}(\Omega)} \to 0,$$

в частности,

$$\| w_3^{(n)} \|_{L_2(\Omega)} \to 0.$$

Теперь соотношения (58.4) дают следующее:

$$\| w_1^{(n)} \|^2 + \| w_2^{(n)} \|^2 \to 1, \tag{58.6}$$

и

$$\left.\begin{aligned} \frac{1}{A_1} \frac{\partial w_1^{(n)}}{\partial \alpha_1} + \frac{1}{A_1 A_2} \frac{\partial A_1}{\partial \alpha_2} w_2^{(n)} \to 0, \\ \frac{1}{2} \left\{ \frac{A_1}{A_2} \frac{\partial}{\partial \alpha_2} \left(\frac{w_1^{(n)}}{A_1} \right) + \frac{A_2}{A_1} \frac{\partial}{\partial \alpha_1} \left(\frac{w_2^{(n)}}{A_2} \right) \right\} \to 0, \\ \frac{1}{A_1 A_2} \frac{\partial A_2}{\partial \alpha_1} w_1 + \frac{1}{A_2} \frac{\partial w_2}{\partial \alpha_2} \to 0; \end{aligned}\right\} \tag{58.7}$$

стремление к пределу здесь в смысле метрики $L_2(\Omega)$ или, что то же, метрики $L_2(\Omega')$.

Будем рассматривать пару функций $\left(w_1^{(n)}, w_2^{(n)}\right)$ как двумерный вектор упругих смещений в плоской области Ω', тогда левые части формул (58.7) суть соответствующие этому вектору деформации. В силу условий закрепления упомянутый вектор обращается в нуль на части контура области Ω'; в таком случае[1]) оператор плоской задачи теории упругости положительно определен, что несовместимо с соотношениями (58.6) и (58.7). Положительная определенность оператора P доказана.

[1]) См. статьи Д. М. Эйдуса [1], [2], а также ПМ, § 43.

§ 59. Оболочки, близкие к плоским пластинам

Пусть даны оболочка с срединной поверхностью Ω и плоская пластина, вырезающая из своей срединной плоскости область Ω^0. Контур поверхности Ω обозначим через L, контур области Ω^0 — через L^0. Введем в Ω и Ω^0 одни и те же координаты α_1 и α_2; для простоты допустим, что эти координаты ортогональные как в Ω, так и в Ω^0. Далее, примем, что область изменения координат в обоих случаях одна и та же. Это позволяет любую функцию точки поверхности Ω (или контура L) рассматривать так же, как функцию точки области Ω^0 (или контура L^0). Допустим, что и к пластине, и к оболочке приложена одна и та же внешняя нагрузка $q(\alpha_1, \alpha_2)$ и что краевые условия, которые мы предполагаем однородными, тождественны в точках контуров L и L^0 с одинаковыми координатами. Пусть при этих условиях нагрузка q определяет векторы смещений: $\boldsymbol{w}$ в оболочке и $\boldsymbol{w}^0$ в пластине.

Оператор теории пластин будем здесь и ниже обозначать через P^0. Напомним, что формально оператор P^0 получается из P, если положить $L = M = N = 0$. Очевидно,

$$(P^0\boldsymbol{w}^0, \boldsymbol{w}^0) = 2\mathcal{E}^0(\boldsymbol{w}^0), \tag{59.1}$$

где $\mathcal{E}^0(\boldsymbol{w}^0)$ — потенциальная энергия деформации пластины, соответствующая смещению $\boldsymbol{w}^0$.

Условимся вообще обозначать поставленным сверху нуликом величины, относящиеся к плоской пластине. Так, если $\boldsymbol{w}$ — вектор смещений, то деформации пластины определяются формулами, чуть более простыми, чем для оболочки:

$$\left.\begin{aligned}
\varepsilon_{11}^0 &= \frac{1}{A_1^0}\frac{\partial w_1}{\partial\alpha_1} + \frac{1}{A_1^0A_2^0}\frac{\partial A_1^0}{\partial\alpha_2}w_2,\\
\varepsilon_{12}^0 &= \frac{1}{2}\left\{\frac{A_1^0}{A_2^0}\frac{\partial}{\partial\alpha_2}\left(\frac{w_1}{A_1^0}\right) + \frac{A_2^0}{A_1^0}\frac{\partial}{\partial\alpha_1}\left(\frac{w_2}{A_2^0}\right)\right\},\\
\varepsilon_{22}^0 &= \frac{1}{A_1^0A_2^0}\frac{\partial A_2^0}{\partial\alpha_1}w_1 + \frac{1}{A_2^0}\frac{\partial w_2}{\partial\alpha_2}.
\end{aligned}\right\} \tag{59.2}$$

Искривления и потенциальная энергия деформации пластины определяются формулами, вполне аналогичными формулам (57.8) и (57.6).

Общие результаты дают возможность оценить величину

$$\frac{\mathcal{E}^0(\boldsymbol{w} - \boldsymbol{w}^0)}{\mathcal{E}^0(\boldsymbol{w}^0)} \tag{59.3}$$

или, что то же, оценить в среднем погрешность при определении напряжений, если для их расчета оболочка Ω заменена пластиной Ω^0.

Оператор P теории оболочек определен на множестве векторов, имеющих необходимые производные и удовлетворяющих на контуре L краевым условиям. Аналогично оператор P^0 теории пластин имеет своей областью определения множество векторов, имеющих те же производные, что и в случае оператора P, но удовлетворяющих краевым условиям на контуре L. Обе области определения совпадают только в том случае, когда контур L (а следовательно, и контур L^0) весь жестко закреплен: в этом случае краевые условия как на контуре L, так и на контуре L^0 имеют вид (58.6). Если же закреплена только часть контура, скажем, L' (соответственно $L^{0'}$), то на L' и $L^{0'}$ краевые условия совпадают и имеют вид (58.6), а на дополнительных частях контуров краевые условия различны. В этом случае области определения операторов P и P^0 различны.

Нетрудно видеть, что энергетические пространства операторов P и P^0 оба состоят из векторов, имеющих обобщенные первые производные по α_1 и α_2, суммируемые с квадратом, и удовлетворяющих условиям (58.6) при значениях α_1 и α_2, соответствующих части контура L' (или, что то же, $L^{0'}$). Отсюда следует, что операторы P и P^0 полусходные.

Мы оценим величину (59.3), если найдем такие постоянные α и β, что

$$\alpha(P^0\boldsymbol{w},\ \boldsymbol{w}) \leqslant (P\boldsymbol{w},\ \boldsymbol{w}) \leqslant \beta(P^0\boldsymbol{w},\ \boldsymbol{w}),$$

или, что равносильно,

$$\alpha Э^0(\boldsymbol{w}) \leqslant Э(\boldsymbol{w}) \leqslant \beta Э^0(\boldsymbol{w}), \tag{59.4}$$

где $\boldsymbol{w}$ — произвольный вектор на H_P.

Будем считать, что оболочка Ω близка по форме к плоской пластине Ω^0, именно, примем, что величины

$$|A_i - A_i^0|, \quad \left|\frac{\partial A_i}{\partial \alpha_k} - \frac{\partial A_i^0}{\partial \alpha_k}\right|, \quad |L|, \quad |M|, \quad |N|$$

не превосходят некоторой достаточно малой величины γ.

Введем обозначения

$$U_\varepsilon = \varepsilon_{11}^2 + 2\sigma\varepsilon_{11}\varepsilon_{22} + \varepsilon_{22}^2 + 2(1-\sigma)\varepsilon_{12}^2,$$
$$U_\varkappa = \varkappa_{11}^2 + 2\sigma\varkappa_{11}\varkappa_{22} + \varkappa_{22}^2 + 2(1-\sigma)\varkappa_{12}^2;$$

через U_ε^0 и $U_\varkappa^0$ мы, в соответствии с принятым выше условием, обозначаем результат замены ε_{ik} на ε_{ik}^0 и $\varkappa_{ik}$ на $\varkappa_{ik}^0$ в выражениях U_ε и $U_\varkappa$. Положим еще

$$2Э_\varepsilon = \frac{Eh}{1-\sigma^2}\iint_\Omega U_\varepsilon\, dS, \qquad 2Э_\varkappa = \frac{Eh^3}{12(1-\sigma^2)}\iint_\Omega U_\varkappa\, dS,$$

$$2Э_\varepsilon^0 = \frac{Eh}{1-\sigma^2}\iint_{\Omega^0} U_\varepsilon^0\, dS^0, \qquad 2Э_\varkappa^0 = \frac{Eh^3}{12(1-\sigma^2)}\iint_{\Omega^0} U_\varkappa^0\, dS,$$

где dS^0 — элемент площади на плоскости (x_1, x_2). Очевидно,

$$Э = Э_\varepsilon + Э_\varkappa, \qquad Э^0 = Э^0_\varepsilon + Э^0_\varkappa.$$

Сравнивая выражения ε_{ik} и ε^0_{ik} (формулы (57.9) и (59.2)), соответствующие одному и тому же смещению $\boldsymbol{w}$, найдем:

$$\left.\begin{aligned}
\varepsilon_{11} &= \varepsilon^0_{11} + \left(\frac{A^0_1}{A_1} - 1\right)\varepsilon^0_{11} + \frac{1}{A_1}\left(\frac{1}{A_2}\frac{\partial A_1}{\partial \alpha_2} - \frac{1}{A^0_2}\frac{\partial A^0_1}{\partial \alpha_2}\right) w_2 - \frac{L}{A_1^2} w_3,\\
\varepsilon_{22} &= \varepsilon^0_{22} + \left(\frac{A^2_2}{A_2} - 1\right)\varepsilon^0_{22} + \frac{1}{A_2}\left(\frac{1}{A_1}\frac{\partial A_2}{\partial \alpha_1} - \frac{1}{A^0_1}\frac{\partial A^0_2}{\partial \alpha_1}\right) w_2 - \frac{N}{A_2^2} w_3,\\
\varepsilon_{12} &= \varepsilon^0_{12} + \frac{1}{2}\left\{\left(\frac{1}{A_2} - \frac{1}{A^0_2}\right)\frac{\partial w_1}{\partial \alpha_2} + \left(\frac{1}{A_1} - \frac{1}{A^0_1}\right)\frac{\partial w_2}{\partial \alpha_1}\right\} -\\
&\quad - \left(\frac{1}{A_1A_2}\frac{\partial A_1}{\partial \alpha_2} - \frac{1}{A^0_1A^0_2}\frac{\partial A^0_1}{\partial \alpha_2}\right) w_1 -\\
&\quad - \left(\frac{1}{A_1A_2}\frac{\partial A_2}{\partial \alpha_1} - \frac{1}{A^0_1A^0_1}\frac{\partial A^0_2}{\partial \alpha_1}\right) w_2 - \frac{M}{A_1A_2} w_3.
\end{aligned}\right\} \tag{59.5}$$

Подставив выражения (59.5) в U_ε, легко получим неравенство

$$\frac{Eh}{1-\sigma^2}\left|U_\varepsilon - U^0_\varepsilon\right| \leqslant C_1\gamma h\left\{\sum_{j,\,k=1}^{2}\left(\frac{\partial w_j}{\partial \alpha_k}\right)^2 + \sum_{j=1}^{3} w_j^2\right\},$$

где C_1 — некоторая постоянная [1]). Проинтегрировав последнее неравенство по Ω^0, получим:

$$\left|\frac{Eh}{1-\sigma^2}\iint_{\Omega} U_\varepsilon\, dS^0 - 2Э^0_\varepsilon\right| \leqslant C_1\gamma h\iint_{\Omega^0}\left\{\sum_{j,\,k=1}^{2}\left(\frac{\partial w_j}{\partial \alpha_k}\right)^2 + \sum_{j=1}^{3} w_j^2\right\} dS^0. \tag{59.6}$$

Для упругого тела, часть границы которого закреплена, верно так называемое *неравенство Корна* [2]), которое для случая плоской пластины принимает вид

$$h\iint_{\Omega^0}\sum_{j,\,k=1}^{2}\left(\frac{\partial w_j}{\partial \alpha_k}\right)^2 dS_0 \leqslant 2C_2 Э^0_\varepsilon. \tag{59.7}$$

---

[1]) Здесь и ниже в настоящей главе буквой C с тем или иным индексом будем обозначать положительную постоянную, зависящую только от вида области Ω^0, условий закрепления ее края и постоянных E и σ.

[2]) См. литературу, цитированную в сноске на стр. 284.

В силу теорем вложения любая функция u, равная нулю на дуге контура $L^{0'}$, удовлетворяет неравенству

$$\iint\limits_{\Omega^0} u^2\, dS^0 \leqslant C_3 \iint\limits_{\Omega^0} \left\{\left(\frac{\partial u}{\partial \alpha_1}\right)^2 + \left(\frac{\partial u}{\partial \alpha_2}\right)^2\right\} dS^0.$$

Последнее неравенство вместе с неравенством Корна дает

$$h \iint\limits_{\Omega^0} (w_1^2 + w_2^2)\, dS^0 \leqslant C_3 h \iint\limits_{\Omega^0} \sum_{j,\,k=1}^{2} \left(\frac{\partial w_j}{\partial \alpha_k}\right)^2 dS^0 \leqslant 2C_4 \mathcal{Э}_\varepsilon^0. \tag{59.8}$$

Неравенства (59.7) и (59.8) позволяют преобразовать формулу (59.6) к виду

$$\left|\frac{Eh}{1-\sigma^2} \iint\limits_{\Omega^0} U_\varepsilon\, dS^0 - 2\mathcal{Э}_\varepsilon^0\right| \leqslant 2C_5 \gamma h \mathcal{Э}_\varepsilon^0 + C_1 \gamma h \iint\limits_{\Omega^0} w_3^2\, dS^0. \tag{59.9}$$

Перейдем к оценке интеграла

$$\iint\limits_{\Omega^0} w_3^2\, dS^0.$$

Функция w_3 и ее первые производные обращаются в нуль на $L^{0'}$. Отсюда и из теорем вложения легко вытекают неравенства

$$\iint\limits_{\Omega^0} w_3^2\, dS^0 \leqslant C_6 \iint\limits_{\Omega^0} \left\{\left(\frac{\partial w_3}{\partial x_1}\right)^2 + \left(\frac{\partial w_3}{\partial x_2}\right)^2\right\} dS^0 \leqslant$$

$$\leqslant C_6^2 \iint\limits_{\Omega^0} \sum_{j,\,k=1}^{2} \left(\frac{\partial^2 w_3}{\partial x_j\, \partial x_k}\right)^2 dS^0,$$

здесь x_1, x_2 — декартовы координаты в области Ω^0. Но в декартовых координатах

$$\varkappa_{jk}^0 = -\frac{\partial^2 w_3}{\partial x_j\, \partial x_k},$$

и так как $U_\varkappa^0$ — положительно определенная квадратичная форма переменных $\varkappa_{jk}^0$, то

$$C_1 \iint\limits_{\Omega^0} w_3^2\, dS^0 \leqslant \frac{2}{h^3} C_7 \mathcal{Э}_\varkappa^0.$$

Подставив это в (59.9), получим окончательно:

$$\left|\frac{Eh}{1-\sigma^2} \iint\limits_{\Omega^0} U_\varepsilon\, dS^0 - 2\mathcal{Э}_\varepsilon^0\right| \leqslant 2C_5 \gamma \mathcal{Э}_\varepsilon^0 + \frac{2C_7\gamma}{h^2} \mathcal{Э}_\varkappa^0. \tag{59.10}$$

Из формул (54.8) нетрудно извлечь, что

$$\varkappa_{jk} = \varkappa^0_{jk} + a_{jk}\varkappa^0_{jk} + b_{jk}\frac{\partial w_3}{\partial \alpha_1} + c_{jk}\frac{\partial w_3}{\partial \alpha_2},$$

где a_{jk}, b_{jk}, c_{jk} суть величины порядка γ. Подставив это в $U_\varkappa$ и интегрируя, получим:

$$\left|\frac{Eh^3}{12(1-\sigma^2)}\iint\limits_{\Omega^0} U_\varkappa\, dS^0 - 2\mathcal{E}^0_\varkappa\right| \leqslant$$
$$\leqslant C_8\gamma\left\{2\mathcal{E}^0_\varkappa + h^3\iint\limits_{\Omega^0}\left[\left(\frac{\partial w_3}{\partial \alpha_1}\right)^2 + \left(\frac{\partial w_3}{\partial \alpha_2}\right)^2\right] dS^0\right\}.$$

Как и выше, последний интеграл оценивается через $\mathcal{E}^0_\varkappa$; отсюда

$$\left|\frac{Eh^3}{12(1-\sigma^2)}\iint\limits_{\Omega^0} U_\varkappa\, dS^0 - 2\mathcal{E}^0_\varkappa\right| \leqslant 2C_9\gamma\mathcal{E}^0_\varkappa. \tag{59.11}$$

Из неравенств (59.10) и (59.11) следует:

$$\left|\frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon + \frac{h^2}{12}U_\varkappa\right) dS^0 - 2\mathcal{E}^0\right| \leqslant 2\left(C_{10} + \frac{C_7}{h^2}\right)\gamma\mathcal{E}^0. \tag{59.12}$$

Рассмотрим теперь величину

$$2\mathcal{E} = \frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon + \frac{h^2}{12}U_\varkappa\right) dS. \tag{59.13}$$

Имеем:

$$dS = A_1A_2\, d\alpha_1\, d\alpha_2, \qquad dS^0 = A^0_1A^0_2\, d\alpha_1\, d\alpha_2.$$

Отсюда

$$dS = \frac{A_1A_2}{A^0_1A^0_2}\, dS^0 = (1+B)\, dS^0,$$

где, как нетрудно видеть, $0 \leqslant B = O(\gamma)$. Пусть $C_{11}\gamma \leqslant B \leqslant C_{12}\gamma$ (величина C_{11} может быть равной нулю). Тогда

$$2\mathcal{E} = \frac{Eh}{(1-\sigma^2)}\iint\limits_{\Omega^0}\left(U_\varepsilon + \frac{h^2}{12}U_\varkappa\right)(1+B)\, dS^0,$$

и следовательно,

$$C_{11}\gamma\frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon + \frac{h^2}{12}U_\varkappa\right) dS^0 \leqslant$$
$$\leqslant 2\mathcal{E} - \frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon + \frac{h^2}{12}U_\varkappa\right) dS^0 \leqslant$$
$$\leqslant C_{12}\gamma\frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon + \frac{h^2}{12}U_\varkappa\right) dS^0.$$

Интеграл в первом и третьем членах последнего неравенства заменяем его оценками, вытекающими из неравенства (59.12):

$$2Э^0(1-\gamma\Gamma)\leqslant\frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon+\frac{h^2}{12}U_\varkappa\right)dS^0\leqslant 2Э^0(1+\gamma\Gamma),$$

где для краткости положено

$$\Gamma=C_{10}+\frac{C_7}{h^2}. \tag{59.14}$$

Это приводит к неравенству

$$2Э^0C_{11}\gamma(1-\gamma\Gamma)\leqslant 2Э-\frac{Eh}{1-\sigma^2}\iint\limits_{\Omega^0}\left(U_\varepsilon+\frac{h^2}{12}U_\varkappa\right)dS^0\leqslant$$
$$\leqslant 2Э^0C_{12}\gamma(1+\gamma\Gamma).$$

Отсюда и из неравенства (59.12) получаем:

$$-\gamma[\Gamma-C_{11}(1-\gamma\Gamma)]Э^0\leqslant Э-Э^0\leqslant\gamma[\Gamma+C_{12}(1+\gamma\Gamma)],$$

и следовательно, в неравенстве (59.4) можно положить

$$\alpha=1-\gamma[\Gamma-C_{11}(1-\gamma\Gamma)],\qquad \beta=1+\gamma[\Gamma+C_{12}(1+\gamma\Gamma)]. \tag{59.15}$$

Если формула (59.15) дает значение $\alpha>0$, то для величины (59.3) верна оценка

$$\frac{Э^0(w-w^0)}{Э^0(w^0)}\leqslant\max\left(\frac{|\alpha-1|}{\alpha},\ \frac{|\beta-1|}{\beta}\right)=$$
$$=\max\left\{\frac{\gamma[\Gamma-C_{11}(1-\gamma\Gamma)]}{1-\gamma[\Gamma-C_{11}(1-\gamma\Gamma)]},\ \frac{\gamma[\Gamma+C_{12}(1+\gamma\Gamma)]}{1+\gamma[\Gamma+C_{12}(1+\gamma\Gamma)]}\right\}. \tag{59.16}$$

§ 60. Чисто моментное напряженное состояние

Пусть величина γ мала. Если приближенно рассчитывать напряжения в оболочке, заменив ее близкой плоской пластинкой Ω^0, то, как это следует из формул (59.14) и (59.16), погрешность такого расчета имеет порядок $\gamma\left(a+\frac{b}{h^2}\right)$, где a и b — некоторые положительные постоянные. Отсюда следует, что упомянутая выше погрешность может быть сколь угодно велика даже при сколь угодно малой γ, если только оболочка достаточно тонкая. Из сказанного вытекает, что в общем случае заменять для расчета напряжений оболочку близкой к ней плоской пластиной нельзя. Можно уничтожить член вида $\frac{b}{h^2}$ в оценке погрешности, если в выражениях деформаций при переходе от формул (57.9) к формулам (59.2) сохранить в этих последних слагаемые $-\frac{L}{A_1^2}w_3$, $-\frac{M}{A_1A_2}w_3$, $-\frac{N}{A_2^2}w_3$ Это приводит в конечном счете к уравнениям равновесия пологих оболочек. Мы не будем здесь на этом останавливаться[1]), а займемся случаем, когда член $\frac{b}{h^2}$ с самого начала не появляется в оценке. Это случай так называемого *чисто моментного* напряженного состояния, когда $\varepsilon_{jk} \equiv 0$. Рассмотрим этот случай подробнее. Имеем $\mathcal{Э} = \mathcal{Э}_\varkappa$, $\mathcal{Э}^0 = \mathcal{Э}^0_\varkappa$. Формула (59.11) принимает вид

$$\left| \frac{Eh^3}{12(1-\sigma^2)} \iint\limits_{\Omega^0} U_\varkappa \, dS^0 - 2\mathcal{Э}^0 \right| \leqslant 2C_9\gamma\mathcal{Э}^0. \tag{60.1}$$

Как и в § 59, полагаем $dS = (1+B)\,dS^0$ и $C_{11}\gamma \leqslant B \leqslant C_{12}\gamma$. Тогда

$$C_{11}\gamma \frac{Eh^3}{12(1-\sigma^2)} \iint\limits_{\Omega^0} U_\varkappa \, dS^0 \leqslant 2\mathcal{Э} - \frac{Eh^3}{12(1-\sigma^2)} \iint\limits_{\Omega^0} U_\varkappa \, dS^0 \leqslant$$
$$\leqslant C_{12}\gamma \frac{Eh^3}{12(1-\sigma^2)} \iint\limits_{\Omega^0} U_\varkappa \, dS^0. \tag{60.2}$$

Из неравенства (60.1) следует

$$2(1-C_8\gamma)\mathcal{Э}^0 \leqslant \frac{Eh^3}{12(1-\sigma^2)} \iint\limits_{\Omega^0} U_\varkappa \, dS^0 \leqslant 2(1+C_8\gamma)\mathcal{Э}^0. \tag{60.3}$$

Подставив это в крайние члены соотношения (60.2), получим:

$$2(1-C_8\gamma)C_{11}\gamma\mathcal{Э}^0 \leqslant 2\mathcal{Э} - \frac{Eh^3}{12(1-\sigma^2)} \iint\limits_{\Omega^0} U_\varkappa \, dS^0 \leqslant 2(1+C_8\gamma)C_{12}\gamma\mathcal{Э}^0.$$

Сложив это с неравенством (60.3), придем к искомой оценке:

$$(1-C_8\gamma)(1+C_{11}\gamma)\mathcal{Э}^0 \leqslant \mathcal{Э} \leqslant (1+C_8\gamma)(1+C_{12}\gamma)\mathcal{Э}^0. \tag{60.4}$$

Неравенства (60.4) дают возможность оценить величину (59.3), если $\gamma < \frac{1}{C_8}$. При этом

$$\frac{\mathcal{Э}^0(w-w^0)}{\mathcal{Э}^0(w^0)} \leqslant \max\left\{ \frac{|C_{11}-C_8-C_8C_{11}\gamma|\,\gamma}{(1-C_8\gamma)(1+C_{11}\gamma)},\ \frac{(C_8+C_{12}+C_8C_{12}\gamma)\,\gamma}{(1+C_8\gamma)(1+C_{11}\gamma)} \right\}. \tag{60.5}$$

Правую часть формулы (60.5) можно, вообще говоря, уменьшить, если воспользоваться формулой (54.12).

§ 61. Геликоидальная оболочка

В связи с вопросом о напряжениях в турбинной лопасти оказалось интересным применить формулу (57.6) к оболочке, серединная поверхность Ω которой есть часть прямого геликоида

$$x = \rho\cos\theta, \qquad y = \rho\sin\theta, \qquad z = \zeta\theta, \tag{61.1}$$

[1]) По этому поводу см. статьи автора [7] и [10].

определяемая неравенствами

$$\rho_0 \leqslant \rho \leqslant \rho_1, \qquad -\theta_0 \leqslant \theta \leqslant \theta_0, \tag{61.2}$$

где ρ_0, ρ_1, θ_0 — положительные постоянные. Пусть часть внутренней границы

$$\rho = \rho_0, \qquad \vartheta_1 < \theta < \vartheta_2 \tag{61.3}$$

жестко закреплена, а вся остальная граница оболочки свободна от действия внешних сил. Линии $\rho = \mathrm{const}$ и $\theta = \mathrm{const}$ на геликоиде ортогональны, и мы примем их за координатные, а ρ и θ — за криволинейные координаты; положим $\alpha_1 = \rho$, $\alpha_2 = \theta$. В координатах ρ и θ элемент длины линии, лежащей на геликоиде, определяется формулой

$$ds^2 = d\rho^2 + (\rho^2 + \zeta^2)\, d\theta^2,$$

так что

$$A_1 = 1, \qquad A_2 = \sqrt{\rho^2 + \zeta^2}.$$

Так как составляющие смещений w_1 и w_2 в последующие выкладки не войдут, то будем писать w вместо w_3. По формулам (57.8) имеем:

$$\left.\begin{aligned} \varkappa_{11} &= -\frac{\partial^2 w}{\partial \rho^2}, \quad \varkappa_{22} = -\frac{1}{\rho^2 + \zeta^2}\frac{\partial^2 w}{\partial \theta^2} - \frac{\rho}{\rho^2 + \zeta^2}\frac{\partial w}{\partial \rho}, \\ \varkappa_{12} &= -\frac{1}{2}\left\{\frac{1}{\sqrt{\rho^2 + \zeta^2}}\frac{\partial^2 w}{\partial \rho\, \partial \theta} + \sqrt{\rho^2 + \zeta^2}\frac{\partial}{\partial \rho}\left(\frac{1}{\rho^2 + \zeta^2}\frac{\partial w}{\partial \theta}\right)\right\}. \end{aligned}\right\} \tag{61.4}$$

Заменим оболочку плоской пластиной Ω^0 той же толщины. Будем считать, что Ω^0 есть часть кругового кольца, определяемая неравенствами (61.2). Имеем:

$$\left.\begin{aligned} &A_1^0 = 1, \qquad A_2^0 = \rho, \\ &\varkappa_{11}^0 = -\frac{\partial^2 w}{\partial \rho^2}, \quad \varkappa_{22}^0 = -\frac{1}{\rho^2}\frac{\partial^2 w}{\partial \theta^2} - \frac{1}{\rho}\frac{\partial w}{\partial \rho}, \\ &\varkappa_{12}^0 = -\frac{1}{\rho}\frac{\partial^2 w}{\partial \theta^2} + \frac{1}{2\rho^2}\frac{\partial w}{\partial \theta}. \end{aligned}\right\} \tag{61.5}$$

В соответствии с общими условиями § 59 примем, что внешняя нагрузка имеет одинаковые значения в точках оболочки и пластины с одинаковыми ρ и θ и что край пластины свободен, за исключением жестко закрепленной части (61.3). Оценим погрешность в решении, возникающую от упомянутой замены. Имеем:

$$2\text{Э} = D \int\limits_{\Omega}\int \left[\varkappa_{11}^2 + 2\sigma\varkappa_{11}\varkappa_{22} + \varkappa_{22}^2 + 2(1-\sigma)\varkappa_{12}^2\right] dS, \tag{61.6}$$

$$2\text{Э}^0 = D \int\limits_{\Omega^0}\int \left[\varkappa_1^{0^2} + 2\sigma\varkappa_{11}^0\varkappa_{22}^0 + \varkappa_{22}^{0^2} + 2(1-\sigma)\varkappa_{12}^{0^2}\right] dS^0, \tag{61.7}$$

где

$$D = \frac{Eh^3}{12(1-\sigma^2)}.$$

Так как в выражении (59.3) множитель D сократится, то будем ниже считать $D=1$.

Из формул (61.4) и (61.5) вытекают соотношения:

$$\left.\begin{aligned} &\varkappa_{11} = \varkappa_{11}^0, \quad \varkappa_{22} = (1-r)\varkappa_{22}^0, \\ &\varkappa_{12} = (1-s)\varkappa_{12}^0 - t\frac{1}{\rho\rho_0}\frac{\partial w}{\partial \theta}, \end{aligned}\right\} \tag{61.8}$$

где

$$\left.\begin{aligned} &r = \frac{\zeta^2}{\rho^2+\zeta^2}, \qquad s = \frac{\zeta^2}{\sqrt{\rho^2+\zeta^2}\,(\sqrt{\rho^2+\zeta^2}+\rho)}, \\ &t = \frac{\rho_0\zeta^4}{\rho(\rho^2+\zeta^2)(\sqrt{\rho^2+\zeta^2}+\rho)^2} = \frac{\rho_0}{\rho}s^2. \end{aligned}\right\} \tag{61.9}$$

Обозначим через U и U^0 подынтегральные функции в интегралах (61.6) и (61.7). По формулам (61.8) находим:

$$\begin{aligned} U = U^0 &- \left[(2r-r^2)\varkappa_{22}^{0^2} + 2\sigma r\varkappa_{11}^0\varkappa_{22}^0\right] - \\ &- 2(1-\sigma)\left[(2s-s^2)\varkappa_{12}^{0^2} + 2(1-s)t\varkappa_{22}^{0^2}\left(\frac{1}{\rho\rho_0}\frac{\partial w}{\partial\theta}\right) - \right. \\ &\left. - t^2\left(\frac{1}{\rho\rho_0}\frac{\partial w}{\partial\theta}\right)^2\right] = U^0 - F_1 - 2(1-\sigma)F_2. \end{aligned} \tag{61.10}$$

Элементарные оценки вместе с последующим интегрированием приводят к неравенству вида

$$\begin{aligned} -\delta'\mathcal{E}^0(w) &- \frac{\delta''}{\rho_0^2}\iint\limits_{\Omega^0}\left(\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right)^2 dS^0 \leqslant \\ &\leqslant \mathcal{E}(w) - \mathcal{E}^0(w) \leqslant \delta_1'\mathcal{E}^0(w) + \frac{\delta_1''}{\rho_0^2}\iint\limits_{\Omega^0}\left(\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right)^2 dS^0. \end{aligned} \tag{61.11}$$

Построение неравенства (61.11) будет показано в следующем параграфе на численном примере. Здесь же мы займемся следующим: чтобы перейти от неравенства (61.11) к неравенству (59.4), нам необходимо получить еще соотношение вида

$$\iint\limits_{\Omega^0}\left(\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right)^2 dS^0 \leqslant k\mathcal{E}^0(w) \tag{61.12}$$

и оценить входящую в это соотношение постоянную k. Формула

$$(\operatorname{grad} w)^2 = \left(\frac{\partial w}{\partial\rho}\right)^2 + \left(\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right)^2$$

позволяет ограничиться получением более сильного неравенства

$$\iint_{\Omega^0} (\operatorname{grad} w)^2\, dS^0 \leqslant k\mathcal{Э}^0(w); \tag{61.13}$$

оно должно быть установлено для функций, которые вместе со своими первыми производными обращаются в нуль на части границы

$$\rho = \rho_0, \qquad \vartheta_1 \leqslant \theta \leqslant \vartheta_2. \tag{61.14}$$

В неравенстве (61.13) перейдем к декартовым координатам

$$\iint_{\Omega^0} \left\{ \left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2 \right\} dS^0 \leqslant \frac{k}{2} \iint_{\Omega^0} \left\{ \left(\frac{\partial^2 w}{\partial x^2}\right)^2 + 2\sigma \frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2} + \left(\frac{\partial^2 w}{\partial y^2}\right)^2 + 2(1-\sigma)\left(\frac{\partial^2 w}{\partial x\, \partial y}\right)^2 \right\} dS^0. \tag{61.15}$$

Постоянная Пуассона σ меняется в пределах $-1 < \sigma < \frac{1}{2}$, поэтому $|\sigma| < 1$; для любых вещественных a и b справедливо неравенство

$$a^2 + 2\sigma ab + b^2 \geqslant (1 - |\sigma|)(a^2 + b^2),$$

и вместо неравенства (61.15) можно доказывать более сильное неравенство

$$\iint_{\Omega^0} \left\{ \left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2 \right\} dS^0 \leqslant \frac{k(1-|\sigma|)}{2} \iint_{\Omega^0} \left\{ \left(\frac{\partial^2 w}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 w}{\partial x\, \partial y}\right)^2 + \left(\frac{\partial^2 w}{\partial y^2}\right)^2 \right\} dS^0. \tag{61.16}$$

Пусть $u(x, y)$ — произвольная функция из $W_2^{(1)}(\Omega^0)$, равная нулю на линии (61.14), и пусть нам удалось найти такую постоянную λ_0, что

$$\iint_{\Omega^0} u^2\, dS^0 \leqslant \frac{1}{\lambda_0} \iint_{\Omega^0} \left\{ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right\} dS^0, \tag{61.17}$$

существование постоянной λ_0 вытекает из теорем вложения. Полагая последовательно $u = \frac{\partial w}{\partial x}$, $u = \frac{\partial w}{\partial y}$ и складывая, мы придем к неравенству (61.16), в котором

$$k = \frac{2}{\lambda_0(1-|\sigma|)}.$$

Вообще же можно положить

$$k \geqslant \frac{2}{\lambda_0 (1 - |\sigma|)}. \qquad (61.18)$$

Наша задача сведена к задаче отыскания числа λ_0. В конкретных случаях его можно построить процессом Ритца как наименьшее собственное число оператора Лапласа для области Ω^0 при следующих краевых условиях: $u = 0$ на линии (61.14) и $\frac{\partial u}{\partial \nu} = 0$ на остальной части границы. Но, с одной стороны, процесс Ритца требует довольно больших вычислений, с другой стороны, он дает значение λ_0 с избытком, что делает применение неравенства (61.17) ненадежным. Мы постараемся поэтому получить оценку для λ_0 снизу, минуя процесс Ритца.

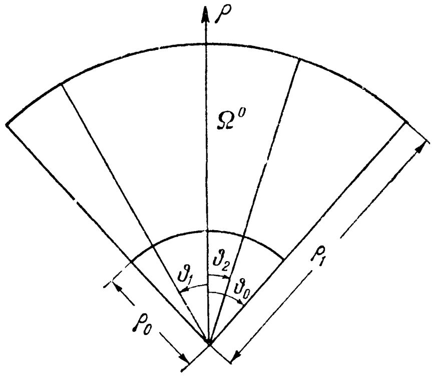

Рис. 5.

Рассмотрим область рис. 5 и пусть $u(\rho, \theta)$ — произвольная функция, непрерывно дифференцируемая в этой области и равная нулю на линии (61.14). Тогда, если $\vartheta_1 \leqslant \theta \leqslant \vartheta_2$, то

$$u(\rho, \theta) = \int_{\rho_0}^{\rho} u_\rho(\rho', \theta)\, d\rho' = \int_{\rho_0}^{\rho} u_\rho(\rho', \theta) \sqrt{\rho'} \frac{d\rho'}{\sqrt{\rho'}}, \qquad u_\rho = \frac{\partial u}{\partial \rho},$$

По неравенству Буняковского,

$$u^2(\rho, \theta) \leqslant \ln\frac{\rho}{\rho_0} \int_{\rho_0}^{\rho} u_\rho^2(\rho', \theta)\, \rho'\, d\rho' \leqslant \ln\frac{\rho}{\rho_0} \int_{\rho_0}^{\rho_1} u_\rho^2(\rho', \theta)\, \rho\, d\rho'.$$

Умножим последнее неравенство на $dS^0 = \rho\, d\rho\, d\theta$ и проинтегрируем по области $\rho_0 \leqslant \rho \leqslant \rho_1$, $\vartheta_1 \leqslant \theta \leqslant \vartheta_2$. После простых подсчетов получим:

$$\int_{\vartheta_1}^{\vartheta_2} \int_{\rho_0}^{\rho_1} u^2(\rho, \theta)\, dS^0 \leqslant \rho_0^2 K \int_{\vartheta_1}^{\vartheta_2} \int_{\rho_0}^{\rho_1} u_\rho^2(\rho, \theta)\, dS^0, \qquad (61.19)$$

где

$$K = \frac{1}{4} \left\{ 2 \left(\frac{\rho_1}{\rho_0}\right)^2 \ln\frac{\rho_1}{\rho_0} - \left(\frac{\rho_1}{\rho_0}\right)^2 + 1 \right\}. \qquad (61.20)$$

Если весь внутренний край жестко закреплен, так что $\vartheta_1 = -\theta_0$, $\vartheta_2 = \theta_0$, то, как легко видеть, $\frac{1}{\lambda_0} \leqslant K$ и можно положить

$$k = \frac{2K}{1 - |\sigma|}.$$

Займемся общим случаем. В формуле (61.19) распространим интегрирование справа на всю область Ω^0. Обозначая для краткости

$$A=\int\int_{\Omega^0} u_\rho^2\,dS^0,$$

будем иметь:

$$\int_{\vartheta_1}^{\vartheta_2}\int_{\rho_0}^{\rho_1} u^2\,dS^0 \leqslant \rho_0^2 KA. \tag{61.21}$$

Из этого неравенства вытекает существование такого угла φ, $\vartheta_1<\varphi<\vartheta_2$, что

$$\int_{\rho_0}^{\rho_1} u^2(\rho,\ \varphi)\,\rho\,d\rho \leqslant \frac{\rho_0^2 K}{\vartheta_2-\vartheta_1}A. \tag{61.22}$$

Пусть $\theta>\varphi$. Имеем, очевидно,

$$u^2(\rho,\ \theta)=u^2(\rho,\ \varphi)+\int_{\varphi}^{\theta}\frac{\partial\,[u^2(\rho,\theta')]}{\partial\theta'}\,d\theta'.$$

Отсюда

$$\int_{\rho_0}^{\rho_1} u^2\rho\,d\rho=\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta}\frac{\partial\,[u^2(\rho,\theta')]}{\partial\theta'}\,dS^0+\int_{\rho_0}^{\rho_1}u^2(\rho,\ \varphi)\,\rho\,d\rho. \tag{61.23}$$

Второй член справа оценим по неравенству (61.22). Далее,

$$\left|\frac{\partial u^2(\rho,\theta')}{\partial\theta'}\right|\leqslant \varepsilon\rho^2u^2(\rho,\ \theta')+\frac{1}{\varepsilon}\left[\frac{1}{\rho}\,u_\theta(\rho,\ \theta')\right]^2,$$

где ε — произвольное положительное число. Теперь

$$\int_{\rho_0}^{\rho_1} u^2\rho\,d\rho\leqslant\varepsilon\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta}\rho^2u^2\,dS^0+\frac{1}{\varepsilon}\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta}\left(\frac{1}{\rho}\,u_\theta(\rho,\ \theta')\right)^2dS^0+\frac{\rho_0^2K}{\vartheta_2-\vartheta_1}A\leqslant$$
$$\leqslant\varepsilon\rho_1^2\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta_0}u^2\,dS^0+\frac{1}{\varepsilon}\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta_0}\left(\frac{1}{\rho}\,u_\theta\right)^2dS^0+\frac{\rho_0^2K}{\vartheta_2-\vartheta_1}A. \tag{61.24}$$

Положим

$$\varepsilon=\frac{1}{2\rho_1^2(\theta_0-\varphi)}.$$

Проинтегрировав последнее неравенство по θ и перенеся налево первый член справа, придем к неравенству

$$\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta_0}u^2\,dS^0\leqslant 4\rho_1^2(\theta_0-\varphi)^2\int_{\rho_0}^{\rho_1}\int_{\varphi}^{\theta_0}\left(\frac{1}{\rho}\,u_\theta\right)^2dS^0+\frac{2\rho_0^2K(\theta_0-\varphi)}{\vartheta_2-\vartheta_1}A. \tag{61.25}$$

Аналогично найдем

$$\int\limits_{\rho_0}^{\rho_1}\int\limits_{-\theta_0}^{\varphi} u^2\,dS^0 \leqslant 4\rho_1^2(\theta_0-\varphi)^2\int\limits_{\rho_0}^{\rho_1}\int\limits_{-\theta_0}^{\varphi}\left(\frac{1}{\rho}u_\theta\right)^2 dS^0+$$

$$+\frac{2\rho_0^2K(\theta_0+\varphi)}{\vartheta_2-\vartheta_1}A. \qquad (61.26)$$

Неравенства (61.25) и (61.26) сложим. Заметим при этом, что

$$\theta_0-\varphi\leqslant\theta_0-\vartheta_1, \qquad \theta_0+\varphi\leqslant\theta_0+\vartheta_2.$$

Обозначая

$$\vartheta=\max(\theta_0-\vartheta_1,\ \theta_0+\vartheta_2) \qquad (61.27)$$

и вспоминая значение A, будем иметь:

$$\iint\limits_{\Omega^0} u^2\,dS^0\leqslant 4\rho_1^2\vartheta^2\iint\limits_{\Omega^0}\left(\frac{1}{\rho}u_\theta\right)^2 dS^0+\frac{4\rho_0^2K\theta_0}{\vartheta_2-\vartheta_1}\iint\limits_{\Omega^0}u_\rho^2\,dS^0. \qquad (61.28)$$

Наконец, полагая

$$m=\max\left(4\rho_1^2\vartheta^2,\ \frac{4\rho_0^2K\theta_0}{\vartheta_2-\vartheta_1}\right), \qquad (61.29)$$

получаем окончательную оценку:

$$\iint\limits_{\Omega^0} u^2\,dS\leqslant m\iint\limits_{\Omega^0}(\operatorname{grad}u)^2\,dS^0. \qquad (61.30)$$

Сравнив соотношения (61.17) и (61.30), найдем, что $\frac{1}{\lambda_0}\leqslant m$. Неравенство (61.18) показывает, что можно положить

$$k=\frac{2m}{1-|\sigma|}. \qquad (61.31)$$

§ 62. Численный пример

В настоящем параграфе мы проведем детальные оценки для геликоидальной оболочки, проектирующейся на область рис. 5, при заданных численных значениях определяющих ее параметров:

$$2\theta_0=85^\circ,\quad \vartheta_1=-29^\circ 30',\quad \vartheta_2=18^\circ 30',\quad \frac{\rho_1}{\rho_0}=2{,}50,\quad \frac{\zeta}{\rho_0}=0{,}50;$$

по формуле (61.20) находим еще

$$K=1{,}55.$$

Примем постоянную Пуассона $\sigma=0{,}29$.

Основное внимание будет уделено подсчету констант, входящих в формулу (61.11).

Оценим входящую в равенство (61.10) форму F_1, сравнив ее с формой

$$A = \varkappa_{11}^{0^2} + 2\sigma\varkappa_{11}^0\varkappa_{22}^0 + \varkappa_{22}^{0^2}. \tag{62.1}$$

С этой целью составим определитель формы $F_1 - \lambda A$, равный

$$\begin{vmatrix} -\lambda & 0{,}29\,(r-\lambda) \\ 0{,}29\,(r-\lambda) & 2r - r^2 - \lambda \end{vmatrix}.$$

Его корни

$$\lambda' = r - 0{,}546r^2 - \sqrt{(r - 0{,}546r^2)^2 + 0{,}092r^2},$$

$$\lambda'' = r - 0{,}546r^2 + \sqrt{(r - 0{,}546r^2)^2 + 0{,}092r^2}.$$

При достаточно малых r корень λ' убывает, а корень λ'' возрастает при возрастании r. Поэтому λ' примет наименьшее, а λ'' — наибольшее значение при

$$r = r_{\max} = \frac{\zeta^2}{\zeta^2 + \rho_0^2} = 0{,}2.$$

Обозначая эти значения через λ_0' и λ_0'', имеем $\lambda_0' = -0{,}010$, $\lambda_0'' = 0{,}366$, и следовательно,

$$-0{,}010A \leqslant F_1 \leqslant 0{,}366A. \tag{62.2}$$

Перейдем к величине F_2. Для краткости обозначим $\varkappa_{12}^0 = u$, $\frac{1}{\rho_0\rho}\frac{\partial w}{\partial\theta} = v$, тогда

$$F_2 = (2s - s^2)\,u^2 + 2\,(1-s)\,tuv - t^2v^2. \tag{62.3}$$

Величина s, очевидно, меньше единицы, поэтому величина $2s - s^2$ принимает наибольшие значения при наибольшем s, которое соответствует наименьшему значению $\rho = \rho_0$ и равно

$$\frac{1}{5 + 2\sqrt{5}} = 0{,}106.$$

Отсюда

$$(2s - s^2)_{\max} = 0{,}200.$$

Наибольшее значение $t = \frac{\rho_0}{\rho}s^2$ равно $s_{\max}^2 = 0{,}011$ и $[2\,(1-s)\,t]_{\max} = 0{,}019$. Наконец, $t_{\max}^2 = 0{,}0001$; этой величиной мы пренебрежем ввиду ее малости. Теперь

$$|F_2| \leqslant 0{,}200u^2 + 0{,}019\,|uv| \leqslant (0{,}200 + 0{,}010\varepsilon)\,u^2 + \frac{0{,}010}{\varepsilon}v^2,$$

где ε — произвольное положительное число. Выберем его так, чтобы $0{,}200+0{,}010\varepsilon=0{,}366$. Тогда $\varepsilon=16{,}6$ и $\frac{0{,}010}{\varepsilon}=0{,}0006$. Таким образом,

$$0{,}366\varkappa_{12}^{02}-0{,}0006\left(\frac{1}{\rho_0\rho}\frac{\partial w}{\partial\theta}\right)^2\leqslant F_2\leqslant$$
$$\leqslant 0{,}366\varkappa_{12}^{02}+0{,}0006\left(\frac{1}{\rho_0\rho}\frac{\partial w}{\partial\theta}\right)^2. \tag{62.4}$$

Из формул (62.2), (62.4) и (61.10) получаем:

$$-0{,}366U^0-0{,}0012(1-\sigma)\left(\frac{1}{\rho_0\rho}\frac{\partial w}{\partial\theta}\right)^2\leqslant U-U^0\leqslant$$
$$\leqslant 0{,}010\left(\varkappa_{11}^{02}+2\sigma\varkappa_{11}^0\varkappa_{22}^0+\varkappa_{22}^{02}\right)+0{,}366\cdot 2(1-\sigma)\varkappa_{12}^{02}+$$
$$+0{,}0012(1-\sigma)\left(\frac{1}{\rho_0\rho}\frac{\partial w}{\partial\theta}\right)^2.$$

Заменив справа множитель 0,010 бо́льшим числом 0,366, получим более простое неравенство

$$|U-U^0|\leqslant\left|0{,}366U^0+0{,}0012(1-\sigma)\left(\frac{1}{\rho_0\rho}\frac{\partial w}{\partial\theta}\right)^2\right|.$$

Проинтегрировав по Ω^0, получим:

$$\left|\frac{1}{2}\iint_{\Omega^0}U\,dS^0-Э^0\right|\leqslant 0{,}366Э^0+\frac{0{,}0006(1-\sigma)}{\rho_0^2}\iint_{\Omega^0}\left(\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right)^2dS^0. \tag{62.5}$$

Воспользуемся теперь формулой (61.12); входящий в нее коэффициент найдем по формуле (61.31). Равенство (61.29) дает

$$m=\rho_0^2\max\left(4\left(\frac{\rho_1}{\rho_0}\right)^2\vartheta^2,\ \frac{4K\theta_0}{\vartheta_2-\vartheta_1}\right)=1{,}6\pi^2\rho_0^2=15{,}79\rho_0^2.$$

Отсюда

$$k=\frac{31{,}58}{1-\sigma}\rho_0^2$$

и

$$\frac{1-\sigma}{\rho_0^2}\iint_{\Omega^0}\left(\frac{1}{\rho}\frac{\partial w}{\partial\theta}\right)^2dS^0\leqslant 31{,}58Э^0.$$

Подставив это в неравенство (62.5), получим:

$$\left|\frac{1}{2}\iint_{\Omega^0}U\,dS^0-Э^0\right|\leqslant 0{,}388Э^0. \tag{62.6}$$

Оценим теперь разность

$$2Э-\iint_{\Omega}U\,dS^0=\iint_{\Omega}U\,dS-\iint_{\Omega^0}U\,dS^0=$$
$$=\iint_{\Omega^0}U\left[\frac{1}{\rho}\sqrt{\rho^2+\zeta^2}-1\right]dS^0.$$

Величина в квадратных скобках оценивается так:

$$\frac{1}{\rho}\sqrt{\rho^2+\zeta^2}-1=\frac{\zeta^2}{\rho\left(\sqrt{\rho^2+\zeta^2}+\rho\right)}.$$

Заменяя ρ его наибольшим значением $\rho=\rho_1=2{,}5\rho_0$, а затем наименьшим значением $\rho=\rho_0$, найдем:

$$0{,}020\leqslant\frac{1}{\rho}\sqrt{\rho^2+\zeta^2}-1\leqslant 0{,}118.$$

Отсюда

$$0{,}010\iint\limits_{\Omega^0}U\,dS^0\leqslant Э-\frac{1}{2}\iint\limits_{\Omega^0}U\,dS^0\leqslant 0{,}059\iint\limits_{\Omega^0}U\,dS^0. \quad (62.7)$$

Из неравенства (62.6) следует

$$1{,}244Э^0\leqslant\iint\limits_{\Omega^0}U\,dS^0=2{,}776Э^0. \quad (62.8)$$

Подставим это в крайние члены неравенства (62.7):

$$0{,}012Э^0\leqslant Э-\frac{1}{2}\iint\limits_{\Omega^0}U\,dS^0\leqslant 0{,}164Э^0. \quad (62.9)$$

Разделив неравенство (62.8) на 2 и сложив с (62.9), получим окончательно:

$$0{,}624Э^0<Э<1{,}552Э^0. \quad (62.10)$$

Итак, входящие в формулу (54.3) постоянные суть в нашем случае

$$\alpha=0{,}624, \qquad \beta=1{,}552.$$

Отсюда следует искомая оценка:

$$\frac{Э^0(w-w^0)}{Э^0(w^0)}\leqslant\max\left(\frac{1-0{,}624}{0{,}624};\ \frac{1{,}552-1}{1{,}552}\right)=0{,}60.$$

Применение формулы (54.12) несколько улучшает оценку: если за приближенное решение принять не w^0, а

$$\tilde{w}^0=\frac{1}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)w^0=1{,}123w^0,$$

то

$$\frac{Э^0(w-\tilde{w}^0)}{Э^0(\tilde{w}^0)}\leqslant 0{,}48.$$

ГЛАВА IX

ВАРИАЦИОННЫЕ МЕТОДЫ В НЕЛИНЕЙНЫХ ЗАДАЧАХ

§ 63. Предварительные замечания и вспомогательные сведения

В этой и в следующей главе мы будем рассматривать нелинейные операторы, действующие из одного банахова (в частности, гильбертова) пространства в другое. Иногда — эти случаи будут особо оговариваться — мы будем рассматривать и операторы, действующие в метрическом ненормированном пространстве. В этом последнем случае мы будем предполагать, что элементы рассматриваемого пространства образуют линейное множество.

Во всех случаях будем рассматривать только вещественные пространства.

Если u — элемент банахова пространства B, а f — элемент сопряженного пространства B^*, то символ $(f, u) = (u, f)$ будет обозначать значение функционала f на элементе u.

Относительно изучаемых в последующем нелинейных операторов сделаем следующие допущения, которые ниже всегда предполагаются выполненными.

I. Область определения оператора линейна и плотна в рассматриваемом пространстве.

II. Пусть $x_0, x_1, x_2, \ldots, x_n$ — линейно независимые элементы данного пространства. Гиперплоскостью n измерений в нем назовем совокупность элементов вида

$$x_0 + \sum_{k=1}^{n} a_k x_k,$$

где a_k — произвольные (вещественные) постоянные.

Потребуем, чтобы при любом n и при любых элементах $x_0, x_1, \ldots, x_n$, принадлежащих области определения данного оператора, значения этого оператора были сильно дифференцируемыми функциями числовых переменных[1]) $a_1, a_2, \ldots, a_n$ при всех конечных

---

[1]) Подробно об абстрактных функциях числовых переменных см., например, М. М. Вайнберг [1] или Л. В. Канторович и Г. П. Акилов [1].

значениях этих переменных. Коротко будем это формулировать так: данный оператор сильно непрерывно дифференцируем на любой гиперплоскости из области своего определения.

Сформулируем некоторые нужные для дальнейшего понятия и теоремы теории нелинейных операторов в банаховых пространствах; достаточно полное изложение этой теории имеется, например, в книгах Л. В. Канторовича и Г. П. Акилова [1] и М. М. Вайнберга [1].

Пусть в банаховом пространстве B действует нелинейный оператор $P(u)$. Пусть u, $h \in B$. В силу требования II существует предел

$$\lim_{t \to 0} \frac{P(u+th)-P(u)}{t} = \frac{d}{dt} P(u+th)\big|_{t=0}.$$

Этот предел обозначим через $P'_u h$:

$$P'_u h = \frac{d}{dt} P(u+th)\big|_{t=0}; \tag{63.1}$$

он называется ***дифференциалом Гато оператора*** P. Дифференциал Гато можно рассматривать как значение в точке h некоторого оператора P'_u, зависящего от точки u. В общем случае этот оператор однороден: $P'_u(ah) \equiv aP'_u h$, но неаддитивен: $P'_u(a_1 h_1 + a_2 h_2) \not\equiv$ $\not\equiv a_1 P'_u h_1 + a_2 P'_u h_2$.

Пусть в некоторой точке u оператор P'_u также и аддитивен, когда h пробегает некоторое множество $M_u \subset D(P)$. Тогда P'_u есть линейный (но необязательно ограниченный) оператор с областью определения $D(P'_u) = M_u$. В этом случае P'_u называют *производной Гато* или просто *производной оператора P в точке u*.

На рассматриваемые нами операторы наложим еще одно требование.

III. Производная P'_u существует при любом $u \in D(P)$.

Оператор и его производная связаны соотношением

$$P(u) - P(v) = \int_0^1 P'_{u+t(u-v)}(u-v)\,dt, \tag{63.2}$$

верным, если обе его части имеют смысл.

Понятие производной естественным образом относится и к функционалам. Пусть для функционала F верны следующие допущения.

I′. Область $D(F)$ определения функционала F линейна и плотна в данном банаховом пространстве B.

II′. На любой конечномерной гиперплоскости из области своего определения функционал F непрерывно дифференцируем.

Из этих допущений вытекает существование предела

$$\lim_{t \to 0} \frac{F(u+th)-F(u)}{t} = \frac{d}{dt} F(u+th)\big|_{t=0} \tag{63.3}$$

для любых u, $h \in D(F)$. Этот предел есть дифференциал Гато функционала F; если при некотором $u \in D(F)$ предел (63.3) есть линейный (хотя, может быть, и неограниченный) функционал над h, то этот предел есть производная Гато функционала F в точке u. Ниже мы не будем пользоваться понятиями дифференциала и производной Гато от функционала; значительно более важным оказывается понятие *градиента функционала*.

Рассмотрим множество $M \subset D(F)$, обладающее тем свойством, что если $u \in M$, то предел (63.3) есть линейный *ограниченный* функционал над h. На множестве M определен некоторый оператор P, действующий из данного банахова пространства B в сопряженное пространство B^* такой, что

$$\frac{d}{dt} F(u + th)\Big|_{t=0} = (Pu, h). \tag{63.4}$$

Ниже мы всегда будем предполагать, что множество $M = D(P)$ линейно.

Оператор P называется *градиентом функционала* F, который в свою очередь называется *потенциалом оператора* P. Между потенциалом и его градиентом существует простое соотношение

$$F(u) = F(u_0) + \int_0^1 (P(u_0 + t(u - u_0)),\ u - u_0))\,dt, \tag{63.5}$$

верное, если u и u_0 — произвольные элементы из области определения градиента. В частности, если нулевой элемент принадлежит этой области, то можно положить $u_0 = 0$, и тогда

$$F(u) = \int_0^1 (P(tu),\ u)\,dt + \text{const}. \tag{63.6}$$

Вид градиента и область его определения зависят от того, в каком пространстве рассматривается данный функционал. Чтобы пояснить это утверждение, рассмотрим функционал

$$\Phi(u) = \frac{1}{2}\int_0^1 u'^2(x)\,dx;$$

за область его определения примем множество функций одной вещественной переменной $x \in [0, 1]$, которые: 1) абсолютно непрерывны на отрезке $[0, 1]$; 2) имеют первые производные, суммируемые с квадратом на этом отрезке; 3) на его концах обращаются в нуль, так что $u(0) = u(1) = 0$,

Имеем:

$$\frac{d}{dt}\Phi(u+th)\big|_{t=0}=\int_0^1 u'(x)\,h'(x)\,dx.$$

Область $D(\Phi)$ определения функционала Φ превращается в полное гильбертово пространство, если в этой области ввести скалярное произведение и норму по формулам

$$[u, v]=\int_0^1 u'(x)\,v'(x)\,dx, \quad |u|^2=\int_0^1 u'^2(x)\,dx;$$

обозначим это пространство через H_0. Если рассматривать Φ как функционал в H_0, то

$$\frac{d}{dt}\Phi(u+th)\big|_{t=0}=[u, h],$$

правая часть этого неравенства при любом u представляет собой ограниченный линейный функционал над h. Отсюда следует, что градиент Φ определен на всем пространстве H_0 и, очевидно, совпадает с тождественным в этом пространстве оператором.

С другой стороны, область определения функционала Φ можно рассматривать, например, как линейное множество в пространстве $H=L_2(0, 1)$. В этом пространстве функционал над $h\in D(\Phi)$, равный

$$\int_0^1 u'(x)\,h'(x)\,dx,$$

в котором $u(x)$ принадлежит области $D(\Phi)$ определения функционала Φ, ограничен тогда и только тогда, когда $u'(x)$ абсолютно непрерывна, а $u''(x)$ суммируема с квадратом на отрезке $[0, 1]$. В этом случае

$$\int_0^1 u'(x)\,h'(x)\,dx=-\int_0^1 u''(x)\,h(x)\,dx=(-u'', h),$$

и ясно, что $P=\operatorname{grad}\Phi=-\frac{d^2}{dx^2}$, причем область $D(P)$ определения оператора P содержит не все функции из $D(\Phi)$, а только те из них, которые, как было только что указано, имеют на отрезке $[0, 1]$ абсолютно непрерывные первые производные и суммируемые с квадратом вторые производные.

Функционал F называется *возрастающим*, если $F(u)\to+\infty$ тогда и только тогда, когда $\|u\|\to\infty$. Если функционал задан не в нормированном, а в произвольном метрическом пространстве с метрической функцией $\rho(u, v)$, то условие $\|u\| \to \infty$ заменяется таким: $\rho(u, u_0) \to \infty$, где u_0 — фиксированный элемент пространства.

Функционал F называется *полунепрерывным снизу* (соответственно *сверху*) в точке u_0, если по данному $\varepsilon > 0$ можно найти такое $\delta > 0$, что если $\rho(u, u_0) < \delta$, то $F(u) - F(u_0) > -\varepsilon$ (соответственно $F(u_0) - F(u) > -\varepsilon$). Иначе это можно сформулировать так: функционал F в точке u_0 полунепрерывен снизу (соответственно сверху), если $\varliminf\limits_{u_n \to u_0} F(u_n) \geqslant F(u_0)$ $\Big($соответственно $\varlimsup\limits_{u_n \to u_0} F(u_n) \leqslant F(u_0)\Big)$.

Функционал F называется *слабо полунепрерывным снизу* (соответственно *сверху*) в точке u_0, если соотношение $\varliminf F(u_n) \geqslant F(u_0)$ (соответственно $\varlimsup F(u_n) \leqslant F(u_0)$) имеет место при условии, что u_n слабо сходится к u_0'.

Функционал полунепрерывен на некотором множестве, если он полунепрерывен в любой точке этого множества.

Функционал F непрерывен в точке u_0, если по данному $\varepsilon > 0$ можно найти такое $\delta > 0$, что при $\|u - u_0\| < \delta$ необходимо $|F(u) - F(u_0)| < \varepsilon$ и функционал непрерывен на некотором множестве, если он непрерывен в каждой точке этого множества. Очевидно, функционал непрерывен тогда и только тогда, когда он одновременно полунепрерывен сверху и снизу.

Приведем теорему В. И. Казимирова [1], которая является важным средством для установления полунепрерывности функционалов довольно широкого класса.

Теорема 63.1. *Пусть*

$$F(U, P) = \int_\Omega G(x, u_1, u_2, \ldots, u_s; p_1, p_2, \ldots, p_k)\, dx = \\ = \int_\Omega G(x, U, P)\, dx, \qquad (63.7)$$

где подынтегральная функция определена при $x \in \Omega$ и при любых значениях переменных u_i и p_j, и пусть во всей области своего определения эта функция обладает следующими свойствами:

1) *функция G непрерывна вместе со своими производными вида $\dfrac{\partial G}{\partial p_j}$;*

2) *функция G неотрицательна;*

3) *имеет место неравенство*

$$G(x, U, P) - G(x, U, \bar{P}) - \sum_{j=1}^{k} (p_j - \bar{p}_j) \frac{\partial G(x, U, \bar{P})}{\partial p_j} \geqslant 0 \quad (63.8)$$

при любых $x \in \Omega$, U, P, $\bar{P}$.

Если функции $u_{in}(x)$ *сильно сходятся к функциям* u_{i0} *в метрике некоторого пространства* $L_r(\Omega)$, $1 < r < \infty$, *а функции* $p_{jn}(x)$ *в том же пространстве слабо сходятся к функциям* $p_{j0}(x)$, *то*

$$\varliminf_{n\to\infty} F(U_n, P_n) \geqslant F(U_0, P_0).$$

Заметим, что неравенство (63.8) выполняется, если матрица из вторых производных вида

$$\frac{\partial^2 G}{\partial p_\alpha \, \partial p_\beta} \qquad (\alpha, \beta = 1, 2, \ldots, k)$$

неотрицательная при всех значениях аргументов $x \in \Omega$, U, P. Это сразу видно из разложения функции G в строку Тейлора.

§ 64. Положительные операторы в банаховых пространствах [1])

Этот параграф, как и предшествующий, является вспомогательным; его назначение — распространить понятие энергетического пространства на случай *линейных* операторов, действующих в банаховых пространствах.

Пусть линейный оператор A с плотной областью определения действует из рефлексивного банахова пространства B в сопряженное с ним пространство B^*, тогда сопряженный оператор A^* также действует из B в B^*. На такие операторы естественным образом распространяются важные для вариационных методов понятия, обычно связываемые с гильбертовым пространством, а именно, оператор A: 1) симметричен, если $A \subset A^*$; 2) самосопряжен, если $A = A^*$; 3) положителен, если он симметричен и если $(Au, u) > 0$, $u \neq 0$; 4) положительно определен, если он симметричен и если $(Au, u) \geqslant \gamma^2 \| u \|^2$, $\gamma = \text{const} > 0$. С положительным оператором A, действующим из B в B^*, можно связать некоторое *гильбертово* пространство H_A, которое будем, как и в предшествующих главах, называть энергетическим. Действительно, пусть $u, v \in D(A)$. Тогда выражение (Au, v), как легко видеть, удовлетворяет всем аксиомам скалярного умножения в вещественном гильбертовом пространстве:

1. $(Au, v) = (u, Av) = (Av, u)$, так как положительный оператор A симметричен.

2. $(A(\alpha u + \beta v), w) = \alpha (Au, w) + \beta (Av, w)$ в силу линейности оператора A.

3. $(Au, u) \geqslant 0$.

4. $(Au, u) = 0$ тогда и только тогда, когда $u = 0$.

[1]) См. М. Ш. Бирман [1].

Аксиомы 3 и 4 вытекают из положительности рассматриваемого оператора.

Введя скалярное («энергетическое») произведение

$$[u, v]_A = (Au, v), \qquad u, v \in D(A), \tag{64.1}$$

мы превратим множество $D(A)$ в гильбертово пространство; если оно окажется неполным — пополним его. Полученное таким образом полное гильбертово пространство будем обозначать через H_A и называть *энергетическим пространством оператора* A.

Если A — положительно определенный оператор, то пространство H_A вкладывается в исходное пространство B; этого не будет, если A — положительный, но не положительно определенный оператор.

Задача о минимуме квадратичного функционала

$$(Au, u) - 2(f, u), \qquad u \in B, \qquad f \in B^*$$

решается слово в слово так же, как и в случае, когда пространство B — гильбертово; мы не станем на этом останавливаться.

§ 65. Некоторые теоремы вариационного исчисления

Будем рассматривать функционал $\Phi(u)$, определенный на линейном множестве $D(\Phi)$, плотном в некотором рефлексивном банаховом пространстве B. В соответствии со сказанным в § 63 будем считать, что названный функционал непрерывно дифференцируем на любой конечномерной гиперплоскости, входящей в $D(\Phi)$. Положим $P = \operatorname{grad} \Phi$ и допустим, что область $D(P)$ определения оператора P также линейна и плотна в B; очевидно, $D(P) \subset D(\Phi)$. Примем также, что оператор P сильно дифференцируем на любой конечномерной гиперплоскости из области его определения, и что производная Гато P'_u определена на плотном в B множестве $D(P'_u)$.

Положим

$$F(u) = \Phi(u) - (f, u), \tag{65.1}$$

где f — фиксированный элемент сопряженного пространства B^*.

Теорема 65.1 [1]. *Если элемент u_0 доставляет относительный минимум функционалу $F(u)$, то $u_0 \in D(P)$ и*

$$Pu_0 = f. \tag{65.2}$$

Действительно, в данном случае для любого $h \in D(\Phi)$

$$\frac{d}{dt} F(u_0 + th)\big|_{t=0} = \frac{d}{dt} \Phi(u_0 + th)\big|_{t=0} - (f, h) = 0.$$

[1] См., например, И. В. Гельман [1].

Отсюда

$$\frac{d}{dt}\Phi(u_0+th)\big|_{t=0}=(f,\ h),$$

и выражение слева есть ограниченный функционал над h. Но тогда $u_0 \in D(P)$ и $Pu_0=f$, что и требовалось доказать.

В некотором смысле обратной к только что доказанной теореме является

Теорема 65.2[1]). *Пусть функционал Φ и его градиент P определены соответственно на линейных множествах $D(\Phi)$ и $D(P) \subset D(\Phi)$, плотных в банаховом пространстве B. Пусть производная P'_u существует и положительна при любом $u \in D(P)$ и элемент $P'_u h$ непрерывно меняется, когда h фиксировано, а u непрерывно меняется вдоль любой прямой, и пусть $D(P'_u) \supset D(P)$ при любом u.*

Если существует решение u_0 уравнения (65.2), то это решение единственно и

$$F(u_0)=\min_{u \in D(P)} F(u), \tag{65.3}$$

где функционал F определен формулой (65.1); если, кроме того, функционал Φ полунепрерывен сверху, то

$$F(u_0)=\min_{u \in D(\Phi)} F(u). \tag{65.4}$$

Пусть u_0 удовлетворяет уравнению (65.1), и пусть это уравнение имеет еще одно решение u_1. Тогда

$$(Pu_1-Pu_0,\quad u_1-u_0)=0.$$

Положим $\xi=u_0+t(u_1-u_0)$, где t — вещественное число. Имеем:

$$\frac{d}{dt}P(\xi)=\frac{d}{d\tau}P(u_0+(t+\tau)(u_1-u_0))\big|_{\tau=0}=P'_\xi(u_1-u_0).$$

Отсюда

$$\int_0^1 P'_\xi(u_1-u_0)\,dt=\int_0^1 \frac{d}{dt}P(\xi)\,dt=P(t(u_1-u_0)+u_0)\big|_{t=0}^{t=1}=Pu_1-Pu_0.$$

Теперь

$$0=(Pu_1-Pu_0,\ u_1-u_0)=\int_0^1 \left(P'_\xi(u_1-u_0),\ u_1-u_0\right)dt,$$

и так как производная P'_ξ положительна, то $u_1=u_0$, и решение уравнения (65.1) единственно.

---

[1]) Эта теорема весьма близка к теореме, доказанной А. Лангенбахом [1]—[3] и ранее, но в более жестких условиях, М. М. Вайнбергом [1].

Докажем теперь, что u_0 реализует минимум функционала (65.2) на множестве $D(P)$. Для этого установим тождество

$$F(u+h)-F(u)=\int_0^1 (P(u+\tau h),\ h)\,d\tau-(f,\ h),\quad u,\ h\in D(P). \tag{65.5}$$

По формуле (63.6)

$$F(u+h)-F(u)=$$
$$=\int_0^1 [(Pt(u+h),\ u+h)-(Ptu,\ u)]\,dt-(f,\ h)=$$
$$=\int_0^1 (Pt(u+h),\ h)\,dt+\int_0^1 (Pt(u+h)-Ptu,\ u)\,dt-(f,\ h). \tag{65.6}$$

Второй интеграл преобразуем так:

$$\int_0^1 (Pt(u+h)-Ptu,\ u)\,dt=\int_0^1 dt\int_0^t \frac{d}{d\tau}(P(tu+\tau h),\ u)\,d\tau=$$
$$=\int_0^1 d\tau\int_\tau^1 \left(P'_{tu+\tau h}h,\ u\right)dt=\int_0^1 d\tau\int_\tau^1 \left(P'_{tu+\tau h}u,\ h\right)dt=$$
$$=\int_0^1 (P(u+\tau h)-P\tau(u+h),\ h)\,d\tau.$$

Подставив это в формулу (65.6), мы получим тождество (65.5).

Пусть u — произвольный элемент области $D(P)$. Полагая $u=$ $=u_0+h$, имеем по формуле (65.5)

$$F(u_0+h)-F(u_0)=\int_0^1 (P(u_0+\tau h),\ h)\,d\tau-(f,\ h)=$$
$$=\int_0^1 (P(u_0+\tau h)-P(u_0),\ h)\,d\tau=$$
$$=\int_0^1 \frac{d\tau}{\tau}\int_0^1 \left(P'_{u_0+t\tau h}\tau h,\ \tau h\right)dt\geqslant 0,$$

что и требовалось доказать.

Допустим теперь, что функционал Φ полунепрерывен сверху, и докажем соотношение (65.4). Если оно неверно, то существует элемент $u_1\in D(\Phi)$ такой, что $F(u_1)<F(u_0)$. Множество $D(P)$ плотно в пространстве B, поэтому при любом $\delta>0$ найдется элемент $v\in D(P)$, удовлетворяющий неравенству $\|v-v_1\|<\delta$.

Вместе с Φ полунепрерывен сверху и функционал F; по заданному $\varepsilon > 0$ найдется такое $\delta > 0$, что если $\|u_1 - w\| < \delta$, то $F(u_1) - F(w) > -\varepsilon$. Возьмем

$$\varepsilon = \frac{1}{2}[F(u_0) - F(u_1)]$$

и $w = v$, где v — упомянутый выше элемент из $D(P)$. Тогда, как легко видеть,

$$F(v) < \frac{1}{2}[F(u_0) + F(u_1)] < F(u_0),$$

что противоречит соотношению (65.3).

§ 66. О существовании решения вариационной задачи

Функционал Φ называется *выпуклым на линейном множестве* $M \subset D(\Phi)$, если

$$\Phi(u) + \Phi(v) - 2\Phi\left(\frac{u+v}{2}\right) \geqslant 0, \qquad u,\ v \in M; \tag{66.1}$$

этот функционал называется *существенно выпуклым* на том же множестве, если знак равенства имеет место лишь при $u = v$.

Теорема 66.1. *Если оператор $P = \operatorname{grad}\Phi$ имеет производную, положительную при любом $u \in D(P)$, причем $D(P'_u) \supset D(P)$ каков бы ни был элемент $u \in D(P)$, то функционал Φ существенно выпуклый на множестве $D(P)$*[1]*. Если, кроме того, названный функционал непрерывен, то он выпуклый на множестве $D(\Phi)$. Наконец, если функционал Φ непрерывен, а производная его градиента равномерно положительно ограничена снизу, т. е. если*

$$(P'_u h,\ h) \geqslant \gamma^2 \|h\|^2, \qquad \gamma = \text{const} > 0, \tag{66.2}$$

и, по-прежнему, $D(P'_u) \supset D(P)$, то функционал Φ — существенно выпуклый на множестве $D(\Phi)$.

1) Обозначим $v = u + h$. Если $u, v \in D(P)$, то по формуле (65.5) имеем:

$$\Phi(u+h) + \Phi(u) - 2\Phi\left(u + \frac{h}{2}\right) =$$
$$= \left[\Phi(u+h) - \Phi\left(u + \frac{h}{2}\right)\right] - \left[\Phi\left(u + \frac{h}{2}\right) - \Phi(u)\right] =$$
$$= \int_0^1 \left(P\left(u + \frac{h}{2} + t\,\frac{h}{2}\right) - P\left(u + t\frac{h}{2}\right),\ \frac{h}{2}\right) dt =$$
$$= \int_0^1 \int_0^1 \left(P'_\xi \frac{h}{2},\ \frac{h}{2}\right) dt\, d\tau \geqslant 0, \quad \xi = u + (t+\tau)\frac{h}{2}, \tag{66.3}$$

[1]) Это утверждение доказано А. Лангенбахом [1]—[3], который рассматривал операторы, действующие в гильбертовом пространстве.

ясно также, что знак равенства имеет место лишь при $h=0$ или $u=v$.

2) Допустим еще, что функционал Φ непрерывен; нам надо доказать, что тогда неравенство (66.1) верно для любых $u \neq v \in D(\Phi)$. Если это не так, то в $D(\Phi)$ найдутся элементы $u_0 \neq v_0$ такие, что

$$\Phi(u_0)+\Phi(v_0)-2\Phi\left(\frac{u_0+v_0}{2}\right)=-k<0. \tag{66.4}$$

Положим

$$\varepsilon=\frac{k}{6}.$$

Функционал Φ непрерывен, поэтому существует такое $\delta>0$, что если $\|u_0-u\|<\delta$, $\|v_0-v\|<\delta$, то

$$\left.\begin{gathered}\Phi(u_0)>\Phi(u)-\varepsilon, \qquad \Phi(v_0)>\Phi(v)-\varepsilon,\\ \Phi\left(\frac{u_0+v_0}{2}\right)<\Phi\left(\frac{u+v}{2}\right)+\varepsilon.\end{gathered}\right\} \tag{66.5}$$

Так как множество $D(P)$ плотно в B, то упомянутые здесь элементы u и v можно взять из этого множества. Из соотношений (66.4) и (66.5) следует, что

$$\Phi(u)+\Phi(v)-2\Phi\left(\frac{u+v}{2}\right)<-k+4\varepsilon=-\frac{k}{3}<0,$$

а это противоречит доказанному в п. 1.

3) Примем теперь, что функционал Φ непрерывен и удовлетворяет соотношению (66.2). Допустим, что наш функционал не существенно выпуклый на $D(\Phi)$. Тогда найдутся элементы u_0, $v_0 \in D(\Phi)$, $u_0 \neq v_0$ такие, что

$$\Phi(u_0)+\Phi(v_0)-2\Phi\left(\frac{u_0+v_0}{2}\right)\leqslant 0.$$

Зададимся произвольно малым числом $\varepsilon>0$ и найдем такое $\delta>0$, что при u, $v \in D(P)$, $\|u-u_0\|<\delta$, $\|v-v_0\|<\delta$, будет

$$\Phi(u)+\Phi(v)-2\Phi\left(\frac{u+v}{2}\right)<\varepsilon. \tag{66.6}$$

Число δ можно взять сколь угодно малым. Возьмем

$$\delta\leqslant\frac{1}{3}\|u_0-v_0\|, \text{ тогда } \|u-v\|>\frac{1}{3}\|u_0-v_0\|.$$

Далее, как в п. 2, можно считать, что u, $v \in D(P)$. Из соотношений (66.2) и (66.3) следует, что

$$\Phi(u)+\Phi(v)-2\Phi\left(\frac{u+v}{2}\right)\geqslant\frac{\gamma^2}{36}\|u_0-v_0\|^2,$$

что противоречит неравенству (66.6), если ε достаточно мало.

Теорема 66.2[1]). *Пусть функционал F определен на некотором банаховом пространстве, в котором сфера слабо компактна, и пусть он возрастающий, существенно выпуклый и слабо полунепрерывный снизу. Тогда он ограничен снизу и его нижняя грань достигается в единственной точке, к которой слабо сходится любая минимизирующая последовательность.*

Доказательство. Прежде всего, функционал F ограничен снизу. Действительно, пусть существует последовательность $\{u_n\}$, $F(u_n) \to -\infty$. Так как функционал возрастающий, то последовательность необходимо ограничена и потому слабо компактна. Пусть $\{u_{n_k}\}$ — слабо сходящаяся подпоследовательность последовательности $\{u_n\}$ и пусть u_0 — слабый предел этой подпоследовательности. Функционал F слабо полунепрерывен снизу, поэтому $F(u_0) \leqslant \leqslant \varliminf (u_{n_k}) = -\infty$, что нелепо.

Обозначим $d = \inf F$ и пусть теперь $\{u_n\}$ обозначает минимизирующую последовательность. Как и выше, убедимся, что она ограничена и потому слабо компактна; некоторая ее подпоследовательность $\{u_{n_k}\}$ имеет слабый предел u_0. Из того, что функционал полунепрерывен снизу, вытекает, что $d = \lim F(u_{n_k}) \geqslant F(u_0)$. Но $d = \inf F(u)$, поэтому $F(u_0) = d$ и нижняя грань функционала F достигается в точке u_0. Такая точка единственная: если бы в точке $u_1 \neq u_0$ было $F(u_1) = d$, то из существенной выпуклости функционала F следовало бы:

$$F\left(\frac{u_0 + u_1}{2}\right) < \frac{1}{2} F(u_0) + \frac{1}{2} F(u_1) = d,$$

что невозможно.

Остается доказать, что минимизирующая последовательность $\{u_n\}$ слабо сходится к u_0. Допустим противное. Тогда найдется такой линейный ограниченный функционал f, что

$$(f, u_n) \not\to f(u_0).$$

Отсюда, в свою очередь, следует существование подпоследовательности $\{u_{n_k}\}$ и числа $\varepsilon_0 > 0$ таких, что

$$\left|(f, u_{n_k}) - (f, u_0)\right| \geqslant \varepsilon_0. \tag{66.7}$$

Последовательность $\{u_{n_k}\}$ также минимизирующая; по доказанному выше, из нее можно выделить подпоследовательность — обозначим ее $\{u_{n_{k_l}}\}$, — которая слабо сходится к u_0. Это противоречит неравенству (66.7).

[1]) См. В. Гельман [1].

В уже цитированной работе [1] И. В. Гельмана доказано существование решения задачи о минимуме функционала

$$\left.\begin{aligned} &\int\limits_{\Omega} F\left(x_1, \ldots, x_m, \ldots, u^{i,\,j}_{j_1, \ldots, j_m}, \ldots\right) dx, \\ &u^{i,\,j}_{j_1, \ldots, j_m}\Big|_S = 0 \qquad (i = 1, 2, \ldots, N, \ j = 0, 1, \ldots, l-1), \end{aligned}\right\} \quad (66.8)$$

здесь Ω — конечная область m-мерного евклидова пространства S — ее граница,

$$u^{i,\,j}_{j_1, \ldots, j_m} = \frac{\partial^j u^{i,\,0}_{0, \ldots, 0}}{\partial x_1^{j_1} \ldots \partial x_m^{j_m}};$$

предполагается, что $u^{i,\,0}_{0, \ldots, 0} \in W_p^{(l)}(\Omega)$, $1 < p < \infty$. На функцию F накладываются следующие ограничения:

1) Эта функция непрерывна вместе со своими первыми и вторыми частными производными по аргументам $u^{i,\,j}_{j_1, \ldots, j_m}$ при любых значениях этих последних и при $x \in \overline{\Omega}$.

2) Имеет место двойное неравенство

$$K_1 \sum_{i=1}^{N} \left[\sum_{j_1 + \ldots + j_m = l} \left| u^{i,\,l}_{j_1, \ldots, j_m} \right|^2 \right]^{\frac{p}{2}} \leqslant F \leqslant$$

$$\leqslant K_2 \sum_{i=1}^{N} \left[\sum_{j=0}^{l} \sum_{j_1 + \ldots + j_m = j} \left| u^{i,\,j}_{j_1, \ldots, j_m} \right|^2 \right]^{\frac{p}{2}}, \qquad K_1, K_2 = \mathrm{const} > 0. \tag{66.9}$$

3) Матрица элементов

$$\frac{\partial^2 F}{\partial u^{i,\,j}_{j_1, \ldots, j_m} \partial u^{k,\,s}_{s_1, \ldots, s_m}} \tag{66.10}$$

— положительно определенная.

Доказательство существования решения этой вариационной задачи опирается, кроме теоремы 66.2, еще на теорему В. И. Казимирова 63.1.

Другой подход к решению вариационных задач разработан в статьях А Лангенбаха[1]) [1] — [3].

Рассмотрим функционал (65.1) в предположении, что производная P'_u, где $P = \operatorname{grad} \Phi$, есть оператор, положительный при любом $u \in D(P)$. Сузим область определения функционала Φ, рассматривая последний только на множестве $D(P)$. Для упрощения записи

---

[1]) В работах А. Лангенбаха принимается, что функционал и его градиент определены в некотором гильбертовом пространстве.

положим $P(0)=0$[1]. Пренебрегая в формуле (63.6) постоянной, безразличной для задачи о минимуме, имеем:

$$F(u)=\int_0^1 (P(tu),\ u)-(f,\ u), \qquad D(F)=D(P). \tag{66.11}$$

Из теоремы 65.2 следует, что решение уравнения $Pu=f$ реализует минимум функционала (66.11), а из теоремы 65.1 вытекает обратное утверждение. Допустим дополнительно, что производная P'_u удовлетворяет еще неравенству (66.2), так что она положительно ограничена снизу равномерно по u. Докажем, что при этом условии функционал (66.11) ограничен снизу. Действительно, так как $P0=0$, то

$$F(u)=\int_0^1 (P(tu)-P0,\ u)\,dt-(f,\ u).$$

По формуле (63.2)

$$P(tu)-P(0)=\int_0^1 P'_{(t+\tau)u}tu\,d\tau.$$

Теперь

$$F(u)=\int_0^1\int_0^1 t\left(P'_{(t+\tau)u}u,\ u\right)dt\,d\tau-(u,\ f)$$

и по неравенству (66.2)

$$F(u)\geqslant\frac{1}{2}\gamma^2\|u\|^2-(u,\ f).$$

Применив неравенство Коши, получим:

$$F(u)\geqslant\frac{1}{2}\gamma^2\|u\|^2-\|u\|\,\|f\|=\frac{\gamma^2}{2}\left(\|u\|-\frac{\|f\|}{\gamma^2}\right)^2-\frac{\|f\|^2}{2\gamma^2}\geqslant-\frac{\|f\|^2}{2\gamma^2},$$

и функционал (66.1) ограничен снизу.

Обозначим

$$d=\inf F(u).$$

Теорема 66.3. *При выполнении неравенства* (66.2) *любая минимизирующая последовательность для функционала* (65.1) *сходится в метрике пространства B к некоторому пределу, который не зависит от выбора минимизирующей последовательности.*

---

[1] Это не умаляет общности: достаточно уравнение $Pu=f$ заменить равносильным ему уравнением $Pu-P0=f-P0$. При этом оператор Pu заменяется на $Pu-P0$, а функционал $\Phi(u)$ — на $\Phi(u)-(P0,\ u)$; такая замена не отражается на свойствах возрастания, полунепрерывности или непрерывности и выпуклости функционала.

По теореме 66.1 функционал Φ существенно выпуклый. Положим

$$\left.\begin{aligned} \rho(u, v) &= \left[\Phi_1(u) + \Phi(v) - 2\Phi\left(\frac{u+v}{2}\right)\right]^{\frac{1}{2}}, \\ \Phi(u) &= \int_0^1 (P(tu), u)\, dt. \end{aligned}\right\} \tag{66.12}$$

По формуле (66.3) имеем:

$$\rho(u, v) = \left[\int_0^1 \int_0^1 \left(P'_\xi \frac{h}{2}, \frac{h}{2}\right) dt\, d\tau\right]^{\frac{1}{2}}, \qquad h = u - v, \tag{66.13}$$

и, в силу неравенства (66.2),

$$\rho(u, v) \geqslant \frac{\gamma}{2}\| u - v \|. \tag{66.14}$$

Пусть $\{u_n\}$ — минимизирующая последовательность для функционала F, так что $F(u_n) \to d$. Оценим величину $\rho(u_m, u_n)$. Заметим прежде всего, что справедлива формула

$$\rho(u, v) = \left[F(u) + F(v) - 2F\left(\frac{u+v}{2}\right)\right]^{\frac{1}{2}}.$$

Положим здесь $u = u_m$, $v = u_n$. Имея в виду, что

$$F\left(\frac{u_m + u_n}{2}\right) \geqslant \inf F(u) = d,$$

получаем:

$$\rho(u_m, u_n) \leqslant [F(u_m) + F(u_n) - 2d]^{\frac{1}{2}}.$$

Возьмем m и n столь большими, чтобы выполнялись неравенства $F(u_m) < d + \varepsilon$, $F(u_n) < d + \varepsilon$, где ε — произвольно заданная положительная величина. Тогда $\rho(u_m, u_n) < \sqrt{2\varepsilon}$ и

$$\lim_{m, n \to \infty} \rho(u_m, u_n) = 0.$$

По неравенству (66.14) $\| u_m - u_n \| \underset{m, n \to \infty}{\to} 0$ и минимизирующая последовательность сходится в метрике пространства B к некоторому элементу u_0. Этот элемент не зависит от выбора минимизирующей последовательности: если $\{u_n\}$ и $\{v_n\}$ — две такие последовательности, то последовательность

$$u_1, v_1, u_2, v_2, \ldots, u_n, v_n, \ldots$$

— также минимизирующая и, по доказанному, имеет предел, а тогда ее подпоследовательности $\{u_n\}$ и $\{v_n\}$ имеют один и тот же предел

Общий предел u_0 минимизирующих последовательностей назовем *обобщенным решением* задачи о минимуме функционала (66.11).

Отметим случай, когда производная P'_u удовлетворяет не только неравенству (66.2), но и более сильному неравенству

$$(P'_u h,\ h) \geqslant \gamma_1^2 (P'_{u_0} h,\ h), \qquad \gamma_1^2 = \text{const} > 0, \tag{66.15}$$

где элемент u_0 — фиксированный. Обозначим через $B_0 = B_{Pu_0}$ энергетическое пространство линейного оператора Pu_0, а через $[\ ,\ ]_0$ и $|\ \ |_0$ — энергетические произведение и норму в этом пространстве. Тогда $(P'_u h,\ h) \geqslant \gamma_1^2 |h|_0^2$ и из тождества (66.13) получаем:

$$\rho(u,\ v) \geqslant \frac{\gamma_1^2}{4} |u - v|_0^2. \tag{66.16}$$

Если $\{u_n\}$ — минимизирующая последовательность для функционала (65.1), то, как мы видели, $\rho(u_m,\ u_n) \underset{m,\ n \to \infty}{\longrightarrow} 0$. В силу неравенства (66.16)

$$|u_m - u_n|_0 \underset{m,\ n \to \infty}{\longrightarrow} 0,$$

и следовательно, обобщенное решение вариационной задачи принадлежит не только исходному пространству B, но и энергетическому пространству B_0.

§ 67. Энергетическое пространство нелинейной задачи

Сохраним предположения конца предшествующего параграфа: пусть область определения функционала Φ совпадает с областью определения его градиента P и пусть производная P'_u удовлетворяет неравенству (66.2). В этом случае можно ввести понятие энергетического пространства для функционала Φ, в известной мере аналогичное тому, которое было введено для линейных задач.

Область $D(\Phi) = D(P)$ определения функционала Φ и его градиента превратим в топологическое пространство, введя в этой области замкнутые множества: объявим таковыми множества вида $\rho(u,\ v) \leqslant \varepsilon$ и $\rho(u,\ v) \geqslant \varepsilon$. Здесь $\rho(u,\ v)$ определяется формулой (66.12), ε — любое положительное число, u — фиксированный, v — переменный элемент из $D(\Phi)$; замкнутыми назовем также объединения конечного числа множеств упомянутого вида и пересечения любых множеств таких множеств.

Если M — любое подмножество из $D(\Phi)$, то его замыканием $\overline{M}$ назовем, как обычно, пересечение всех замкнутых множеств, содержащих M; если M — пустое множество, или $M = D(\Phi)$, то положим $\overline{M} = M$. Как известно, так определенная операция замыкания удовлетворяет известным аксиомам Хаусдорфа[1]).

[1]) Ф. Хаусдорф [1], стр. 110—112.

Пусть последовательность $u_n \in D(P)$ $(n=1,\ 2,\ \ldots)$ такова, что $\rho(u_n,\ u_m) \xrightarrow[n,\ m\to\infty]{} 0$. Неравенство (66.2) показывает, что наша последовательность сходится в себе в метрике пространства B и, следовательно, существует такой элемент $\tilde{u} \in B$, что $\|u_n - \tilde{u}\| \to 0$. Присоединим множество всех таких «предельных» элементов $\tilde{u}$ к области $D(\Phi)$; полученное таким образом новое множество назовем «энергетическим пространством» оператора P и будем обозначать его через B_P. Название «энергетическое пространство» весьма условно, так как на множестве B_P мы не вводим в общем случае никакой топологии.

Отметим случай, когда в энергетическом пространстве можно ввести топологию (и даже метрику).

Из формулы (66.12) видно, что величина $\rho(u,\ v)$ удовлетворяет двум из трех аксиом метрики: 1) $\rho(u,\ v)=\rho(v,\ u)$; 2) $\rho(u,\ v)\geqslant 0$, причем $\rho(u,\ v)=0$ тогда и только тогда, когда $v=u$.

Пусть функционал $\rho(u,\ v)$ удовлетворяет «ослабленному неравенству треугольника»

$$\rho(u,\ v)\leqslant a\,[\rho(u,\ w)+\rho(w,\ v)],\qquad a=\text{const}. \tag{67.1}$$

Очевидно, $a\geqslant 1$. Если $a=1$, то $\rho(u,\ v)$ — метрическая функция и энергетическое пространство B_P есть попросту пополнение пространства $D(P)$ в метрике ρ. Если $a>1$, то[1]) существует метрическая функция $\tilde{\rho}(u,\ v)$, топологически эквивалентная функции $\rho(u,\ v)$; в этом случае B_P есть пополнение пространства $D(P)$ в метрике $\tilde{\rho}$.

Условие (67.1) выполнено, например, в следующем случае: существуют такие положительные постоянные α и β, что для любых элементов $u,\ h \in D(P)$

$$\alpha^2(P_0'h,\ h)\leqslant(P_u'h,\ h)\leqslant\beta^2(P_0'h,\ h); \tag{67.2}$$

здесь P_0' — означает производную оператора P при $u=0$. Обозначим через $H_0=H_{P_0'}$ энергетическое пространство линейного оператора P_0' (см. § 64). Неравенство (67.2) можно переписать так:

$$\alpha^2\,|h|_0^2\leqslant(P_u'h,\ h)\leqslant\beta^2\,|h|_0^2.$$

---

[1]) См. Е. У. Читтенден [1], где доказано следующее более общее утверждение: пусть на множестве M определена функция $\rho(x,\ y)$ пары элементов $x,\ y$, обладающая следующими свойствами: $\rho(x,\ y)=\rho(y,\ x)$; $\rho(x,\ y)\geqslant 0$; $\rho(x,\ y)=0$ тогда и только тогда, когда $x=y$; существует функция $f(t)$, определенная при $t\geqslant 0$, причем $\lim\limits_{t\to 0} f(t)=0$, такая, что если $\rho(x,\ z)<t$ и $\rho(z,\ y)<t$, то $\rho(x,\ y)<f(t)$. Тогда на множестве M можно определить метрическую функцию $\tilde{\rho}(x,\ y)$, топологически эквивалентную функции $\rho(x,\ y)$. Мы называем функцию $\tilde{\rho}(x,\ y)$ метрической, если она удовлетворяет известным аксиомам расстояния: $\tilde{\rho}(x,\ y)=\tilde{\rho}(y,\ x)$; $\tilde{\rho}(x,\ y)\geqslant 0$, причем $\tilde{\rho}(x,\ y)=0$, тогда и только тогда, когда $x=y$; $\rho(x,\ y)\leqslant\rho(x,\ z)+\rho(z,\ y)$.

Из формулы (63.13) следует теперь:

$$\frac{\alpha}{2}|u-v|_0 \leqslant \rho(u, v) \leqslant \frac{\beta}{2}|u-v|_0,$$

и неравенство (67.1) выполняется при значении постоянной $a=\frac{\beta}{\alpha}$.

§ 68. Функционалы теории пластичности и их обобщение[1])

Некоторые из задач теории пластичности могут быть сведены к задаче о минимуме функционала

$$F(u)=\int_\Omega \sum_{j=1}^{s} \int_0^{\tau_j(u)} g_j(\xi)\,d\xi\,dx - \int_\Omega fu\,dx. \tag{68.1}$$

Здесь Ω — конечная область m-мерного евклидова пространства, $f(x)$ — функция, определенная почти всюду в Ω и принадлежащая тому или иному функциональному пространству, u — скалярная или векторная функция, подчиненная некоторым однородным краевым условиям, $\tau_j(u)$ $(j=1, 2, \ldots, s)$ — неотрицательные квадратичные формы относительно функции u и ее производных; коэффициенты форм $\tau_j(u)$ могут быть и функциями от x. Далее, $g_j(\xi)$ $(j=1, 2, \ldots, s)$ — неотрицательные функции от ξ, определенные почти всюду и локально суммируемые на полуоси $0 \leqslant \xi < \infty$. Существует хотя бы один номер k, $1 \leqslant k \leqslant s$, такой, что

$$g_k(\xi) \geqslant a, \qquad a=\text{const}>0, \tag{68.2}$$

и если функция u удовлетворяет краевым условиям задачи, то

$$\int_\Omega \tau_k(u)\,dx \geqslant \gamma_0^2 \int_\Omega u^2\,dx, \qquad \gamma_0=\text{const}>0. \tag{68.3}$$

Приведем несколько примеров.

1. Кручение упрочняющихся стержней[2]). Если Ω — односвязная область плоскости (x, y), то уравнение и краевое условие задачи упруго-пластического кручения можно привести к виду

$$P(u)=-\frac{\partial}{\partial x}\left[\bar{g}(T^2)\frac{\partial u}{\partial x}\right]-\frac{\partial}{\partial y}\left[\bar{g}(T^2)\frac{\partial u}{\partial y}\right]=\omega, \tag{68.4}$$

$$u|_S=0. \tag{68.5}$$

Здесь, как обычно, S — граница области Ω. Далее,

$$T^2=(\operatorname{grad} u)^2=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2.$$

---

[1]) См. А. Лангенбах [2].

[2]) См. Л. М. Качанов [3].

ω — угол закручивания стержня на единицу его длины, $g(T^2)$ — функция, характерная для материала стержня в состоянии упрочнения и входящая в уравнение связи

$$\Gamma = \bar{g}(T^2)\,T \tag{68.6}$$

между максимальным касательным напряжением T и интенсивностью деформаций сдвига Γ.

Хорошо известно, что Γ — возрастающая функция от T, поэтому $\frac{d\Gamma}{dT} \geqslant 0$ или

$$\bar{g}(\xi^2) + 2\bar{g}'(\xi^2)\,\xi^2 \geqslant 0.$$

Мы примем, что функция $\bar{g}$ удовлетворяет более сильному неравенству

$$\bar{g}(\xi^2) + 2\bar{g}'(\xi^2)\,\xi^2 \geqslant C_1, \quad C_1 = \text{const} > 0. \tag{68.7}$$

Известно также, что

$$\bar{g}(\xi^2) \geqslant C_2, \qquad C_2 = \text{const} > 0, \tag{68.8}$$

именно, можно положить $C_2 = \frac{1}{G}$, где G — модуль сдвига материала в упругом состоянии.

Допустим, что функция $\bar{g}$ имеет две непрерывные производные. Будем рассматривать P как оператор в пространстве $L_2(\Omega)$; за область его определения примем линейное множество функций, дважды непрерывно дифференцируемых в Ω и удовлетворяющих условию (68.5); мы принимаем при этом, что граница S достаточно гладкая. Очевидно, условие II § 63 выполнено. Докажем, что оператор P имеет производную, которая является линейным оператором и удовлетворяет неравенству (66.2). Непосредственное вычисление дает:

$$\begin{gathered}\lim_{t\to 0}\frac{P(u+th)-P(u)}{t} = -\frac{\partial}{\partial x}\left\{\left[\bar{g}(T^2(u)) + 2\bar{g}'(T^2(u))\left(\frac{\partial u}{\partial x}\right)^2\right]\frac{\partial h}{\partial x} + \right.\\ \left. + 2\bar{g}'(T^2(u))\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{\partial h}{\partial y}\right\} - \frac{\partial}{\partial y}\left\{2\bar{g}'(T^2(u))\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{\partial h}{\partial x} + \right.\\ \left. + \left[\bar{g}(T^2(u)) + 2\bar{g}'(T^2(u))\left(\frac{\partial u}{\partial y}\right)^2\right]\frac{\partial h}{\partial y}\right\}, \quad h|_S = 0,\end{gathered} \tag{68.9}$$

что представляет собой оператор, линейный относительно h. Обозначим этот оператор, как обычно, через $P'_h h$; он, очевидно, симметричен. Далее,

$$\begin{gathered}(P'_u h, h) = \iint_\Omega \{\bar{g}(T^2(u))(\operatorname{grad} h)^2 + \\ + 2\bar{g}'(T^2(u))(\operatorname{grad} u \operatorname{grad} h)^2\}\,dx\,dy,\end{gathered} \tag{68.10}$$

где Ω — область сечения стержня.

Функцию $\bar{g}'$ представим как разность ее положительной и отрицательной частей:

$$\bar{g}' = g_1 - g_2, \qquad g_1 = \begin{cases} \bar{g}', & \bar{g}' \geqslant 0, \\ 0, & \bar{g}' < 0, \end{cases} \qquad g_2 = \begin{cases} 0, & \bar{g}' \geqslant 0, \\ -\bar{g}', & \bar{g}' < 0. \end{cases}$$

По неравенству Коши

$$(\operatorname{grad} u \operatorname{grad} h)^2 \leqslant (\operatorname{grad} u)^2 (\operatorname{grad} h)^2 = T^2(u)(\operatorname{grad} h)^2,$$

и из неравенства (68.10) следует:

$$\left(P'_u h,\ h\right) \geqslant \int\limits_{\Omega}\int \{\bar{g}(T^2(u)) - 2g_2(T^2(u))\,T^2(u)\}(\operatorname{grad} h)^2\,dx\,dy. \qquad (68.11)$$

Если $\bar{g}'(T^2(u)) \geqslant 0$, то $g_2 = 0$ и

$$\bar{g}(T^2(u)) - 2g_2(T^2(u))\,T^2(u) = \bar{g}(T^2(u)) \geqslant C_2,$$

если же $\bar{g}'(T^2(u)) < 0$, то $g_2(T^2(u)) = -\bar{g}'(T^2(u))$ и

$$\bar{g}(T^2(u)) - 2g_2(T^2(u))\,T^2(u) = \bar{g}(T^2(u)) + 2\bar{g}'(T^2(u))\,T^2(u) \geqslant C_1.$$

В обоих случаях

$$\bar{g}(T^2(u)) - 2g_2(T^2(u))\,T^2(u) \geqslant \mu = \min(C_1,\ C_2),$$

и формула (68.11) дает:

$$\left(P'_u h,\ h\right) \geqslant \mu \int\limits_{\Omega}\int (\operatorname{grad} u)^2\,dx\,dy. \qquad (68.12)$$

Теперь, в силу неравенства Фридрихса (ВМ, стр. 136),

$$\left(P'_u h,\ h\right) \geqslant \gamma^2 \|u\|^2, \qquad \gamma^2 = \mu\varkappa, \qquad (68.13)$$

где $\varkappa$ — постоянная неравенства Фридрихса.

Заметим, что производная P'_u удовлетворяет не только неравенству (66.2), но и неравенству (66.15), последнему — при $u_0 = 0$. Действительно, если $u = u_0 \equiv 0$, то и $T^2(u) = (\operatorname{grad} u)^2 \equiv 0$, и соотношение (68.10) принимает вид:

$$\left(P'_0 h,\ h\right) = \bar{g}(0) \int\limits_{\Omega}\int (\operatorname{grad} h)^2\,dx\,dy.$$

В силу неравенства (68.12)

$$\left(P'_u h,\ h\right) \geqslant \gamma_1^2 \left(P'_0 h,\ h\right), \qquad \gamma_1^2 = \frac{\mu}{\bar{g}(0)}. \qquad (68.14)$$

Из теорем 65.1 и 65.2 вытекает, что задача (68.4), (68.5) равносильна задаче о минимуме функционала

$$F(u) = \int\limits_0^1 (P(tu),\ u)\,dt - (\omega,\ u), \qquad u|_S = 0.$$

Имеем:

$$\int_0^1 (P(tu),\ u)\,dt =$$

$$= -\int_0^1 t\,dt \int\int_\Omega u\left\{\frac{\partial}{\partial x}\left[\bar g(t^2T^2(u))\,\frac{\partial u}{\partial x}\right] + \frac{\partial}{\partial y}\left[\bar g(t^2T^2(u))\,\frac{\partial u}{\partial y}\right]\right\}dx\,dy,$$

или, если взять внутренний интеграл по частям и воспользоваться краевым условием $u|_S = 0$,

$$\int_0^1 (P(tu),\ u)\,dt = \int_0^1 t\,dt \int\int_\Omega \bar g(t^2T^2(u))\,T^2(u)\,dx\,dy =$$

$$= \int\int_\Omega dx\,dy \int_0^1 t\bar g(t^2T^2(u))\,T^2(u)\,dt.$$

Замена $\xi = t^2T^2(u)$ дает:

$$\int_0^1 (P(tu),\ u)\,dt = \int\int_\Omega dx\,dy \int_0^{T^2(u)} \frac{1}{2}\bar g(\xi)\,d\xi,$$

и следовательно,

$$F(u) = \int\int_\Omega dx\,dy \int_0^{T^2(u)} \frac{1}{2}\bar g(\xi)\,d\xi - \omega\int\int_\Omega u\,dx\,dy. \qquad (68.15)$$

Функционал (68.15) получается из общего выражения (68.1) при $s = 1$, $g_1(\xi) = \frac{1}{2}\bar g(\xi)$, $f(x) = \omega$ и $\tau_1(u) = T^2(u)$; при этом выполняются все допущения, сделанные в начале настоящего параграфа.

2. Упруго-пластический изгиб плоской пластинки, жестко закрепленной на краю[1]. Уравнение упруго-пластического изгиба пластин имеет вид:

$$Pw = \frac{\partial^2}{\partial x^2}\left[g(H)\left(\frac{\partial^2 w}{\partial x^2} + \frac{1}{2}\,\frac{\partial^2 w}{\partial y^2}\right)\right] + \frac{\partial^2}{\partial x\,\partial y}\left[g(H)\,\frac{\partial^2 w}{\partial x\,\partial y}\right] +$$

$$+ \frac{\partial^2}{\partial y^2}\left[g(H)\left(\frac{\partial^2 w}{\partial y^2} + \frac{1}{2}\,\frac{\partial^2 w}{\partial x^2}\right)\right] = f(x,\ y). \qquad (68.16)$$

---

[1]) Эта задача поставлена Л. М. Качановым [1] и рассмотрена в работах А. Лангенбаха [1]—[4]; в последней изучены также случаи, когда край пластины свободно оперт или свободен.

Здесь $f(x, y)$ — величина, пропорциональная внешней нормальной нагрузке, рассчитанной на единицу площади, g — функция, характерная для данного материала,

$$H = H(w) = \left(\frac{\partial^2 w}{\partial x^2}\right)^2 + \left(\frac{\partial^2 w}{\partial y^2}\right)^2 + \left(\frac{\partial^2 w}{\partial x\, \partial y}\right)^2 + \frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2}. \quad (68.17)$$

Очевидно, $H(w)$ — положительно определенная квадратичная форма относительно вторых производных функции w. Прогиб w должен удовлетворять граничным условиям

$$w\Big|_S = \frac{\partial w}{\partial \nu}\Big|_S = 0, \quad (68.18)$$

где S — контур пластинки и ν — нормаль к S.

Допустим, что функция $g(\xi)$ имеет три непрерывные производные и удовлетворяет неравенствам

$$g(\xi) \geqslant c_1, \qquad g(\xi) + 2g'(\xi)\,\xi \geqslant c_2, \quad (68.19)$$

где c_1 и c_2 — положительные постоянные. Будем рассматривать P (формула (68.16)) как оператор в пространстве $L_2(\Omega)$, где Ω — область пластинки; за область $D(P)$ определения этого оператора примем множество функций, четвертые производные которых непрерывны в замкнутой области $\overline{\Omega} = \Omega \cup S$ и которые удовлетворяют краевым условиям (68.18). Оператор P имеет производную P'_w, которая представляет собой линейный симметричный оператор, удовлетворяющий неравенствам (66.2) и (66.15); последнее неравенство имеет место при $\overset{*}{w}_0 = 0$. Докажем эти утверждения.

Пусть w и h — произвольные элементы из $D(P)$, а t — вещественная переменная. Простой подсчет дает:

$$\begin{aligned} P'_w h = \lim_{t\to 0} \frac{P(w+th) - P(w)}{t} = \frac{\partial}{\partial x^2}\Big\{ g(H(w))\left(\frac{\partial^2 h}{\partial x^2} + \frac{1}{2}\frac{\partial^2 h}{\partial y^2}\right) + \\ + 2g'(H(w))\left(\frac{\partial^2 w}{\partial x^2} + \frac{1}{2}\frac{\partial^2 w}{\partial y^2}\right) H(w, h)\Big\} + \\ + \frac{\partial^2}{\partial y^2}\Big\{ g(H(w))\left(\frac{\partial^2 h}{\partial y^2} + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}\right) + \\ + 2g'(H(w))\left(\frac{\partial^2 w}{\partial y^2} + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}\right) H(w, h)\Big\} + \\ + \frac{\partial^2}{\partial x\, \partial y}\Big\{ g(H(w))\frac{\partial^2 h}{\partial x\, \partial y} + 2g'(H(w))\frac{\partial^2 w}{\partial x\, \partial y} H(w, h)\Big\}. \end{aligned} \quad (68.20)$$

Здесь

$$\begin{aligned} H(w, h) = \frac{\partial^2 w}{\partial x^2}\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 w}{\partial x\, \partial y}\frac{\partial^2 h}{\partial x\, \partial y} + \\ + \frac{1}{2}\left(\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 w}{\partial y^2}\frac{\partial^2 h}{\partial x^2}\right) \end{aligned} \quad (68.21)$$

— билинейная форма, соответствующая квадратичной форме $H(w)$. Будучи элементом области $D(P)$, функция h удовлетворяет краевым условиям

$$h\Big|_S = \frac{\partial h}{\partial \nu}\Big|_S = 0. \tag{68.22}$$

Очевидно, оператор P'_w, определяемый формулами (68.20) — (68.22), — линейный и симметричный. Составим теперь выражение

$$\left(P'_w h,\ h\right) = \iint\limits_{\Omega} hP'_w h\, dx\, dy.$$

Интегрируя дважды по частям и используя условия (68.22), получим:

$$\left(P'_w h,\ h\right) = \iint\limits_{\Omega} \{g(H(w))\,H(h) + 2g'(H(w))\,H^2(w,\ h)\}\, dx\, dy. \tag{68.23}$$

Повторив рассуждения примера 1, получим неравенство

$$\left(P'_w h,\ h\right) \geqslant c' \iint\limits_{\Omega} H(h)\, dx\, dy, \qquad c' = \text{const} > 0. \tag{68.24}$$

Условия (68.22) позволяют написать тождество [1])

$$\iint\limits_{\Omega} H(h)\, dx\, dy = \iint\limits_{\Omega} \left\{\left(\frac{\partial^2 w}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 w}{\partial x\, \partial y}\right)^2 + \left(\frac{\partial^2 w}{\partial y^2}\right)^2\right\} dx\, dy$$

и следовательно [2]),

$$\left(P'_w h,\ h\right) \geqslant c'\varkappa^2 \|w\|^2, \tag{68.25}$$

где $\varkappa$ — постоянная неравенства Фридрихса. Тем самым доказано, что P'_w удовлетворяет неравенству (66.2).

При $w = 0$ имеем $H(w) = 0$, $H(w,\ h) = 0$ и из соотношений (68.23) и (68.24) вытекает

$$\left(P'_w h,\ h\right) \geqslant \frac{c'}{g(0)} \left(P'_0 h,\ h\right), \tag{68.26}$$

что равносильно неравенству (66.15) при $w_0 = 0$.

Задача (68.16) — (68.18) равносильна задаче о минимуме функционала

$$F(w) = \int\limits_0^1 (P(tw),\ w)\, dt - (f,\ w).$$

---

[1]) Ср. ВМ, стр. 167, формула (6_2).

[2]) См. ВМ, стр. 172.

Действуя, как в примере 1, легко найдем:

$$\int_0^1 (P(tw),\, w)\,dt = \int_\Omega\int dx\,dy \int_0^{H(w)} \frac{1}{2}\,g(\xi)\,d\xi,$$

теперь

$$F(w) = \int_\Omega\int dx\,dy \int_0^{H(w)} \frac{1}{2}\,g(\xi)\,d\xi - \int_\Omega fw\,dx\,dy, \qquad (68.27)$$

что подходит под формулу (68.1).

3. В статьях К. Грёгера [1], [2] рассмотрены случаи, когда формы $\tau_i(u)$ могут вырождаться, а также случаи бесконечных областей Ω. В статье [5] А. Лангенбаха рассмотрена общая задача нелинейной теории упругих пластин в предположении, что потенциальная энергия деформации определенным образом зависит от первых двух инвариантов тензора деформации; это приводит к функционалам вида (68.1) при $s=2$.

§ 69. Функционалы теории пластичности и их обобщение (продолжение)

В настоящем параграфе мы займемся вопросами существования обобщенного решения задачи о минимуме функционала вида (68.1).

Для упрощения выкладок ограничимся случаем, когда сумма в формуле (68.1) содержит только одно слагаемое, так что функционал (68.1) имеет вид

$$F(u) = \int_\Omega dx \int_0^{\tau(u)} g(\xi)\,d\xi - (f,\, u). \qquad (69.1)$$

Вместо допущений § 68 примем приводимые ниже допущения о функционале (69.1); одни из них являются новыми, другие аналогичны приведенным в § 68.

1) $\tau(u)$ — неотрицательная квадратичная форма относительно функции u (скалярной или векторной) и ее производных до некоторого порядка l включительно. При этом $\tau(u) = \tau^{(1)}(u) + \tau^{(2)}(u)$, где $\tau^{(1)}u$ зависит только от производных порядка l, а $\tau^{(2)}(u)$ — только от самой функции u и ее производных порядка меньшего, чем l. Коэффициенты формы $\tau(u)$ могут зависеть от x.

2) Функция $g(\xi)$ удовлетворяет неравенствам

$$g(\xi) \geqslant a, \qquad (69.2)$$

$$g(\xi) + 2\xi g'(\xi) \geqslant a_1, \qquad (69.3)$$

где a и a_1 — положительные постоянные.

3) При $\xi \to \infty$ отношение $g(\xi) : \xi^{\frac{p}{2}-1}$, где p — постоянная, $2 \leqslant p < \infty$, ограничено сверху и снизу положительными числами; мы ограничиваемся значениями $p \geqslant 2$, так как при $p < 2$ было бы нарушено неравенство (69.2). Допущение 3) равносильно неравенству

$$\alpha_0 + \alpha_1 \xi^{\frac{p}{2}-1} \leqslant g(\xi) \leqslant A_0 + A_1 \xi^{\frac{p}{2}-1}, \qquad 0 \leqslant \xi < \infty, \tag{69.4}$$

где α_0, α_1, A_0, A_1 — положительные постоянные.

4) При $\xi \to \infty$

$$g'(\xi) = O\left(\xi^{\frac{p}{2}-2}\right). \tag{69.5}$$

5) При любом конечном $\xi \geqslant 0$ функция $g(\xi)$ имеет непрерывную вторую производную.

Прежде чем сформулировать остальные допущения, введем функциональное пространство, которое в последующем будет играть важную роль. Как обычно, задачу о минимуме функционала (69.1) мы будем решать в предположении, что функция u удовлетворяет некоторым краевым условиям, которые мы будем считать *линейными однородными*. Введем в рассмотрение пространство $\overline{W}_p^{(l)}(\Omega)$, которое является подпространством соболевского пространства $W_p^{(l)}(\Omega)$ и представляет собой замыкание в метрике $W_p^{(l)}(\Omega)$ множества достаточно гладких функций, удовлетворяющих краевым условиям нашей вариационной задачи.

Допустим теперь, что:

6) Норма в $W_p^{(l)}(\Omega)$ на подпространстве $\overline{W}_p^{(l)}(\Omega)$ эквивалентна величине

$$\left\{ \int_\Omega [\tau(u)]^{\frac{p}{2}} \, dx \right\}^{\frac{1}{p}},$$

это позволяет принять указанную величину за норму в $\overline{W}_p^{(l)}(\Omega)$. Обозначая для краткости $B = \overline{W}_p^{(l)}(\Omega)$, имеем, следовательно,

$$\|u\|_B = \left\{ \int_\Omega [\tau(u)]^{\frac{p}{2}} \, dx \right\}^{\frac{1}{p}}. \tag{69.6}$$

В последующем знак B у нормы будем опускать.

7) В заключение примем, что f в формуле (69.1) есть элемент пространства B^*, сопряженного с B.

В силу допущений 1) и 3) функционал

$$\Phi(u) = \int_\Omega dx \int_0^{\tau(u)} g(\xi) \, d\xi \tag{69.7}$$

определен на всем пространстве B; из допущения 7) вытекает, что функционал

$$F(u)=\Phi(u)-(f,\ u)$$

определен на том же пространстве[1]).

Ниже формулируется и доказывается ряд теорем о функционале (69.1). Мы предполагаем при этом, для упрощения формулировок, что все допущения 1) — 7) имеют место, хотя для каждой в отдельности взятой теоремы все эти допущения не всегда необходимы.

Теорема 69.1. *Функционал* (69.1) *непрерывен в пространстве* B.

Теорему достаточно доказать для функционала (69.7). Пусть $\|u_n-u\|\to 0$. Имеем:

$$\Phi(u_n)-\Phi(u)=\int\limits_{\Omega} dx\int\limits_{\tau(u)}^{\tau(u_n)} g(\xi)\,d\xi.$$

Из неравенства (69.4) следует, что

$$|\Phi(u_n)-\Phi(u)|\leqslant\int\limits_{\Omega}\left[A_0\,|\tau_n-\tau|+A_2\left|\tau_n^{\frac{p}{2}}-\tau^{\frac{p}{2}}\right|\right]dx; \qquad (69.8)$$

для краткости здесь положено $\tau=\tau(u)$, $\tau_n=\tau(u_n)$, $A_2=\frac{2}{p}A_1$.

Оценим каждое слагаемое отдельно. Имеем:

$$\tau_n-\tau=\tau(u_n)-\tau(u)=\tau(u_n-u,\ u_n+u),$$

где $\tau(u,\ v)$ — билинейная форма, соответствующая квадратичной форме $\tau(u)$. Эта последняя неотрицательна, для нее верно неравенство Коши:

$$|\tau(u_n-u,\ u_n+u)|\leqslant\sqrt{\tau(u_n-u)}\,\sqrt{\tau(u_n+u)},$$

а следовательно, и неравенство

$$\int\limits_{\Omega}|\tau_n-\tau|\,dx\leqslant\int\limits_{\Omega}[\tau(u_n-u)]^{\frac{1}{2}}\,[\tau(u_n+u)]^{\frac{1}{2}}\cdot 1\cdot dx.$$

---

[1]) В статье А. Лангенбаха [2] рассмотрен более общий случай, когда сумма (68.1) содержит любое число слагаемых, а функции $g_j(\xi)$ удовлетворяют соотношению, более общему, чем (69.3); при этом оказывается необходимым ввести в рассмотрение пространства Орлича.

К интегралу справа применим неравенство Гёльдера с показателями p, p, r, где r таково, что $\frac{2}{p}+\frac{1}{r}=1$. Мы получим тогда:

$$\int\limits_\Omega |\tau_n-\tau|\,dx \leqslant$$
$$\leqslant\left\{\int\limits_\Omega [\tau_n(u_n-u)]^{\frac{p}{2}}\,dx\right\}^{\frac{1}{p}}\left\{\int\limits_\Omega [\tau(u_n+u)]^{\frac{p}{2}}\,dx\right\}^{\frac{1}{p}}\left\{\int\limits_\Omega dx\right\}^{\frac{1}{r}}=$$
$$=(\mu\Omega)^{\frac{1}{r}}\|u_n-u\|\,\|u_n+u\|,$$

через $\mu\Omega$ обозначена мера области Ω. Так как $\|u_n-u\|\to 0$, то норма $\|u_n+u\|$ ограничена, и

$$\lim_{u_n\to u}\int\limits_\Omega |\tau_n-\tau|\,dx=0. \tag{69.9}$$

Оценим теперь второй интеграл в (69.8). По формуле Лагранжа имеем:

$$\frac{\tau_n^{\frac{p}{2}}-\tau^{\frac{p}{2}}}{\tau_n-\tau}=\frac{p}{2}\,\sigma_n^{\frac{p}{2}-1},\qquad \sigma_n=(1-\theta_n)\tau+\theta_n\tau_n,\qquad 0<\theta_n<1.$$

При любых неотрицательных числах a, b, k справедливо неравенство

$$(a+b)^k\leqslant 2^k\max(a^k,\ b^k)\leqslant 2^k(a^k+b^k). \tag{69.10}$$

Отсюда

$$\sigma_n^{\frac{p}{2}-1}\leqslant 2^{\frac{p}{2}-1}\left(\tau_n^{\frac{p}{2}-1}+\tau^{\frac{p}{2}-1}\right)$$

и

$$\int\limits_\Omega \left|\tau_n^{\frac{p}{2}}-\tau^{\frac{p}{2}}\right|dx\leqslant 2^{\frac{p}{2}-2}p\int\limits_\Omega |\tau_n-\tau|\left(\tau_n^{\frac{p}{2}-1}+\tau^{\frac{p}{2}-1}\right)dx\leqslant$$
$$\leqslant 2^{\frac{p}{2}-2}p\int\limits_\Omega [\tau(u_n-u)]^{\frac{1}{2}}[\tau(u_n+u)]^{\frac{1}{2}}\left(\tau_n^{\frac{p}{2}-1}+\tau^{\frac{p}{2}-1}\right)dx.$$

К интегралу справа применим неравенство Гёльдера с прежними показателями p, p, r:

$$\int\limits_\Omega \left|\tau_n^{\frac{p}{2}}-\tau^{\frac{p}{2}}\right|dx\leqslant$$
$$\leqslant 2^{\frac{p}{2}-2}p\|u_n-u\|\,\|u_n+u\|\left\{\int\limits_\Omega\left(\tau_n^{\frac{p-2}{2}}+\tau^{\frac{p-2}{2}}\right)^r dx\right\}^{\frac{1}{r}}.$$

Замечая, что $r=\frac{p}{p-2}$, имеем по неравенству (69.10):

$$\int_\Omega \left(\tau_n^{\frac{p-2}{2}}+\tau^{\frac{p-2}{2}}\right)^r dx \leqslant 2^r \int_\Omega \left(\tau_n^{\frac{p}{2}}+\tau^{\frac{p}{2}}\right) dx = 2^r (\|u_n\|^p+\|u\|^p).$$

Приняв во внимание, что величина $\|u_n\|$ ограничена, находим окончательно:

$$\lim_{u_n\to u}\int_\Omega \left|\tau_n^{\frac{p}{2}}-\tau^{\frac{p}{2}}\right| dx = 0. \tag{69.11}$$

Теперь из формул (69.8), (69.9) и (69.11) следует:

$$\lim_{u_n\to u} [\Phi(u_n)-\Phi(u)] = 0,$$

что и требовалось доказать.

Теорема 69.2. *Функционал* (69.1) — *существенно выпуклый в пространстве B.*

Достаточно доказать, что существенно выпуклым является функционал (69.7). Доказательство же этого последнего утверждения разобьем на несколько этапов.

1°. Функционал Φ имеет градиент $P=\operatorname{grad}\Phi$, определенный на всем пространстве B. Действительно,

$$\frac{d}{dt}\Phi(u+th)\big|_{t=0} = \frac{d}{dt}\int_\Omega dx \int_0^{\tau(u+th)} g(\xi)\,d\xi\,\big|_{t=0} =$$

$$= 2\int_\Omega g(\tau(u))\,\tau(u,\,h)\,dx. \tag{69.12}$$

Докажем, что правая часть равенства (69.12) есть функционал над h, ограниченный в B при любом $u\in B$. Имеем:

$$\left|\int_\Omega g(\tau(u))\,\tau(u,\,h)\,dx\right| \leqslant \int_\Omega g(\tau(u))\,[\tau(u)]^{\frac{1}{2}}\,[\tau(h)]^{\frac{1}{2}}\,dx.$$

К интегралу справа применим неравенство Гёльдера с показателями $r=\frac{p}{p-2}$, p, p:

$$\left|\int_\Omega g(\tau(u))\,\tau(u,\,h)\,dx\right| \leqslant \left\{\int_\Omega [g(\tau(u))]^r\,dx\right\}^{\frac{1}{r}} \|u\|\cdot\|h\|.$$

По неравенству (69.4)

$$\int_\Omega [g(\tau(u))]^r\,dx \leqslant \int_\Omega \left\{A_0 + A_1[\tau(u)]^{\frac{p-2}{2}}\right\}^r dx \leqslant$$
$$\leqslant 2^r \int_\Omega \left\{A_0^r + A_1^r[\tau(u)]^{\frac{p}{2}}\right\} dx = 2^r\left\{A_0^r\mu\Omega + A_1^r\|u\|^p\right\}, \quad (69.13)$$

и наше утверждение доказано.

2°. Производная P'_u существует при любом $u \in B$; ее область определения $D(P'_u) = B$. Для доказательства напишем равенство, определяющее оператор P и вытекающее из тождества (69.12):

$$(P(u), h) = 2\int_\Omega g(\tau(u))\,\tau(u, h)\,dx$$

Отсюда

$$\frac{d}{dt}(P(u + th_1), h)|_{t=0} =$$
$$= 2\int_\Omega \{g(\tau(u))\,\tau(h_1, h) + 2g'(\tau(u))\,\tau(u, h_1)\,\tau(u, h)\}\,dx. \quad (69.14)$$

Докажем, что правая часть равенства (69.14) есть билинейный оператор над h и h_1, определенный на всем пространстве B и ограниченный при любом $u \in B$. Имеем:

$$\left|\int_\Omega \{g(\tau(u))\,\tau(h_1, h) + 2g'(\tau(u))\,\tau(u, h_1)\,\tau(u, h)\}\,dx\right| \leqslant$$
$$\leqslant \int_\Omega g(\tau(u))\,[\tau(h_1)]^{\frac{1}{2}}\,[\tau(h)]^{\frac{1}{2}}\,dx +$$
$$+ 2\int_\Omega |g'(\tau(u))|\,\tau(u)\,[\tau(h_1)]^{\frac{1}{2}}\,[\tau(h)]^{\frac{1}{2}}\,dx. \quad (69.15)$$

Применив к первому члену неравенство Гёльдера с показателями $r = \frac{p}{p-2}$, p, p, найдем, что

$$\int_\Omega g(\tau(u))\,[\tau(h_1)]^{\frac{1}{2}}\,[\tau(h)]^{\frac{1}{2}}\,dx \leqslant C_1\|h_1\|\,\|h\|, \quad (69.16)$$

где, в силу формулы (69.13),

$$C_1 = 2^r\left\{A_0^r\mu\Omega + A_1^r\|u\|^p\right\}.$$

Далее, из соотношения (69.5) следует, что

$$|g'(\tau(u))|\,\tau(u) \leqslant b_0\tau(u) + b_1[\tau(u)]^{\frac{p}{2}-1}; \quad b_0,\ b_1 = \mathrm{const} > 0.$$

Применив ко второму интегралу в (69.15) неравенство Гёльдера с теми же показателями, мы легко получим оценку

$$\int_\Omega |g'(\tau(u))|\,\tau(u)\,[\tau(h_1)]^{\frac{1}{2}}\,[\tau(h)]^{\frac{1}{2}}\,dx \leqslant C_2\,\|h_1\|\,\|h\|, \tag{69.17}$$

где C_2 — постоянная. Из неравенств (69.16) и (69.17) вытекает ограниченность билинейного функционала (69.14); в свою очередь отсюда следует, что производная P'_u существует при любом $u \in B$ и ее область определения совпадает с B при любом u.

3°. Производная P'_u положительна; при $p=2$ она равномерно положительно определенная. Для доказательства напишем равносильное формуле (69.14) соотношение

$$\left(P'_u h_1,\ h\right) = 2\int_\Omega \{g(\tau(u))\,\tau(h_1,\ h) + 2g'(\tau(u))\,\tau(u,\ h_1)\,\tau(u,\ h)\}\,dx.$$

Из этого соотношения видно, что P'_u — симметричный оператор. Положив $h_1 = h$, получим:

$$\left(P'_u h,\ h\right) = 2\int_\Omega \{g(\tau(u))\,\tau(h) + 2g'(\tau(u))\,\tau^2(u,\ h)\}\,dx.$$

Повторив рассуждения примера 1 § 68 и применив неравенство (69.3), найдем, что

$$g(\tau(u))\,\tau(h) + g'(\tau(u))\,\tau^2(u,\ h) \geqslant c\tau(h), \qquad c = \mathrm{const} > 0.$$

Отсюда следует, что при любом $p \geqslant 2$ оператор P'_u положителен; если же $p=2$, то

$$\left(P'_u h,\ h\right) \geqslant c\int_\Omega \tau(u)\,dx = c\,\|u\|^2,$$

и наш оператор равномерно положительно определенный.

4°. Существенная выпуклость функционала (69.7) в пространстве B теперь вытекает из первого утверждения теоремы 66.1, так как в нашем случае $D(P) = B$.

Теорема 69.3. *Функционал* (69.1) — *возрастающий в пространстве* B.

По неравенству (69.4)

$$\Phi(u) \geqslant \frac{2\alpha_1}{p}\int_\Omega [\tau(u)]^p\,dx = \frac{2\alpha_1}{p}\,\|u\|^p.$$

Отсюда

$$F(u) \geqslant \frac{2\alpha_1}{p}\,\|u\|^p - \|f\|\,\|u\| \xrightarrow[\|u\|\to\infty]{} \infty.$$

С другой стороны, по тому же неравенству (69.4)

$$F(u) \leqslant A_0 \int_\Omega \tau(u)\, dx + \frac{2A_1}{p} \| u \|^p + \| f \| \| u \|.$$

По неравенству Гёльдера

$$\int_\Omega \tau(u)\, dx \leqslant \left\{ \int_\Omega [\tau(u)]^{\frac{p}{2}}\, dx \right\}^{\frac{2}{p}} (\mu\Omega)^{\frac{p-2}{p}} = (\mu\Omega)^{\frac{p-2}{p}} \| u \|^2,$$

и ясно, что функционал $F(u)$ ограничен, если ограничена норма функции u.

Теорема 69.4. *Функционал* (69.1) *слабо полунепрерывен снизу в пространстве* B.

Достаточно доказать это утверждение для функционала (69.7); мы покажем, что этот функционал подходит под теорему 63.1. В данном случае мы примем за $U = (u_1, u_2, \ldots, u_s)$ совокупность из самой функции u и ее производных до порядка $l-1$ включительно, а за $P = (p_1, p_2, \ldots, p_k)$ — совокупность производных порядка l от функции u. В нашем случае

$$G(x, U, P) = \int_0^{\tau(u)} g(\xi)\, d\xi.$$

Выполнение условий 1) и 2) теоремы 63.1 очевидно. Проверим условие 3): как было указано в § 63, для этого достаточно проверить, что матрица вторых производных

$$\frac{\partial^2 G}{\partial p_\alpha \partial p_\beta} \qquad (\alpha, \beta = 1, 2, \ldots, k)$$

или, что то же, квадратичная форма

$$\sum_{\alpha, \beta = 1}^{k} \frac{\partial^2 G}{\partial p_\alpha \partial p_\beta} t_\alpha t_\beta \tag{69.18}$$

от переменных $t_1, t_2, \ldots, t_k$ — неотрицательная. Напомним, что $\tau(u) = \tau^{(1)}(u) + \tau^{(2)}(u)$, где от переменных $p_1, p_2, \ldots, p_k$ зависит только величина $\tau^{(1)}(u)$, которая представляет собой неотрицательную форму упомянутых переменных.

Выражение $\tau^{(1)}(u)$ есть неотрицательная квадратичная форма независимых переменных $p_1, p_2, \ldots, p_k$; пусть

$$\tau^{(1)}(u) = \sum_{\alpha, \beta = 1}^{k} a_{\alpha\beta}\, p_\alpha p_\beta.$$

Заметим, что коэффициенты $a_{\alpha\beta}$ могут зависеть от x. Имеем теперь:

$$\frac{\partial G}{\partial p_i} = 2g(\tau(u)) \sum_{\beta=1}^{k} a_{i\beta} p_\beta,$$

$$\frac{\partial^2 G}{\partial p_i \, \partial p_j} = 2a_{ij} g(\tau(u)) + 4g'(\tau(u)) \sum_{\beta=1}^{k} a_{i\beta} p_\beta \sum_{\gamma=1}^{k} a_{j\gamma} p_\gamma.$$

Отсюда

$$\sum_{i, j=1}^{k} \frac{\partial^2 G}{\partial p_i \, \partial p_j} t_i t_j = 2\tau^{(1)}(t) g(\tau(u)) + 4g'(\tau(u)) \tau^{(1)\,2}(u, t); \qquad (69.19)$$

здесь положено для краткости

$$\tau^{(1)}(t) = \sum_{\alpha, \beta=1}^{k} a_{\alpha\beta} t_\alpha t_\beta, \qquad \tau^{(1)}(u, t) = \sum_{\alpha, \beta=1}^{k} a_{\alpha\beta} t_\alpha u_\beta.$$

Если $g'(\tau(u)) \geqslant 0$, то, как показывает формула (69.19), выражение (69.18) неотрицательно. Пусть теперь $g'(\tau(u)) < 0$. Форма $\tau^{(1)}(u)$ неотрицательна, и по неравенству Коши

$$\tau^{(1)\,2}(u, t) \leqslant \tau^{(1)}(u) \tau^{(1)}(t).$$

Теперь из формулы (69.19) следует

$$\sum_{i, j=1}^{k} \frac{\partial^2 G}{\partial p_i \, \partial p_j} t_i t_j \geqslant 2\tau^{(1)}(u) [g(\tau(u)) + 2\tau^{(1)}(u) g'(\tau(u))];$$

так как форма $\tau^{(2)}(u)$ неотрицательна, то $\tau^{(1)}(u) \leqslant \tau(u)$ и

$$\sum_{i, j=1}^{k} \frac{\partial^2 G}{\partial p_i \, \partial p_j} t_i t_j \geqslant 2\tau^{(1)}(u) [g(\tau(u)) + 2\tau(u) g'(\tau(u))].$$

Используя рассуждения примера 1 § 65 и неравенство (69.3), мы легко найдем теперь, что форма (69.18) неотрицательна при любых значениях $p_1, p_2, \ldots, p_k$.

Пусть теперь u_n слабо сходится к u в пространстве B. Отсюда следует, что в пространстве $L_p(\Omega)$ производные порядка l от функций u_n слабо сходятся к аналогичным производным от функции u; из теорем вложения С. Л. Соболева следует тогда, что в том же пространстве $L_p(\Omega)$ производные порядка $< l$ от функций u_n сильно сходятся к аналогичным производным от функции u. Все условия теоремы 63.1 оказываются выполненными, а тогда

$$\underline{\lim}\, \Phi(u_n) \geqslant \Phi(u_0)$$

и функционал $\Phi(u)$ слабо полунепрерывен снизу.

Из теорем 69.1—69.4 и 63.1, а также из соболевских теорем вложения вытекает

Теорема 69.5. *Пусть имеют место допущения* 1)—7). *Тогда функционал* (69.1) *в пространстве B имеет абсолютный минимум, который достигается в единственной точке. Если u_0 — эта точка и $\{u_n\}$ — любая минимизирующая последовательность, то производные порядка l от u_n стремятся к аналогичным производным от u_0 слабо в $L_p(\Omega)$, а производные меньших порядков от u_n в том же пространстве $L_p(\Omega)$ сильно сходятся к аналогичным производным от u_0.*

В случае, когда $p=2$, теорему 69.5 можно усилить следующим образом.

Теорема 69.6. *Пусть u_0 — точка абсолютного минимума функционала* (69.1), *$\{u_n\}$ — минимизирующая последовательность для этого функционала. Если $p=2$, то производные порядка l от функций u_n сильно сходятся в метрике $L_2(\Omega)$ к аналогичным производным от функции u_0.*

Действительно, в ходе доказательства теоремы 69.2 (п. 3°) было установлено, что при $p=2$ производная P'_u — равномерно положительно определенная в пространстве B, а тогда из теоремы 66.3 вытекает, что

$$\| u_n - u_0 \| = \| u_n - u_0 \|_{W_2^{(l)}(\Omega)} \xrightarrow[n\to\infty]{} 0.$$

Утверждение теоремы 69.5 является прямым следствием последнего соотношения.

Обратимся к примерам § 68. В задаче о кручении упрочняющих стержней

$$\tau(u) = T^2(u) = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2, \quad g(\xi) = \frac{1}{2}\bar{g}(\xi).$$

Если мы допустим, что при $\xi \to \infty$ отношение $\bar{g}(\xi) : \xi^{\frac{p}{2}-1}$ положительно ограничено сверху и снизу, а отношение $|\bar{g}'(\xi)| : \xi^{\frac{p}{2}-2}$ ограничено сверху, причем $p \geqslant 2$, то все допущения настоящего параграфа будут выполнены при $l=1$. Задача о минимуме функционала (68.15) имеет решение в $W_p^{(1)}(\Omega)$. Если u_0 — это решение, а $\{u_n\}$ — минимизирующая последовательность, то в метрике $L_p(\Omega)$ $\operatorname{grad} u_n \to \operatorname{grad} u_0$ слабо, а $u_n \to u_0$ сильно. Если $p=2$, то $\operatorname{grad} u_n \to \operatorname{grad} u_0$ сильно в $L_2(\Omega)$.

В задаче об упруго-пластическом изгибе пластинки, жестко закрепленной по краю, $l=2$,

$$\tau(w) = H(w) = \left(\frac{\partial^2 w}{\partial x^2}\right)^2 + \left(\frac{\partial^2 w}{\partial y^2}\right)^2 + \left(\frac{\partial^2 w}{\partial x\,\partial y}\right)^2 + \frac{\partial^2 w}{\partial x^2}\,\frac{\partial^2 w}{\partial y^2}.$$

Если функция $g(\xi)$ удовлетворяет условиям, которые только что были поставлены для функции $\bar{g}(\xi)$, то наша задача имеет решение $w_0 \in W_2^{(2)}(\Omega)$. Если $\{w_n\}$ — минимизирующая последовательность, то

$$\frac{\partial^2 w_n}{\partial x^2} \to \frac{\partial^2 w_0}{\partial x^2}, \qquad \frac{\partial^2 w_n}{\partial x\, \partial y} \to \frac{\partial^2 w_0}{\partial x\, \partial y}, \qquad \frac{\partial^2 w_n}{\partial y^2} \to \frac{\partial^2 w_0}{\partial y^2} \tag{69.20}$$

слабо в $L_p(\Omega)$, а

$$\frac{\partial w_n}{\partial x} \to \frac{\partial w_0}{\partial x}, \qquad \frac{\partial w_n}{\partial y} \to \frac{\partial w_0}{\partial y}, \qquad w_n \to w_0,$$

сильно в $L_p(\Omega)$. Если $p=2$, то предельные соотношения (69.20) выполняются сильно в $L_2(\Omega)$.

ГЛАВА X

ЧИСЛЕННОЕ РЕШЕНИЕ НЕЛИНЕЙНЫХ ВАРИАЦИОННЫХ ЗАДАЧ

§ 70. Процессы Ритца и Бубнова — Галёркина

Как показывают результаты предшествующей главы, решение вариационной задачи в широком классе случаев можно получить, если построить последовательность, минимизирующую для данного функционала. Если минимизирующая последовательность сходится, то ее любой достаточно далекий член можно принять за приближенное решение данной вариационной задачи; тем самым вопрос о численном ее решении можно свести к тому же вопросу о построении минимизирующей последовательности. Как и для линейных задач, такое построение во многих случаях можно выполнять с помощью процесса Ритца.

Поставим задачу о минимуме функционала F с линейной областью определения $D(F)$. Примем, что множество $D(F)$ плотно в некотором линейном сепарабельном метрическом (не обязательно нормированном) пространстве M с метрической функцией $\rho(u, v)$. Зададим последовательность элементов («координатную систему»)

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, \tag{70.1}$$

подчиненных обычным условиям: 1) $\varphi_n \in D(F)$; 2) элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ линейно независимы при любом n; 3) система (70.1) полна в M — последнее, как всегда, означает, что множество всевозможных конечных линейных комбинаций элементов системы (70.1) плотно в M. Как это обычно для процесса Ритца, задаем число n и заменяем задачу о минимуме функционала F на множестве $D(F)$ задачей о минимуме того же функционала на n-мерном подпространстве, натянутом на элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$. Пусть у новой задачи существует решение; оно необходимо имеет вид

$$u_n = \sum_{k=1}^{n} a_k \varphi_k; \tag{70.2}$$

коэффициенты a_k на самом деле зависят еще и от n, но мы этого отмечать не будем. Элемент u_n называется приближенным по Ритцу решением вариационной задачи для функционала F.

Допустим, что функционал F непрерывно дифференцируем на любой конечномерной гиперплоскости в области своего определения. Тогда коэффициенты Ритца a_k необходимо удовлетворяют следующей системе уравнений:

$$\frac{\partial F\left(\sum_{k=1}^{n} a_k\varphi_k\right)}{\partial a_j} = 0 \qquad (j = 1,\ 2,\ \ldots,\ n). \tag{70.3}$$

Теорема 70.1. *Пусть выполнены сформулированные в настоящем параграфе условия относительно функционала F и координатной системы* (70.1). *Если функционал F в метрике ρ — возрастающий и полунепрерывный сверху, то приближенные решения по Ритцу можно построить при любом n, и для функционала F эта последовательность — минимизирующая*[1]).

Доказательство. Выражение $F\left(\sum_{k=1}^{n} a_k\varphi_k\right)$ есть функция переменных $a_1,\ a_2,\ \ldots,\ a_n$, непрерывно дифференцируемая при всех значениях этих переменных и стремящаяся к $+\infty$, если $a_1^2 + a_2^2 + \ldots$ $\ldots + a_n^2 \to \infty$. Эта функция достигает своего абсолютного минимума по крайней мере в одной точке, которая находится на конечном расстоянии от начала, и в этой точке необходимо

$$\frac{\partial F\left(\sum_{k=1}^{n} a_k\varphi_k\right)}{\partial a_j} = 0 \qquad (j = 1,\ 2,\ \ldots,\ n).$$

Таким образом, при любом натуральном n система Ритца (70.3) имеет по крайней мере одно решение и, следовательно, при любом натуральном n существует хотя бы одно приближение по Ритцу решения для задачи о минимуме функционала F.

Докажем теперь, что эти приближенные решения образуют минимизирующую последовательность для функционала F. Обозначим $\inf F(u) = d$. Построим минимизирующую последовательность $\{u^{(n)}\}$ такую, что

$$F(u^{(n)}) \leqslant d + \frac{1}{n},$$

[1]) См. статью автора [19]. В формулировке теоремы в этой статье допущена описка: вместо «полунепрерывный сверху» написано «полунепрерывный снизу».

В силу условия полноты координатной системы, для каждого $u^{(n)}$ можно подобрать такую линейную комбинацию

$$v^{(N_n)} = \sum_{k=1}^{N_n} \alpha_k^{(n)} \varphi_k,$$

что $\rho\left(u^{(n)},\ v^{(N_n)}\right) < \delta_n$; число δ_n выберем столь малым, чтобы для любого v, удовлетворяющего неравенству $\rho(u^{(n)}, v) < \delta_n$, было $F(u^{(n)}) - F(v) \geqslant -\frac{1}{n}$. Это возможно, потому что функционал F полунепрерывен сверху. Теперь

$$F\left(v^{(N_n)}\right) \leqslant F\left(u^{(n)}\right) + \frac{1}{n}$$

и, тем более,

$$F\left(v^{(N_n)}\right) \leqslant d + \frac{2}{n}.$$

Отсюда следует, что $\left\{v^{(N_n)}\right\}$ есть минимизирующая последовательность.

Пусть

$$u_p = \sum_{k=1}^{p} a_k \varphi_k$$

означает p-е приближенное решение по Ритцу. Тогда

$$F\left(u_{N_n}\right) \leqslant F\left(v^{(N_n)}\right) \leqslant d + \frac{2}{n}$$

и $F\left(u_{N_n}\right)$ также есть минимизирующая последовательность. Наконец, так как $F(u_n)$ монотонно убывает с возрастанием n, а последовательность $F\left(u_{N_n}\right) \to d$, то и $F(u_n) \to d$, что и требовалось доказать.

В задачах § 68 функционал F не только полунепрерывен сверху, но и просто непрерывен в соответствующем пространстве B; в обеих задачах указанный функционал также и возрастающий. Поэтому в каждой из этих задач приближенное решение по Ритцу можно построить при любом n; о сходимости этих приближенных решений можно повторить все сказанное в конце прошлого параграфа о сходимости минимизирующей последовательности.

Элемент, реализующий минимум функционала F, удовлетворяет уравнению

$$Pu = 0, \qquad P = \operatorname{grad} F. \tag{70.4}$$

Допустим, что область $D(F)$ определения функционала F принадлежит некоторому сепарабельному банахову пространству B и плотна в нем. Далее, примем, что $\operatorname{grad} F$ имеет линейную область определения, также плотную в B. Тогда уравнение (70.4) можно решать процессом Бубнова — Галёркина. Зададим координатную систему

(70.1), удовлетворяющую поставленным выше условиям 2) и 3); условие 1) заменим более жестким: 1′) $\varphi_n \in D(P)$. Возьмем элемент вида (70.2) и определим коэффициенты a_k из системы уравнений Бубнова — Галёркина

$$\left(P\left(\sum_{k=1}^{n} a_k\varphi_k\right),\ \varphi_j\right)=0 \qquad (j=1,\ 2,\ \ldots,\ n). \tag{70.5}$$

Если система (70.5) имеет решение, то, подставив его в формулу (70.2), получим элемент, который называется приближенным решением по Бубнову — Галёркину уравнения (70.5).

Заметим сразу же, что процесс Бубнова — Галёркина можно применять и тогда, когда оператор P не потенциальный: достаточно, чтобы $D(P)$ было линейным множеством, плотным в B, чтобы значения оператора P лежали в сопряженном пространстве B^* и чтобы координатная система удовлетворяла условиям 1), 2), 3). Более того, нет необходимости, чтобы оператор P действовал из B в B^*; достаточно, чтобы этот оператор действовал из B в любое банахово пространство B_1. В этом случае уравнения Бубнова — Галёркина принимают другой вид; подробнее об этом см. М. А. Красносельский [1]. В случае, когда $P=I+T$, где I — тождественный, а T — вполне непрерывный нелинейный оператор в некотором банаховом пространстве, сходимость процесса Бубнова — Галёркина исследована М. А. Красносельским [1].

Теорема 70.2 [1]. *Пусть область $D(F)$ определения функционала F и область $D(P)$ определения оператора $P=\operatorname{grad} F$ линейны и плотны в сепарабельном банаховом пространстве B и пусть указанный функционал непрерывно дифференцируем в любой конечномерной гиперплоскости из области своего определения. Если при этом координатная система* (70.1) *удовлетворяет условиям* 1′), 2) *и* 3) *настоящего параграфа, то система* (70.3) *Ритца и система* (70.4) *Бубнова — Галёркина равносильны.*

Доказательство очень просто. Уравнения (70.3) можно записать в виде

$$\frac{\partial F(u_n)}{\partial a_j}=0,$$

где u_n определяется формулой (70.2). Но

$$\frac{\partial F(u_n)}{\partial a_j}=\frac{d}{dt}F(u_n+t\varphi_j)\big|_{t=0}=(P(u_n),\ \varphi_j),$$

и системы (70.3) и (70.4) оказываются тождественными.

---

[1]) См. статью Л. Н. Довбыш и автора [1].

Важнейший момент в реализации вариационного метода для нелинейных задач — это выбор способа численного решения системы Ритца.

В последующих параграфах этой главы мы рассмотрим три таких способа: метод Ньютона — Канторовича, метод дифференцирования по параметру и метод Л. М. Качанова.

§ 71. Применение метода Ньютона — Канторовича

Техника применения метода Ньютона — Канторовича, условия сходимости этого метода и оценка быстроты сходимости достаточно подробно изложены в книге Л. В. Канторовича и Г. П. Акилова [1]; здесь мы укажем лишь, как проводятся вычисления по этому методу для системы (70.3), и приведем численный пример.

Обозначим

$$F\left(\sum_{k=1}^{n} a_k\varphi_k\right)=\tilde{F}(a_1, a_2, \ldots, a_n).$$

По методу Ньютона — Канторовича мы исходим из некоторого начального приближения $a_{10}, a_{20}, \ldots, a_{n0}$ и строим следующее приближение $a_{11}, a_{21}, \ldots, a_{n1}$ как решение линейной системы:

$$\sum_{k=1}^{n} \frac{\partial^2\tilde{F}}{\partial a_j\,\partial a_k}\bigg|_0 (a_{k1}-a_{k0})=-\frac{\partial\tilde{F}}{\partial a_j}\bigg|_0 \qquad (j=1, 2, \ldots, n);$$

символ $|_0$ означает, что аргументы $a_1, a_2, \ldots, a_n$ следует заменить на $a_{10}, a_{20}, \ldots, a_{n0}$. Вообще, если построено s-е приближение $a_{1s}, a_{2s}, \ldots, a_{ns}$, то $(s+1)$-е приближение $a_{1,\,s+1}, a_{2,\,s+1}, \ldots, a_{n,\,s+1}$ строится как решение линейной системы

$$\sum_{k=1}^{n} \frac{\partial^2\tilde{F}}{\partial a_j\,\partial a_k}\bigg|_s (a_{k,\,s+1}-a_{ks})=-\frac{\partial\tilde{F}}{\partial a_j}\bigg|_s \qquad (j=1, 2, \ldots, n);$$

аргументы в выражениях, написанных слева от символа $|_s$, следует заменить на $a_{1s}, a_{2s}, \ldots, a_{ns}$.

Применение метода Ньютона — Канторовича к системам Ритца наталкивается на ту трудность, что далеко не всегда можно дать простой способ надежного выбора начального приближения, между тем, от этого выбора зависит сходимость процесса. Об одном способе выбора будет сказано ниже, в § 72. В примере, который мы сейчас приведем, случайно выбранное простое начальное приближение оказалось удачным.

Пример. Найдем функцию, удовлетворяющую краевым условиям

$$y(0)=1, \qquad y(1)=0$$

и реализующую минимум интеграла

$$F(y)=\int_0^1\left(\frac{1}{48}\,y'^4+y'^2+y^6-\,6y\right)dx.$$

Если положить $y=z+1-x$, то $z(0)=z(1)=0$, и z будет пробегать линейное множество. В качестве координатных можно взять функции

$$\sin k\pi x \qquad (k=1,\ 2,\ 3,\ \ldots).$$

В данном случае известно точное решение задачи $y=1-x^2$. Если функцию $z=y-(1-x)=x(1-x)$ разложить в ряд Фурье по синусам, то члены, содержащие $\sin 2k\pi x$ $(k=1,\ 2,\ \ldots)$, в этом ряде будут отсутствовать; пробные вычисления также показали, что члены с четными k входят в приближенное решение с очень малыми коэффициентами, поэтому в окончательном расчете сохранены только координатные функции

$$\sin(2k-1)\pi x \qquad (k=1,\ 2,\ 3,\ \ldots).$$

Были взяты координатные функции, соответствующие значениям $k=1,\ 2,\ 3,\ 4,\ 5$, так что приближенное решение имеет вид

$$y_5=1-x+\sum_{k=1}^{5}a_k\sin(2k-1)\pi x.$$

Подставив это в функционал, получим:

$$\begin{aligned}
&F(y_5)=F(a_1,\ a_2,\ a_3,\ a_4,\ a_5)=\\
&=-1{,}6458333-3{,}1830989a_1-1{,}0610330a_2-0{,}63661978a_3-\\
&-0{,}45472841a_4-0{,}35367765a_5+6{,}0516525a_1^2+50{,}464872a_2^2+\\
&+139{,}29131a_3^2+272{,}53097a_4^2+450{,}18385a_5^2+0{,}76100853a_1^4+\\
&+61{,}641691a_2^4+475{,}63033a_3^4+1827{,}1815a_4^4+\\
&+4992{,}9770a_5^4+27{,}396307a_1^2a_2^2+76{,}100853a_1^2a_3^2+\\
&+149{,}15767a_1^2a_4^2+246{,}56676a_1^2a_5^2+684{,}90767a_2^2a_3^2+\\
&+1342{,}4190a_2^2a_4^2+2219{,}1009a_2^2a_5^2+3728{,}9418a_3^2a_4^2+\\
&+6164{,}1691a_3^2a_5^2+12081{,}771a_4^2a_5^2+45{,}660512a_1^2a_2a_3+\\
&+106{,}54119a_1^2a_3a_4+191{,}77415a_1^2a_4a_5+136{,}98153a_1a_2^2a_3+\\
&+191{,}77415a_1a_2^2a_4+1598{,}1179a_2a_3^2a_4+684{,}90767a_1a_3^2a_5+\\
&+6712{,}0952a_3a_4^2a_5+3{,}0440341a_1^3a_2+246{,}56676a_2^3a_5+\\
&+319{,}62358a_1a_2a_3a_4+410{,}94460a_1a_2a_3a_5+\\
&\qquad\qquad\qquad+575{,}32244a_1a_2a_4a_5+2876{,}6122a_2a_3a_4a_5.
\end{aligned}$$

Продифференцировав это по переменным $a_j\,(j=1,\ 2,\ 3,\ 4,\ 5)$ и приравняв нулю полученные результаты, мы придем к следующей системе Ритца:

$$\frac{\partial F(y_5)}{\partial a_1} = -3,1830989 + 12,103305a_1 + 3,0440341a_1^3 + \\ + 54,792614a_1a_2^2 + 152,20171a_1a_3^2 + 298,31534a_1a_4^2 + \\ + 493,13352a_1a_5^2 + 91,321024a_1a_2a_3 + 213,08238a_1a_3a_4 + \\ + 383,54830a_1a_4a_5 + 136,98153a_2^2a_3 + 191,77415a_2^2a_4 + \\ + 684,90767a_3^2a_5 + 9,1321023a_1^2a_2 + 319,62358a_2a_3a_4 + \\ + 410,94460a_2a_3a_5 + 575,32244a_2a_4a_5 = 0;$$

$$\frac{\partial F(y_5)}{\partial a_2} = -1,0610330 + 100,92974a_2 + 246,56676a_2^3 + \\ + 1369,8153a_2a_3^2 + 2684,8380a_2a_4^2 + 4438,2018a_2a_5^2 + \\ + 45,660512a_1^2a_3 + 273,96306a_1a_2a_3 + 383,5483a_1a_2a_4 + \\ + 1598,1179a_3^2a_4 + 739,70028a_2^2a_5 + 319,62358a_1a_3a_4 + \\ + 410,94460a_1a_3a_5 + 575,32244a_1a_4a_5 + 2876,6122a_3a_4a_5 + \\ + 54,792614a_1^2a_2 + 3,0440341a_1^3 = 0;$$

$$\frac{\partial F(y_5)}{\partial a_3} = -0,63661978 + 278,58262a_3 + 1902,5213a_3^3 + \\ + 7457,8836a_3a_4^2 + 12328,338a_3a_5^2 + 152,20171a_1^2a_3 + \\ + 1369,8153a_2^2a_3 + 45,660512a_1^2a_2 + 106,54119a_1^2a_4 + \\ + 136,98153a_1a_2^2 + 3196,2358a_2a_3a_4 + 1369,8153a_1a_3a_5 + \\ + 6712,0952a_4^2a_5 + 319,62358a_1a_2a_4 + 410,94460a_1a_2a_5 + \\ + 2876,6122a_2a_4a_5 = 0;$$

$$\frac{\partial F(y_5)}{\partial a_4} = -0,45472841 + 545,06194a_4 + 7308,7260a_4^3 + \\ + 298,31534a_1^2a_4 + 2684,8380a_2^2a_4 + 7457,8836a_3^2a_4 + \\ + 24163,542a_4a_5^2 + 106,54119a_1^2a_3 + 191,77415a_1^2a_5 + \\ + 191,77415a_1a_2^2 + 1598,1179a_2a_3^2 + 13424,190a_3a_4a_5 + \\ + 319,62358a_1a_2a_3 + 575,32244a_1a_2a_5 + 2876,6122a_2a_3a_5 = 0;$$

$$\frac{\partial F(y_5)}{\partial a_5} = -0,35367765 + 900,36770a_5 + 19971,908a_5^3 + \\ + 493,13352a_1^2a_5 + 4438,2018a_2^2a_5 + 12328,338a_3^2a_5 + \\ + 24163,542a_4^2a_5 + 191,77415a_1^2a_4 + 684,90767a_1a_3^2 + \\ + 6712,0952a_3a_4^2 + 410,94460a_1a_2a_3 + 575,32244a_1a_2a_4 + \\ + 2876,6122a_2a_3a_4 + 246,56676a_2^3 = 0.$$

Система эта решалась по методу Ньютона — Канторовича с начальным приближением

$$a_{10}=a_{20}=a_{30}=a_{40}=a_{50}=0.$$

Ниже приведены результаты третьего приближения; для сравнения приведено и второе приближение:

n	Второе приближение	Третье приближение	n	Второе приближение	Третье приближение
a_1	0,25801631	0,25801628	a_4	0,00075839	0,00075838
a_2	0,00956008	0,00956007	a_5	0,00036394	0,00036394
a_3	0,00206908	0,00206907			

Сравним это с точным решением $y=1-x^2$. Легко получить разложение

$$1-x^2=1-x+\sum_{k=1}^{\infty}\frac{8}{(2k-1)^3\pi^3}\sin(2k-1)\pi x.$$

Приведем значения первых пяти коэффициентов с 8 десятичными знаками:

k	$\frac{8}{(2k-1)^3\pi^3}$	k	$\frac{8}{(2k-1)^3\pi^3}$
1	0,25801227	4	0,00075222
2	0,00955601	5	0,00035393
3	0,00206410		

В приближенных значениях первых четырех коэффициентов верны, следовательно, 5 знаков после запятой; погрешность пятого коэффициента — одна единица пятого знака после запятой.

Значение функционала $F(y_5)$, соответствующее третьему приближению, равно

$$F(y_5)=-2{,}0666036.$$

Точное решение нашей вариационной задачи $y=1-x^2$ дает значение

$$F_{\min}=-2\frac{1}{15}=-2{,}0666667.$$

§ 72. Дифференцирование по параметру

Метод дифференцирования по параметру был предложен почти одновременно несколькими авторами[1]). Этот метод можно распространить на широкий класс уравнений в метрических пространствах, но мы здесь ограничимся его изложением для систем конечного числа

[1]) См., например, С. Кирия [1], Д. Ф. Давиденко [1].

нелинейных уравнений с числовыми неизвестными, какими являются, в частности, системы Ритца.

Пусть требуется решить систему уравнений

$$f_i(a_1, a_2, \ldots, a_n) = 0 \qquad (i = 1, 2, \ldots, n). \tag{72.1}$$

Построим функции $F_i(a_1, a_2, \ldots, a_n; \lambda)$ от переменных $a_1, a_2, \ldots, a_n$ и вспомогательного параметра λ; потребуем, чтобы эти функции были непрерывно дифференцируемы при всех значениях переменных $a_1, a_2, \ldots, a_n$ и при $0 \leqslant \lambda \leqslant 1$. Далее, потребуем, чтобы

$$F_i(a_1, a_2, \ldots, a_n; 1) \equiv f_i(a_1, a_2, \ldots, a_n) \qquad (i = 1, 2, \ldots, n) \tag{72.2}$$

и чтобы система уравнений

$$F_i(a_1, a_2, \ldots, a_n; 0) = 0 \qquad (i = 1, 2, \ldots, n) \tag{72.3}$$

достаточно просто решалась; пусть $a_i^{(0)}$ $(i = 1, 2, \ldots, n)$ — ее решение.

Вместо системы (72.1) будем решать систему

$$F_i(a_1, a_2, \ldots, a_n; \lambda) = 0 \qquad (i = 1, 2, \ldots, n), \tag{72.4}$$

которая, вообще говоря, определяет $a_1, a_2, \ldots, a_n$ как функции от λ; заметим, что нам достаточно знать значения этих функций при $\lambda = 1$.

Система (72.4) очевидным образом равносильна системе обыкновенных дифференциальных уравнений

$$\sum_{k=1}^{n} \frac{\partial F_i}{\partial a_k} \frac{da_k}{d\lambda} + \frac{\partial F_i}{\partial \lambda} = 0, \tag{72.5}$$

которые получаются дифференцированием уравнений (72.4) по λ, с начальными условиями

$$a_i|_{\lambda=0} = a_i^{(0)}. \tag{72.6}$$

Мы пришли, таким образом, к задаче Коши (72.5), (72.6) для определения функций $a_i(\lambda)$; методы ее приближенного решения, как известно, хорошо разработаны. Если задача (72.5), (72.6) имеет решение на отрезке $0 \leqslant \lambda \leqslant 1$, то, построив значение этого решения при $\lambda = 1$, мы тем самым решим систему (72.1).

Дело сводится к тому, чтобы выяснить условия, при которых задача Коши (72.5), (72.6) имеет решение на *конечном* промежутке $0 \leqslant \lambda \leqslant 1$. Ниже мы приведем условия [1]), достаточные для разрешимости такой задачи Коши, к которой приводит система Ритца. Тем самым будет установлена и разрешимость системы Ритца при тех же условиях.

Рассмотрим функционал $F(x)$, заданный на линейном множестве $D(F)$, плотном в некотором сепарабельном банаховом пространстве B; как обычно, допустим, что этот функционал непрерывно

---

[1]) В несколько менее общей форме эти условия даны в статье Л. Н. Довбыш (Гаген-Торн) и автора [1].

дифференцируем на любой конечномерной гиперплоскости из области $D(F)$. Пусть оператор $Q = \operatorname{grad} F$ обладает аналогичными свойствами: его область определения $D(Q)$ плотна в B и линейна, а самый оператор непрерывно дифференцируем на любой конечномерной гиперплоскости из $D(Q)$. Пусть, наконец, при любом $u \in D(Q)$ существует производная Q'_u, определенная на плотном множестве, не зависящем от u и содержащем $D(Q)$, и пусть эта производная симметрична и положительна:

$$\left(Q'_u h_1,\ h_2\right) = \left(h_1,\ Q'_u h_2\right), \qquad \left(Q'_u h,\ h\right) > 0, \qquad h \neq 0.$$

Пусть $\{\varphi_k\}$ — координатная система. Допустим, что $\varphi_k \in D(Q)$ $(k = 1, 2, \ldots)$. Тогда систему Ритца можно записать в виде (70.5). Следуя методу дифференцирования по параметру, построим новую систему, зависящую от параметра λ, так, чтобы при $\lambda = 0$ система решалась без затруднений, а при $\lambda = 1$ она совпадала с системой (70.5). Эту новую систему возьмем в виде

$$a_j + \lambda\left[(Q(u_n), \varphi_j) - a_j\right] = 0 \ (j = 1, 2, \ldots, n), \ u_n = \sum_{k=1}^{n} a_k\varphi_k, \tag{72.7}$$

при $\lambda = 0$ имеем

$$a_j\big|_{\lambda=0} = 0 \qquad (j = 1, 2, \ldots, n). \tag{72.8}$$

Уравнения (72.7) продифференцируем по λ:

$$\frac{da_j}{d\lambda} + \lambda\left\{\sum_{k=1}^{n}\left[\left(Q'_{u_n}\varphi_k,\ \varphi_j\right) - \delta_{jk}\right]\frac{da_k}{d\lambda}\right\} + (Q(u_n), \varphi_j) - a_j = 0 \tag{72.9}$$

$$(j = 1, 2, \ldots, n).$$

Оператор Q'_{u_n} положителен, поэтому можно ввести энергетическое произведение

$$[u, v] = \left(Q'_{u_n}u,\ v\right).$$

Тогда матрицу коэффициентов при производных $\frac{da_k}{d\lambda}$ можно записать в виде

$$(1-\lambda)I + \lambda\begin{Vmatrix} [\varphi_1, \varphi_1] & [\varphi_2, \varphi_1] & \ldots & [\varphi_n, \varphi_1] \\ [\varphi_1, \varphi_2] & [\varphi_2, \varphi_2] & \ldots & [\varphi_n, \varphi_2] \\ \cdot & \cdot & \cdot & \cdot \\ [\varphi_1, \varphi_n] & [\varphi_2, \varphi_n] & \ldots & [\varphi_n, \varphi_n] \end{Vmatrix}, \tag{72.10}$$

где I — единичная матрица. Вторая матрица в (72.10) есть матрица Грама линейно независимых элементов $\varphi_1, \varphi_2, \ldots, \varphi_n$, поэтому она положительно определенная. Но тогда положительно определенной будет и матрица (72.10) при любом $\lambda \in [0, 1]$. Отсюда вытекает, что при указанных значениях λ определитель Δ_n матрицы (72.10) отличен от нуля, и уравнения (72.9) можно разрешить относительно производных $\frac{da_j}{d\lambda}$, которые можно представить по формулам Крамера в виде

$$\frac{da_j}{d\lambda} = \frac{\Delta_n^{(j)}}{\Delta_n} \equiv g_j(a_1, a_2, \ldots, a_n; \lambda) \qquad (j = 1, 2, \ldots, n). \tag{72.11}$$

Примем следующие допущения:

1) Функции $(Q(u_n), \varphi_j)$ и $(Q'_{u_n}\varphi_k, \varphi_j)$ имеют полиномиальный порядок роста:

$$|(Q(u_n), \varphi_j)| \leqslant P_m(|a_1|, |a_2|, \ldots, |a_n|), \tag{72.12}$$

$$|(Q'_{u_n}\varphi_k, \varphi_j)| \leqslant P_{m-1}(|a_1|, |a_2|, \ldots, |a_n|). \tag{72.13}$$

В формулах (72.12) и (72.13) P_m и P_{m-1} означают некоторые полиномы степеней m и $m-1$ соответственно; эти полиномы, так же как и число m, могут зависеть от n.

2) В каждом промежутке $\delta \leqslant \lambda \leqslant 1$, где $\delta > 0$, имеет место оценка снизу

$$(Q'_{u_n}h, h) \geqslant N\left(\sum_{j=1}^{n} a_j^2\right)^{\frac{m-1}{2}} \|h\|^2; \tag{72.14}$$

постоянная N может зависеть от n и от δ.

Из неравенств (72.13) и (72.14) вытекают следующие оценки для определителя Δ_n:

$$N_1\left(\sum_{k=1}^{n} a_k^2\right)^{\frac{n(m-1)}{2}} \leqslant \Delta_n \leqslant P_{n(m-1)}(|a_1|, |a_2|, \ldots, |a_n|), \tag{72.15}$$

где N_1 — положительная постоянная и $P_{n(m-1)}$ — полином степени $n(m-1)$. Правая оценка (72.15) сразу вытекает из неравенства (72.13), и нам достаточно установить левую оценку. Имеем:

$$\sum_{j,k=1}^{n} [\varphi_k, \varphi_j]\, t_k t_j = \left|\sum_{j=1}^{n} \varphi_j t_j\right|^2 = (Q'_{u_n}v_n, v_n),$$

где

$$v_n = \sum_{j=1}^{n} \varphi_j t_j.$$

По неравенству (72.14)

$$(Q'_{u_n}v_n, v_n) \geqslant N\left(\sum_{k=1}^{n} a_k^2\right)^{\frac{m-1}{2}} \left\|\sum_{k=1}^{n} t_k\varphi_k\right\|^2.$$

На единичной сфере выражение

$$\left\|\sum_{k=1}^{n} t_k\varphi_k\right\|^2$$

есть непрерывная неотрицательная функция от $t_1, t_2, \ldots, t_n$, которая ни в какой точке не обращается в нуль, — в противном случае элементы $\varphi_1, \varphi_2, \ldots, \varphi_n$ были бы линейно зависимы. Но тогда эта функция имеет положительную нижнюю грань

$$\left\| \sum_{k=1}^{n} t_k \varphi_k \right\|^2 \geqslant L = \text{const} > 0, \qquad \sum_{k=1}^{n} t_k^2 = 1.$$

Теперь

$$\sum_{j,\,k=1}^{n} [\varphi_k, \varphi_j]\, t_k t_j \geqslant LN \left(\sum_{k=1}^{n} a_k^2 \right)^{\frac{m-1}{2}}, \qquad \sum_{k=1}^{n} t_k^2 = 1.$$

Это неравенство показывает, что все собственные числа матрицы

$$\begin{Vmatrix} [\varphi_1, \varphi_1] & [\varphi_2, \varphi_1] & \ldots & [\varphi_n, \varphi_1] \\ [\varphi_1, \varphi_2] & [\varphi_2, \varphi_2] & \ldots & [\varphi_n, \varphi_2] \\ \cdot & \cdot & \cdot & \cdot \\ [\varphi_1, \varphi_n] & [\varphi_2, \varphi_n] & \ldots & [\varphi_n, \varphi_n] \end{Vmatrix}$$

при $\delta \leqslant \lambda \leqslant 1$ больше, чем $LN \left(\sum\limits_{k=1}^{n} a_k^2 \right)^{\frac{m-1}{2}}$. Но тогда все собственные числа матрицы (72.10) больше, чем

$$\delta LN \left(\sum_{k=1}^{n} a_k^2 \right)^{\frac{m-1}{2}},$$

а определитель Δ_n, равный произведению этих собственных чисел, больше, чем

$$N_1 \left(\sum_{k=1}^{n} a_k^2 \right)^{\frac{n(m-1)}{2}}, \qquad N_1 = (\delta LN)^n.$$

Левая оценка (72.15) доказана.

Докажем теперь, что задача Коши (72.8), (72.9) имеет решение при $0 \leqslant \lambda \leqslant 1$. Прежде всего заметим, что правые части уравнений (72.9) непрерывны и непрерывно дифференцируемы по всем аргументам при $0 \leqslant \lambda \leqslant 1$ и $-\infty < a_j < \infty$ $(j = 1, 2, \ldots, n)$; это следует из того факта, что при указанных значениях аргументов определитель Δ_n положительно ограничен снизу.

Выберем числа $\lambda^* \in (0, 1)$, $a^* \in (0, \infty)$ и рассмотрим область

$$0 \leqslant \lambda \leqslant \lambda^*, \qquad -a^* \leqslant a_j \leqslant a^* \qquad (j = 1, 2, \ldots, n). \tag{72.16}$$

В этой области выполнены все условия классической теоремы существования решения задачи Коши и, следовательно, задача (72.8), (72.9) имеет единственное решение при $\lambda \in [0, \sigma]$, где $\sigma = \min\left(\lambda^*, \frac{a^*}{M}\right)$ и M есть наибольший из максимумов функций g_j в области (72.16).

Возьмем теперь некоторое число δ, $0 < \delta < \sigma$, и пусть при $\lambda = \delta$ будет $a_j = \tilde{a}_j$ $(j = 1, 2, \ldots, n)$. Рассмотрим область

$$\delta \leqslant \lambda \leqslant 1, \qquad -\infty < a_j < +\infty \qquad (j = 1, 2, \ldots, n) \qquad (72.17)$$

и поставим задачу Коши для системы (72.9) при начальных условиях

$$a_j|_{\lambda=\delta} = \tilde{a}_j \qquad (j = 1, 2, \ldots, n). \qquad (72.18)$$

Из оценок (72.12)—(72.15) вытекает, что производные $\frac{\partial g_j}{\partial a_k}$ ограничены в области (72.17), а тогда, как известно[1]), существует единственное решение задачи (72.9), (72.18), определенное в промежутке $\delta \leqslant \lambda \leqslant 1$.

Теперь ясно, что задача (72.8), (72.9) имеет решение, определенное при $0 \leqslant \lambda \leqslant 1$; отсюда вытекает, в частности, разрешимость системы Ритца (67.5).

В статье Л. Н. Довбыш (Гаген-Торн) [1] показано, что для функционала (68.1) в весьма широких условиях имеют место оценки (72.12)—(72.14) при $m = 1$.

Замечание 1. В нашем изложении система дифференциальных уравнений, полученная дифференцированием по параметру, служила не только средством для приближенного решения данной системы, но также и средством для доказательства существования самого решения. Можно действовать иначе: можно установить условия, при которых система (72.4) имеет решение, определенное и непрерывно дифференцируемое при $0 \leqslant \lambda \leqslant 1$; тем самым автоматически доказывается существование решения задачи Коши (72.5), (72.6), и остается только построить это решение, используя какой-либо из приближенных методов. Такой прием использован в работе З. А. Власовой [1], посвященной методу сеток; об этой работе у нас пойдет речь в следующем параграфе. К системам Ритца этот прием применил М. Н. Яковлев[2]) в статье [1].

Приведем основные результаты этой статьи. Для краткости обозначим через a, f, F векторы, составляющие которых суть a_j, f_j, F_j $(j = 1, 2, \ldots, n)$, через J — якобиеву матрицу вектора f. Система $a + \lambda(f - a) = 0$ имеет решение, непрерывно дифференцируемое на отрезке $0 \leqslant \lambda \leqslant 1$, если выполнены следующие условия:

---

[1]) См., например, Дж. Сансоне [1].

[2]) Результаты М. Н. Яковлева опираются на одну теорему С. М. Лозинского [1], полное доказательство которой, насколько автору известно, до сих пор еще не опубликовано.

а) Существует число r^*, $0 < r^* \leqslant \infty$, такое, что в сфере $\|a\| \leqslant r$, $0 \leqslant r < r^*$, имеет место неравенство

$$(J(a)h,\ h) \geqslant m(r)\|h\|^2,$$

где h — произвольный вектор, а $m(r)$ — функция, положительная и невозрастающая в промежутке $0 \leqslant r < r^*$.

б) Справедливо неравенство

$$\|f(0)\| \leqslant \int_0^{r^*} m^*(\rho)\,d\rho, \qquad m^*(\rho) = \min\{1,\ m(\rho)\}.$$

При тех же условиях, очевидно, имеет решение и задача Коши

$$\frac{da_j}{d\lambda} + \lambda\left[\sum_{k=1}^{n} \frac{\partial f_j}{\partial a_k} - \delta_{jk}\right]\frac{da_k}{d\lambda} + f_j - a_j = 0, \qquad a_j|_{\lambda=0} = 0.$$

В статье [1] М. Н. Яковлев рассмотрел еще один способ введения параметра: вектор-функция $F(x,\ \lambda)$ определяется формулой

$$F(x,\ \lambda) = f(x) - (1-\lambda)f(0).$$

Существование и непрерывная дифференцируемость решения уравнения $F(x,\ \lambda) = 0$, а с ним и соответствующей задачи Коши устанавливается, если выполнены приведенные выше условия а) и формулируемые ниже условия:

в) При $0 \leqslant r \leqslant r^*$ верно неравенство $\|J^{-1}(a)\| \leqslant D(r)$, где $D(r)$ — непрерывная неубывающая функция.

г) $\|f(0)\| \leqslant \rho^*$, где $\rho^* = \int_0^{r^*} \frac{1}{D(\rho)}\,d\rho$.

Замечание 2. Как показывают численные примеры, точность метода дифференцирования по параметру невелика. Целесообразно, по-видимому, использовать приближение, полученное методом продолжения по параметру, как начальное для метода Ньютона — Канторовича.

§ 73. Применение к сеточным уравнениям

Обычно задачи вариационного исчисления представляют собой задачи об экстремуме некоторого интеграла; в таком случае приближенное решение задачи можно строить и по методу сеток. Для этого интеграл заменяем конечной суммой по той или иной квадратурной формуле, а производные под знаком интеграла заменяем разностными отношениями. Далее, ставится задача об экстремуме полученного таким образом нового функционала, определенного уже в конечномерном пространстве; уравнение Эйлера — Лагранжа для этого нового функционала представляет собой систему сеточных уравнений для функционала исходного. Если этот последний — не квадратичный, то сеточная система нелинейна, и для ее приближенного решения можно использовать метод дифференцирования по параметру. Для так называемой «простейшей задачи вариационного исчисления» это сделано в статье З. А. Власовой [1], которую мы в существенном воспроизводим ниже.

Рассмотрим задачу: найти кривую, проходящую через точки $A(a, a_1)$ и $B(b, b_1)$, на которой интеграл

$$\int_a^b \varphi(x, y, y')\,dx \tag{73.1}$$

принимает минимальное значение; функцию $\varphi(x, y, y')$ считаем непрерывной вместе с ее производными до второго порядка включительно при $x \in [a, b]$ и любых конечных y и y'.

Выберем на отрезке $[a, b]$ $n+2$ точки

$$x_0 = a,\ x_1,\ \ldots,\ x_n,\ x_{n+1} = b.$$

Используя их как узлы квадратурной формулы с коэффициентами $C_k > 0$ и выражая производную в узлах квадратурной формулы через первую разность, можем написать приближенное равенство

$$\int_a^b \varphi(x, y, y')\,dx \approx \sum_{k=0}^{n} C_k \varphi\left(x_k,\ y_k,\ \frac{y_{k+1} - y_k}{\Delta x_k}\right),$$

где

$$\Delta x_k = x_{k+1} - x_k, \qquad y_0 = a_1, \qquad y_{n+1} = b_1.$$

Замена производной $y'(x_k)$ разностным отношением $\frac{y_{k+1} - y_k}{\Delta x_k}$ является простейшей, но не очень точной. В примере (см. ниже, § 74) будет использована более точная замена.

Введем в рассмотрение какую-либо достаточно гладкую функцию $C(x)$, принимающую в узлах x_k значения C_k.

Обозначим

$$\psi(x, y, y') = C(x)\,\varphi(x, y, y').$$

Сформулированную выше точную вариационную задачу заменяем приближенной: найти вектор $Y = (y_1, \ldots, y_n)$, на котором сумма

$$\sum_{k=0}^{n} \psi\left(x_k,\ y_k,\ \frac{y_{k+1} - y_k}{\Delta x_k}\right) = f(Y) \tag{73.2}$$

принимает минимальное значение.

Необходимое условие минимума приводит к следующей системе для нахождения Y:

$$\left.\frac{\partial\psi}{\partial y}\right|_k-\frac{1}{\Delta x_k}\left.\frac{\partial\psi}{\partial y'}\right|_k+\frac{1}{\Delta x_{k-1}}\left.\frac{\partial\psi}{\partial y'}\right|_{k-1}=0. \tag{73.3}$$

Здесь и ниже $\left.\frac{\partial\psi}{\partial y}\right|_k$ означает частную производную функции $\psi(x, y, y')$ по y, вычисленную в точке $\left(x_k, y_k, \frac{y_{k+1}-y_k}{\Delta x_k}\right)$. Аналогичные обозначения будут использованы и для других производных.

Как хорошо известно, для того чтобы вектор Y, удовлетворяющий системе (73.3), реализовал относительный минимум функционала (73.2), достаточно, чтобы производная его градиента была положительно определенной при любом Y и для любого вектора $H=(h_1, h_2, \ldots, h_n)$. Обозначим $P=\operatorname{grad} f$. Произведя необходимые вычисления, получим:

$$(P'_Y H, H)=\sum_{k=0}^{n}\left[\left(p_k-q_k+q_{k-1}\right)h_k^2+\left(r_k-q_k\right)\left(h_{k+1}-h_k\right)^2\right], \tag{73.4}$$

где

$$p_k=\left.\frac{\partial^2\psi}{\partial y^2}\right|_k,\qquad q_k=\frac{1}{\Delta x_k}\left.\frac{\partial^2\psi}{\partial y\,\partial y'}\right|_k,\qquad r_k=\frac{1}{(\Delta x_k)^2}\left.\frac{\partial^2\psi}{\partial y'^2}\right|_k.$$

Пусть подынтегральная функция $\varphi(x, y, y')$ и квадратурная формула, определяемая узлами x_k и коэффициентами C_k, таковы, что при любых конечных y_k выполнены неравенства

$$r_k-q_k>0, \tag{73.4$'$}$$

$$p_k-q_k+q_{k-1}\geqslant\delta\qquad(k=1, 2, \ldots, n), \tag{73.4$''$}$$

где δ — некоторая положительная постоянная. Тогда

$$\left(P'_Y H, H\right)\geqslant\delta\sum_{k=1}^{n}{}' h_k^2>0$$

при любом векторе Y.

Поэтому, если выполнены условия (73.4′) и (73.4″), то для нахождения экстремального вектора достаточно решить систему (73.3).

Замечание 1. При достаточно малых Δx_k условие (73.4′) вытекает из усиленного условия Лежандра[1]). Условие (73.4″) при малых Δx_k и $C_1=C_2=\ldots=C_n$ вытекает из условия С. Н. Бернштейна[2]), написанного для уравнения Эйлера нашей задачи, и из того же усиленного условия Лежандра. Для уравнения вида

[1]) См., например, Н. И. Ахиезер [1], стр. 68.

[2]) С. Н. Бернштейн [1], стр. 37, а также Н. И. Ахиезер [1], стр. 36.

$y'' = f(x, y, y')$ условие С. Н. Бернштейна имеет вид $\frac{\partial f}{\partial y} \geqslant K$, где K — положительная постоянная.

Для решения системы (73.3), в общем случае нелинейной, применим метод дифференцирования по параметру.

Введем систему, содержащую параметр λ, $0 \leqslant \lambda \leqslant 1$,

$$N_k(\lambda, y_{k-1}, y_k, y_{k+1}) = (1-\lambda)(-\delta^2 y_k + T_k y_k) + \\ + \lambda \left(\left.\frac{\partial \psi}{\partial y}\right|_k - \frac{1}{\Delta x_k} \left.\frac{\partial \psi}{\partial y'}\right|_k + \frac{1}{\Delta x_{k-1}} \left.\frac{\partial \psi}{\partial y'}\right|_{k-1} \right) = 0 \qquad (73.5)$$
$$(k = 1, 2, \ldots, n),$$

где

$$\delta^2 y_k = R_k y_{k+1} - (R_k + R_{k-1}) y_k + R_{k-1} y_{k-1},$$

а R_k и T_k — произвольно выбранные положительные постоянные.

При $\lambda = 1$ система (73.5) совпадает с системой (73.3), а при $\lambda = 0$ она вырождается в линейную систему

$$-\delta^2 y_k + T_k y_k = 0 \qquad (k = 1, 2, \ldots, n). \qquad (73.6)$$

Квадратичная форма, соответствующая матрице коэффициентов системы (73.6) и равная выражению

$$\sum_{k=0}^{n} \left[R_k (h_{k+1} - h_k)^2 + T_k h_k^2 \right],$$

— положительно определенная. Поэтому решение системы (73.6) существует и единственно.

В частности, если $a_1 > b_1 \geqslant 0$, то можно взять

$$R_k = \frac{1}{2^{k-n}} \frac{1}{\Delta x_k}, \qquad T_k = \frac{a_1 - b_1}{2^{k-n} [a_1 (b - x_k) + b_1 (x_k - a)]}.$$

Тогда вектор Y с компонентами, определяемыми формулой

$$y_k = \frac{a_1 (b - x_k) + b_1 (x_k - a)}{b - a} \qquad (k = 1, 2, \ldots, n),$$

удовлетворяет системе (73.6).

Продифференцируем систему (73.5) по параметру λ. Это приведет к системе линейных уравнений относительно неизвестных $\frac{\partial y_l}{\partial \lambda}$:

$$\left. \begin{aligned} &\sum_{l=k-1}^{k+1} \frac{\partial N_k}{\partial y_l} \frac{d y_l}{d\lambda} = -\frac{\partial N_k}{\partial \lambda} \qquad (k = 1, 2, \ldots, n), \\ &\frac{d y_0}{d\lambda} = \frac{d y_{n+1}}{d\lambda} = 0. \end{aligned} \right\} \qquad (73.7)$$

Введем следующие обозначения:

$$\left.\begin{aligned}
\alpha_{k-1} &= -\frac{\partial N_k}{\partial y_{k-1}} = \lambda(r_{k-1} - q_{k-1}) + (1-\lambda) R_{k-1},\\
\beta_k &= \frac{\partial N_k}{\partial y_k} = \lambda(p_k - 2q_k + r_k + r_{k-1}) +\\
&\qquad\qquad + (1-\lambda)[T_k + R_k + R_{k-1}],\\
\alpha_k &= -\frac{\partial N_k}{\partial y_{k+1}} = \lambda(r_k - q_k) + (1-\lambda) R_k,\\
\gamma_k &= \frac{\partial N_k}{\partial \lambda} = \delta^2 y_k - T_k y_k +\\
&\qquad + \left.\frac{\partial\psi}{\partial y}\right|_k - \frac{1}{\Delta x_k}\left.\frac{\partial\psi}{\partial y'}\right|_k + \frac{1}{\Delta x_{k-1}}\left.\frac{\partial\psi}{\partial y'}\right|_{k-1}.
\end{aligned}\right\} \qquad (73.8)$$

Очевидно, матрица A коэффициентов системы (73.7) — симметричная трехдиагональная и имеет вид

$$A_n = \begin{vmatrix} \beta_1 & -\alpha_1 & 0 \ldots & 0 & 0\\ -\alpha_1 & \beta_2 & -\alpha_2 \ldots & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 \ldots & -\alpha_{n-1} & \beta_n \end{vmatrix}.$$

Покажем, что система (73.7) разрешима относительно неизвестных $\frac{dy_l}{d\lambda}$, для чего рассмотрим квадратичную форму, соответствующую матрице A_n и равную выражению

$$(A_n H, H) = \sum_{k=1}^{n} \{(1-\lambda) T_k + \lambda(p_k - q_k + q_{k-1})\} h_k^2 + \\ + \sum_{k=0}^{n} \{(1-\lambda) R_k + \lambda(r_k - q_k)\}(h_k - h_{k+1})^2.$$

Из условий (73.4′) и (73.4″) следует неравенство:

$$(A_n H, H) \geqslant \delta_1 \sum_{k=1}^{n} h_k^2, \qquad 0 < \delta_1 \leqslant \delta\lambda + (1-\lambda) \min_{1 \leqslant k \leqslant n} T_k,$$

обеспечивающее положительную определенность квадратичной формы $(A_n H, H)$, а также оценка снизу для определителя Δ_n матрицы A_n:

$$\Delta_n \geqslant \delta_1^n > 0. \qquad (73.9)$$

Таким образом, систему (73.7) можно разрешить относительно неизвестных $\frac{dy_l}{d\lambda}$. По формуле Крамера имеем:

$$\frac{dy_l}{d\lambda} = \frac{\Delta_n^l}{\Delta_n}, \qquad \lambda \in [0, 1] \qquad (l = 1, 2, \ldots, n). \qquad (73.10)$$

Теперь убедимся, что на всем сегменте [0, 1] существует единственное решение системы (73.10), выходящее из точки

$$\lambda = 0, \quad y_0 = a_1, \quad y_k = \bar{y}_k \quad (k = 1, 2, \ldots, n), \quad y_{n+1} = b_1,$$

где $\bar{y}_k$ $(k = 1, 2, \ldots, n)$ есть решение системы (73.6) при выбранных значениях R_k, T_k.

Для этого докажем, что система (73.5) при любом $\lambda \in [0, 1]$ разрешима единственным образом относительно неизвестных $y_1, y_2, \ldots$ $\ldots, y_n$; при этом функции $y_k(\lambda)$ $(k = 1, 2, \ldots, n)$ будут непрерывны и непрерывно дифференцируемы по λ.

Предварительно установим априорную оценку решений системы (73.5).

Лемма 73.1. *Если система* (73.5) *имеет решение, то оно равномерно ограничено при* $\lambda \in [0, 1]$.

Пусть $y_k(\lambda)$, $\bar{y}_k$ — решения систем (73.5) и (73.6) соответственно, $\delta y_k(\lambda) = y_k(\lambda) - \bar{y}_k$. Фиксируем некоторое $\lambda = \lambda_0 \neq 0$.

По формуле конечных приращений для любого k имеем:

$$-N_k(\lambda_0, \bar{y}_{k-1}, \bar{y}_k, \bar{y}_{k+1}) = \\ = \frac{\partial \tilde{N}_k}{\partial y_{k-1}} \delta y_{k-1}(\lambda_0) + \frac{\partial \tilde{N}_k}{\partial y_k} \delta y_k(\lambda_0) + \frac{\partial \tilde{N}_k}{\partial y_{k+1}} \delta y_{k+1}(\lambda_0).$$

Знак $\sim$ обозначает здесь, что все входящие в выражения $\frac{\partial N_k}{\partial y_{k-1}}$, $\frac{\partial N_k}{\partial y_k}$, $\frac{\partial N_k}{\partial y_{k+1}}$ частные производные функции $\varphi(x, y, y')$ берутся в точках

$$\left(x_k, \ \bar{y}_k + \theta_k \delta y_k(\lambda_0), \ \frac{\bar{y}_{k+1} - \bar{y}_k}{\Delta x_k} + \theta_k \frac{\delta y_{k+1}(\lambda_0) - \delta y_k(\lambda_0)}{\Delta x_k}\right)$$

и

$$\left(x_{k-1}, \ \bar{y}_{k-1} + \theta_k \delta y_{k-1}(\lambda_0), \ \frac{\bar{y}_k - \bar{y}_{k-1}}{\Delta x_{k-1}} + \theta_k \frac{\delta y_k(\lambda_0) - \delta y_{k-1}(\lambda_0)}{\Delta x_k}\right), \\ 0 < \theta_k < 1.$$

Перепишем предыдущее равенство, пользуясь обозначениями (73.8):

$$-N_k(\lambda_0, \bar{y}_{k-1}, \bar{y}_k, \bar{y}_{k+1}) = \\ = \tilde{\alpha}_{k-1}[\delta y_k(\lambda_0) - \delta y_{k-1}(\lambda_0)] + (\tilde{\beta}_k - \tilde{\alpha}_{k-1} - \tilde{\alpha}_k)\delta y_k(\lambda_0) + \\ + \tilde{\alpha}_k[\delta y_k(\lambda_0) - \delta y_{k+1}(\lambda_0)]. \quad (73.11)$$

Из условий (73.4′), (73.4″) и положительности постоянных T_k, R_k следует, что при $\lambda \in [0, 1]$

$$\left.\begin{aligned} &\lambda_k \geqslant 0, \\ \beta_k = \alpha_k + \alpha_{k-1} + \lambda(p_k - q_k + q_{k-1}) + (1-\lambda) T_k \geqslant \\ &\geqslant \alpha_k + \alpha_{k+1} + \delta_1. \end{aligned}\right\} \quad (73.12)$$

Пусть при $k=t$ функция $|\delta y_k(\lambda_0)|$ достигает максимума, например,

$$\delta y_t(\lambda_0)>0, \quad \delta y_t(\lambda_0)\geqslant\delta y_{t-1}(\lambda_0),$$
$$\delta y_t(\lambda_0)\geqslant\delta y_{t+1}(\lambda_0).$$

Из (73.11) и (73.12) следует

$$[y_t(\lambda_0)-\bar{y}_t]\leqslant-\frac{1}{\delta_1}N_t(\lambda_0, \bar{y}_{t-1}, \bar{y}_t, \bar{y}_{t+1}).$$

В случае отрицательного минимума функции $\delta y_k(\lambda_0)$ аналогичное неравенство выводится для $-\delta y_t(\lambda_0)$.

Поскольку λ_0 — произвольное число из сегмента $[0,1]$, можно написать оценку решений системы (73.5):

$$|y_k(\lambda)-\bar{y}_k|\leqslant\frac{1}{\delta_1}\max_{\substack{1\leqslant k\leqslant n\\ 0\leqslant\lambda\leqslant 1}}|N_k(\lambda, \bar{y}_{k-1}, \bar{y}_k, \bar{y}_{k+1})|=m.$$

Лемма доказана.

Для доказательства существования решения системы (73.5) нам придется использовать теорему о неявных функциях для случая уравнения вида $F(x, y, z)=0$.

Как известно, эта теорема носит локальный характер и утверждает разрешимость уравнения $F(x, y, z)=0$ относительно z в некоторой окрестности начальной точки.

Уточним применительно к системе (73.8) размеры этой окрестности.

Лемма 73.2. *Пусть в области D, определяемой неравенствами*

$$0\leqslant x-x_0\leqslant\Delta x, \quad \Delta x>0,$$
$$|y-y_0|\leqslant\Delta y, \quad \Delta y>0,$$
$$|z-z_0|\leqslant\Delta z, \quad \Delta z>0,$$

функция $F(x, y, z)$ непрерывна и имеет непрерывные частные производные, причем

$$\left.\begin{aligned} |F_x|\leqslant M, \quad |F_y|\leqslant M,\\ F_z\geqslant k>0,\\ F(x_0, y_0, z_0)=0.\end{aligned}\right\} \tag{73.13}$$

Тогда утверждения теоремы о неявных функциях справедливы в области

$$\left.\begin{aligned} 0\leqslant x-x_0\leqslant\min\left\{\frac{k}{2M}\Delta z, \ \Delta x\right\},\\ |y-y_0|\leqslant\min\left\{\frac{k}{2M}\Delta z, \ \Delta y\right\}.\end{aligned}\right\} \tag{73.14}$$

Для доказательства леммы применим формулу конечных приращений к функции $F(x_0, y_0, z)$:

$$F(x_0, y_0, z_0+\Delta z)=F_z(x_0, y_0, z_0+\theta_1\Delta z)\Delta z, \qquad 0<\theta_1<1.$$

Отсюда и из оценки (73.13) следует соотношение:

$$F(x_0, y_0, z_0+\Delta z)\geqslant k\Delta z.$$

Определим сегмент на полупрямой

$$x\geqslant x_0, \qquad y=y_0, \qquad z=z_0+\Delta z,$$

в котором справедливо неравенство

$$F(x, y_0, z_0+\Delta z)\geqslant \frac{k}{2}\Delta z. \tag{73.15}$$

Пусть в точке $\xi\in[x_0, x_0+\Delta x]$ достигается равенство

$$F(\xi, y_0, z_0+\Delta z)=\frac{k}{2}\Delta z.$$

По формуле конечных приращений имеем:

$$F(\xi, y_0, z_0+\Delta z)=F(x_0, y_0, z_0+\Delta z)+$$
$$+F_x[x_0+\theta_2(\xi-x_0), y_0, z_0+\Delta z](\xi-x_0), \qquad 0<\theta_2<1.$$

Если $\xi-x_0<\Delta x$, то

$$\xi-x_0\geqslant\frac{k}{2M}\Delta z.$$

Следовательно, неравенство (73.15) выполняется при

$$0\leqslant x-x_0\leqslant\min\left\{\frac{k}{2M}\Delta z, \Delta x\right\}.$$

Подобным же путем находим, что в области d, определяемой неравенствами (73.14), функция $F(x, y, z_0+\Delta z)$ неотрицательна.

Повторяем рассуждения для плоскости $z=z_0-\Delta z$, убеждаемся, что в области d функция $F(x, y, z_0-\Delta z)$ неположительна.

Из доказанных утверждений вытекает существование неявной функции в области d. Ее непрерывность и непрерывная дифференцируемость доказываются обычными средствами.

Теорема 73.1. *При сформулированных выше требованиях гладкости подынтегральной функции* $\varphi(x, y, y')$ *и выполнении условий* (73.4′) *и* (73.4″) *система* (73.5) *разрешима на сегменте* [0, 1] *единственным образом и определяет функции*

$$y_k=Q_k(\lambda) \qquad (k=1, 2, \ldots, n),$$

непрерывные вместе со своими первыми производными на том же сегменте.

Доказательство. Из непрерывности функции $\varphi(x, y, y')$ и ее производных, входящих в выражение (73.8), а также из оценки (73.9) вытекает существование таких постоянных S и δ_0, что при $\lambda \in [0, 1]$, $|y_k(\lambda) - \bar{y}_k| \leqslant 3m$ для всех k имеют место неравенства

$$\left.\begin{gathered} |\gamma_k| \leqslant S, \quad \alpha_k \leqslant S, \quad \beta_k \leqslant S, \quad \Delta_k \leqslant S, \\ \frac{\alpha_{k-1}\Delta_{k-2}}{\Delta_{k-1}} \leqslant S, \qquad \frac{\Delta_k}{\Delta_{k-1}} \geqslant \delta_0, \\ \left|\gamma_k + \frac{\alpha_{k-1}\Delta_{k-2}}{\Delta_{k-1}}\left(\gamma_{k-1} + \ldots + \frac{\alpha_1\Delta_0}{\Delta_1}\gamma_1\right)\right| \leqslant S. \end{gathered}\right\} \tag{73.16}$$

Выбрав S достаточно большим, можно считать

$$\frac{1}{S}\cdot\frac{\delta_0}{2} \leqslant 1.$$

При $\lambda = 0$ система (73.5) имеет решение, равное $\bar{y}_k$ ($k = 1, 2, \ldots, n$). Продолжим это решение по параметру λ.

Из леммы 73.2 и теоремы о неявных функциях следует, что в области, задаваемой неравенствами

$$0 \leqslant \lambda \leqslant \min\left\{2m\left(\frac{\delta_0}{2S}\right)^{n-1}, \ 1\right\},$$

$$|y_{j+1} - \bar{y}_{j+1}| \leqslant 2m\left(\frac{\delta_0}{2S}\right)^j \qquad (j = 0, 1, 2, \ldots, n-1),$$

система (73.5) определяет рекуррентные соотношения

$$y_j = \varphi_j(\lambda, y_{j+1}), \qquad \varphi_j(0, \bar{y}_{j+1}) = \bar{y}_j \qquad (j = 1, 2, \ldots, n-1),$$

при этом функции $\varphi_j(\lambda, y_{j+1})$ непрерывны и имеют непрерывные частные производные

$$\left.\begin{gathered} \frac{\partial\varphi_j}{\partial\lambda} = -\frac{\Delta_{j-1}}{\Delta_j}\left\{\gamma_j + \frac{\alpha_{j-1}\Delta_{j-2}}{\Delta_{j-1}}\left[\gamma_{j-1} + \ldots + \frac{\alpha_1\Delta_0}{\Delta_1}\gamma_1\right]\right\}, \\ \frac{\partial\varphi_j}{\partial y_{j+1}} = \frac{\alpha_j\Delta_{j-1}}{\Delta_j}. \end{gathered}\right\} \tag{73.17}$$

Действительно, допустим, что высказанное выше утверждение справедливо для $n = k$ (для $n = 2$ это получается непосредственным применением леммы 73.2 и теоремы о неявных функциях к первому уравнению системы (73.5)).

Применим лемму 73.2 и теорему о неявных функциях к $(k+1)$-му уравнению системы (73.5).

В качестве $F(x, y, z)$ берем функцию

$$N_k[\lambda, \varphi_{k-1}(\lambda, y_k), y_k, y_{k+1}] = \Phi_k(\lambda, y_k, y_{k+1})$$

и полагаем

$$x = \lambda, \qquad y = y_{k+1}, \qquad z = y_k.$$

Зададим область D соотношениями:

$$0 \leqslant \lambda \leqslant \min\left\{2m\left(\frac{\delta_0}{2S}\right)^{k-1},\ 1\right\},$$
$$|y_k - \bar{y}_k| \leqslant 2m\left(\frac{\delta_0}{2S}\right)^{k-1},$$
$$|y_{k+1} - \bar{y}_{k+1}| \leqslant 2m.$$

Вычислим используемые в лемме 73.2 частные производные функции $\Phi_k(\lambda,\ y_k,\ y_{k+1})$, пользуясь выражениями $\frac{\partial\varphi_{k-1}}{\partial\lambda}$, $\frac{\partial\varphi_{k-1}}{\partial y_k}$:

$$\frac{\partial\Phi_k}{\partial\lambda} = \frac{\partial N_k}{\partial\lambda} + \frac{\partial N_k}{\partial y_{k-1}}\frac{\partial\varphi_{k-1}}{\partial\lambda} =$$
$$= \gamma_k + \frac{\alpha_{k-1}\Delta_{k-2}}{\Delta_{k-1}}\left\{\gamma_{k-1} + \ldots + \frac{\alpha_1\Delta_0}{\Delta_1}\gamma_1\right\},$$
$$\frac{\partial\Phi_k}{\partial y_k} = \frac{\partial N_k}{\partial y_k} + \frac{\partial N_k}{\partial y_{k-1}}\frac{\partial\varphi_{k-1}}{\partial y_k} = \frac{\Delta_k}{\Delta_{k-1}},$$
$$\frac{\partial\Phi_k}{\partial y_{k+1}} = -\alpha_k.$$

На основании неравенств (73.13) и (73.16) можем положить $M = S$, $k = \delta_0$.

По лемме 73.2 и теореме о неявных функциях находим в области

$$0 \leqslant \lambda \leqslant \min\left\{2m\left(\frac{\delta_0}{2S}\right)^{k},\ 1\right\}, \qquad |y_{k+1} - \bar{y}_{k+1}| \leqslant 2m\left(\frac{\delta_0}{2S}\right)^{k}$$

непрерывную однозначную функцию $y_k = \varphi_k(\lambda,\ y_{k+1})$, $\varphi_k(0,\ \bar{y}_{k+1}) = \bar{y}_k$, частные производные которой выражаются формулами (73.17), написанными для индекса k.

Продолжая этот процесс, доходим до последнего уравнения системы (73.5)

$$N_n[\lambda,\ \varphi_n(\lambda,\ y_n),\ y_n,\ b_1] = \Phi_n(\lambda,\ y_n,\ b_1) = 0,$$

разрешая которое, определим в области $d^{(n)}$:

$$0 \leqslant \lambda \leqslant \min\left\{2m\left(\frac{\delta_0}{2S}\right)^{n},\ 1\right\} = \Delta\lambda \tag{73.18}$$

функцию $y_n = \varphi_n(\lambda,\ b_1) = Q_n(\lambda)$, обладающую всеми свойствами, указанными в лемме 73.2 и теореме о неявных функциях.

Далее, проделывая обратный ход, строим последовательно в области $d^{(n)}$ непрерывные однозначные функции

$$y_k = Q_k(\lambda), \qquad Q_k(0) = \bar{y}_k \qquad (k = 1,\ 2,\ \ldots,\ n),$$

имеющие непрерывные первые производные.

Таким образом, мы доказали справедливость теоремы 73.1 для сегмента $[0, \Delta\lambda]$, где $\Delta\lambda > 0$ и определяется соотношением (73.18). Этот сегмент может или совпасть со всем сегментом $[0, 1]$, или явиться его частью.

В последнем случае полученное решение надо продолжить на весь сегмент $[0, 1]$, что и будет выполнено ниже.

Возьмем уже найденные значения функции $y_k(\lambda)$ при $\lambda = \frac{\Delta\lambda}{2}$ за новые начальные данные

$$y_k\left(\frac{\Delta\lambda}{2}\right) = \overline{\overline{y}}_k \qquad (k = 1, 2, \ldots, n).$$

Из леммы 73.1 следует, что $|\overline{\overline{y}}_k - \overline{y}_k| \leqslant m$, $|\overline{y}_k - \overline{\overline{y}}_k| \leqslant 2m$.

Так как оценки (73.16), используемые при доказательстве теоремы, справедливы для любого $\lambda \in [0, 1]$, то аналогичными рассуждениями докажем справедливость теоремы либо на всем сегменте $[0, 1]$, либо на сегменте $\left[0, \frac{\Delta\lambda}{2} + \Delta\lambda\right]$.

Из соотношения (73.18) следует, что

$$\Delta\lambda \geqslant 2m\left(\frac{\delta_0}{2S}\right)^n.$$

Следовательно, проделав по λ меньше чем $\frac{1}{m}\left(\frac{2S}{\delta_0}\right)^n$ шагов, дойдем до точки $\lambda = 1$.

Замечание 2. Если точка $x_{n+1} = b$ является узлом квадратурной формулы, т. е. мы совершаем следующую замену интеграла на сумму

$$\int_a^b \varphi(x, y, y')\,dx =$$

$$= \sum_{k=0}^{n} C_k\varphi\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x_k}\right) + C_{n+1}\varphi\left(x_{n+1}, y_{n+1}, \frac{y_{n+1} - y_n}{\Delta x_n}\right),$$

то в последнем уравнении системы (73.3) добавится слагаемое, и уравнение примет форму

$$\psi_y\Big|_n - \frac{1}{\Delta x_n}\psi_{y'}\Big|_n + \frac{1}{\Delta x_{n-1}}\psi_{y'}\Big|_{n-1} -$$

$$- \frac{1}{\Delta x_n}\psi_{y'}\left(x_{n+1}, y_{n+1}, \frac{y_{n+1} - y_n}{\Delta x_n}\right) = 0.$$

Кроме того, изменятся выражения β_n и γ_n:

$$\beta_n = \lambda(p_n - 2q_n + r_n + r_{n-1}) + (1-\lambda)(R_n + R_{n-1} + T_n) + \\ + \frac{1}{(\Delta x_n)^2}\psi_{y'y'}\left(x_{n+1},\ y_{n+1},\ \frac{y_{n+1}-y_n}{\Delta x_n}\right),$$

$$\gamma_n = \delta^2 y_n - T_n y_n + \psi_y\Big|_n - \frac{1}{\Delta x_n}\psi_{y'}\Big|_n + \frac{1}{\Delta x_{n-1}}\psi_{y'}\Big|_{n-1} - \\ - \frac{1}{\Delta x_n}\psi_{y'}\left(x_{n+1},\ y_{n+1},\ \frac{y_{n+1}-y_n}{\Delta x_n}\right).$$

Все остальные уравнения систем (73.3) и (73.7) полностью сохраняются. Лемма 73.1 и теорема 73.1 также остаются справедливыми.

§ 74. Пример

В настоящем параграфе мы рассмотрим численный пример, в котором применены соображения предшествующего параграфа [1]).

Прежде чем перейти к самому примеру, изложим некоторые общие соображения, использованные при его решении. Мы несколько сузим общность построений предшествующего параграфа и откажемся от разбиения основного промежутка на части произвольной длины; будем считать, что промежуток $[a, b]$ разбит на n равных частей длины $h = \frac{b-a}{n}$ точками деления $x_k = a + kh$ $(k = 0, 1, \ldots, n)$. Одновременно мы применим более точную аппроксимацию производной, чем в § 73, а именно, мы положим:

$$\left.\begin{aligned} y'_k &= y'(x_k) \approx \frac{y_{k+1} - y_{k-1}}{2h} \qquad (k = 1, 2, \ldots, n-1), \\ y'_0 &= y'(x_0) \approx \frac{-3y_0 + 4y_1 - y_2}{2h}, \\ y'_n &= y'(x_n) \approx \frac{y_{n-2} - 4y_{n-1} + 3y_n}{2h}. \end{aligned}\right\} \tag{74.1}$$

Как и в § 73, заменяем точную вариационную задачу приближенной: найти вектор $Y = (y_1, y_2, \ldots, y_{n-1})$, на котором сумма

$$f(y_1, y_2, \ldots, y_{n-1}) = \sum_{k=0}^{n} C_k \varphi(x_k, y_k, y'_k) = \\ = \psi\left(x_0, y_0, \frac{-3y_0 + 4y_1 - y_2}{2h}\right) + \sum_{k=1}^{n-1}\psi\left(x_k, y_k, \frac{y_{k+1} - y_{k-1}}{2h}\right) + \\ + \psi\left(x_n, y_n, \frac{y_{n-2} - 4y_{n-1} + 3y_n}{2h}\right) \tag{74.2}$$

принимает минимальное значение.

---

[1]) Этот пример дан в статье З. А. Власовой и Ю. В. Рыбаковой [1], которую мы в основном воспроизводим в настоящем параграфе.

Из необходимого условия минимума получаем систему $n-1$ нелинейных уравнений относительно неизвестных $y_1, y_2, \ldots, y_{n-1}$:

$$\left.\begin{aligned}
&\frac{\partial f}{\partial y_1}=\left.\frac{\partial\psi}{\partial y}\right|_1-\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_2+\frac{2}{h}\left.\frac{\partial\psi}{\partial y'}\right|_0=0,\\
&\frac{\partial f}{\partial y_2}=\left.\frac{\partial\psi}{\partial y}\right|_2-\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_3+\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_1-\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_0=0,\\
&\frac{\partial f}{\partial y_k}=\left.\frac{\partial\psi}{\partial y}\right|_k-\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_{k+1}+\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_{k-1}=0\\
&\qquad\qquad (k=3,\ 4,\ \ldots,\ n-3),\\
&\frac{\partial f}{\partial y_{n-2}}=\left.\frac{\partial\psi}{\partial y}\right|_{n-2}-\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_{n-3}+\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_{n-1}+\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_{n}=0,\\
&\frac{\partial f}{\partial y_{n-1}}=\left.\frac{\partial\psi}{\partial y}\right|_{n-1}-\frac{2}{h}\left.\frac{\partial\psi}{\partial y'}\right|_{n}+\frac{1}{2h}\left.\frac{\partial\psi}{\partial y'}\right|_{n-2}=0,
\end{aligned}\right\}\tag{74.3}$$

где, например,

$$\left.\frac{\partial\psi}{\partial y}\right|_1=\frac{\partial\psi\left(x_1,\ y_1,\ \dfrac{y_2-y_0}{2h}\right)}{\partial y}.$$

Будем решать систему (74.3) методом дифференцирования по параметру. Введем в рассмотрение систему, содержащую параметр λ, $0\leqslant\lambda\leqslant 1$,

$$N_k(\lambda,\ y_{k-2},\ y_{k-1},\ y_k,\ y_{k+1},\ y_{k+2})=(1-\lambda)\Delta_k+\lambda\frac{\partial f}{\partial y_k}=0 \tag{74.4}$$
$$(k=1,\ 2,\ \ldots,\ n-1),$$

где $\Delta_k=\Delta_k(y_{k-1},\ y_k,\ y_{k+1})$ — выбранные нами линейные функции такие, что система

$$\left.\begin{aligned}
&\Delta_k(y_{k-1},\ y_k,\ y_{k+1})=0 \qquad (k=1,\ 2,\ \ldots,\ n-1),\\
&y_0=A, \qquad y_n=B
\end{aligned}\right\}\tag{74.5}$$

симметрична и имеет единственное решение $y_k=\bar{y}_k$ ($k=1,\ 2,\ \ldots,\ n-1$) при $\lambda=0$. Таким образом, если обозначить решение системы (74.4) через $y_k(\lambda)$ ($k=1,\ 2,\ \ldots,\ n-1$), то

$$y_k(0)=\bar{y}_k.$$

При $\lambda=1$ система (74.4) переходит в систему (74.3), решения которой мы ищем.

Продифференцируем систему (74.4) по λ. Имеем:

$$\sum_{l=-2}^{2}\frac{\partial N_k}{\partial y_{k+l}}\frac{dy_{k+l}}{d\lambda}=-\frac{\partial N_k}{\partial\lambda} \qquad (k=1,\ 2,\ \ldots,\ n-1), \tag{74.6}$$

или, короче,

$$A_{n-1}V_{n-1}=\Gamma_{n-1}, \tag{74.7}$$

где V_{n-1} и Γ_{n-1} — векторы, равные соответственно:

$$V_{n-1}=\left(\frac{dy_1}{d\lambda},\ \frac{dy_2}{d\lambda},\ \ldots,\ \frac{dy_{n-1}}{d\lambda}\right),$$

$$\Gamma_{n-1}=(\gamma_1,\ \gamma_2,\ \ldots,\ \gamma_{n-1}),\qquad \gamma_k=-\frac{\partial N_k}{\partial\lambda}\qquad (k=1,\ 2,\ \ldots,\ n-1),$$

а A_{n-1} — матрица системы (74.6). Как легко видеть, она симметричная пятидиагональная.

Обозначим:

$$\left.\begin{aligned}
b_1&=\frac{\partial N_1}{\partial y_1}=(1-\lambda)\frac{\partial\Delta_1}{\partial y_1}+\lambda\,(p_1+r_2+16r_0),\\
b_2&=\frac{\partial N_2}{\partial y_2}=(1-\lambda)\frac{\partial\Delta_2}{\partial y_2}+\lambda\,(p_2+r_3+r_1+r_0),\\
b_k&=\frac{\partial N_k}{\partial y_k}=(1-\lambda)\frac{\partial\Delta_k}{\partial y_k}+\lambda\,(p_k+r_{k+1}+r_{k-1})\\
&\qquad\qquad (k=3,\ 4,\ \ldots,\ n-3),\\
b_{n-2}&=\frac{\partial N_{n-2}}{\partial y_{n-2}}=(1-\lambda)\frac{\partial\Delta_{n-2}}{\partial y_{n-2}}+\lambda\,(p_{n-2}+r_{n-2}+r_{n-3}+r_n),\\
b_{n-1}&=\frac{\partial N_{n-1}}{\partial y_{n-1}}=(1-\lambda)\frac{\partial\Delta_{n-1}}{\partial y_{n-1}}+\lambda\,(p_{n-1}+16r_n+r_{n-2}),\\
a_1&=\frac{\partial N_1}{\partial y_2}=\frac{\partial N_2}{\partial y_1}=(1-\lambda)\frac{\partial\Delta_2}{\partial y_1}+\lambda\,(-q_2+q_1-4r_0),\\
a_k&=\frac{\partial N_k}{\partial y_{k+1}}=\frac{\partial N_{k+1}}{\partial y_k}=(1-\lambda)\frac{\partial\Delta_k}{\partial y_{k+1}}+\lambda(-q_{k+1}+q_k)\\
&\qquad\qquad (k=2,\ 3,\ \ldots,\ n-3),\\
a_{n-2}&=\frac{\partial N_{n-2}}{\partial y_{n-1}}=\frac{\partial N_{n-1}}{\partial y_{n-2}}=\\
&\qquad=(1-\lambda)\frac{\partial\Delta_{n-2}}{\partial y_{n-1}}+\lambda\,(q_{n-2}-q_{n-1}-4r_n),\\
c_k&=\frac{\partial N_{k-1}}{\partial y_{k+1}}=\frac{\partial N_{k+1}}{\partial y_{k-1}}=-\lambda r_k\qquad (k=2,\ 3,\ \ldots,\ n-2),
\end{aligned}\right\} \tag{74.8}$$

где на этот раз введены обозначения

$$r_k=\frac{1}{4h^2}\,\frac{\partial^2\psi}{\partial y'\,\partial y'}\bigg|_k,\qquad q_k=\frac{1}{2h}\,\frac{\partial^2\psi}{\partial y'\,\partial y}\bigg|_k,\qquad p_k=\frac{\partial^2\psi}{\partial y\,\partial y}\bigg|_k.$$

Тогда

$$A_{n-1}=\left\|\begin{array}{ccccccccccc}
b_1 & a_1 & c_2 & & & & & & & & \\
a_1 & b_2 & a_2 & c_3 & & & & & & & \\
c_2 & a_2 & b_3 & a_3 & c_4 & & & & & & \\
 & \cdot & \cdot & \cdot & \cdot & \cdot & & & & & \\
 & & \cdot & \cdot & \cdot & \cdot & \cdot & & & & \\
 & & & \cdot & \cdot & \cdot & \cdot & \cdot & & & \\
 & & & & c_{k-1} & a_{k-1} & b_k & a_k & c_{k+1} & & \\
 & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \\
 & & & & & & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & & & & & & & \cdot & \cdot & \cdot & \cdot \\
 & & & & & & c_{n-4} & a_{n-4} & b_{n-3} & a_{n-3} & c_{n-2} \\
 & & & & & & & c_{n-3} & a_{n-3} & b_{n-2} & a_{n-2} \\
 & & & & & & & & c_{n-2} & a_{n-2} & b_{n-1}
\end{array}\right\|. \tag{74.9}$$

Искомым решением системы (74.3) является решение задачи Коши при $\lambda=1$ для системы (74.6), удовлетворяющее условиям

$$y_k(0)=\overline{y}_k \qquad (k=1,\ 2,\ \ldots,\ n-1).$$

Это решение мы находим численно. Именно, разбиваем интервал $[0,\ 1]$ на m частичных интервалов точками λ_i $(i=1,\ 2,\ \ldots,\ m)$ и применяем, например, метод Рунге — Кутта.

Для системы дифференциальных уравнений вида

$$\frac{dy_l}{d\lambda}=Z_l(\lambda,\ y_1,\ y_2,\ \ldots,\ y_{n-1}) \qquad (l=1,\ 2,\ \ldots,\ n-1) \tag{74.10}$$

известные формулы Рунге — Кутта принимают вид:

$$y_l(\lambda_{i+1})=y_l(\lambda_i)+\frac{1}{6}\left[k_1^{(li)}+2k_2^{(li)}+2k_3^{(li)}+k_4^{(li)}\right],$$

где

$$\left.\begin{aligned}
k_1^{(li)}&=\Delta\lambda\cdot Z_l\left[\lambda_i,\ y_1(\lambda_i),\ y_2(\lambda_i),\ \ldots,\ y_{n-1}(\lambda_i)\right],\\
k_2^{(li)}&=\Delta\lambda\cdot Z_l\left[\lambda_i+\frac{\Delta\lambda}{2},\ y_1(\lambda_i)+\frac{k_1^{(1i)}}{2},\ \ldots\right.\\
&\qquad\qquad\left.\ldots,\ y_{n-1}(\lambda_i)+\frac{k_1^{(n-1,\ i)}}{2}\right],\\
k_3^{(li)}&=\Delta\lambda\cdot Z_l\left[\lambda_i+\frac{\Delta\lambda}{2},\ y_1(\lambda_i)+\frac{k_2^{(1i)}}{2},\ \ldots\right.\\
&\qquad\qquad\left.\ldots,\ y_{n-1}(\lambda_i)+\frac{k_2^{(n-1,\ i)}}{2}\right],\\
k_4^{(li)}&=\Delta\lambda\cdot Z_l\left[\lambda_i+\Delta\lambda,\ y_1(\lambda_i)+k_3^{(1i)},\ \ldots\right.\\
&\qquad\qquad\left.\ldots,\ y_{n-1}(\lambda_i)+k_3^{(n-1,\ i)}\right],\\
&(l=1,\ 2,\ \ldots,\ n-1;\ l=1,\ 2,\ \ldots,\ n-1).
\end{aligned}\right\} \quad 74.11)$$

Мы видим, что для использования формул (74.11) нам нужно представить (74.6) в разрешенном относительно производных виде (74.10). Однако разрешить (74.6) в общем виде весьма затруднительно.

При численном интегрировании можно обойти эту трудность: мы будем при каждом фиксированном λ_i разрешать числовую систему, полученную из (74.6) после подстановки в коэффициенты при неизвестных и правые части значений $y_k(\lambda_i)$, найденных ранее.

При решении полученной числовой системы мы используем разложение матрицы A_{n-1} в произведение двух треугольных [1]).

Так как матрица A_{n-1} является пятидиагональной, то

$$A_{n-1} = C_{n-1} \cdot B_{n-1}, \tag{74.12}$$

где

$$C_{n-1} = \left\| \begin{array}{ccccccc} \beta_1 & & & & & & \\ \alpha_1 & \beta_2 & & & & & \\ \delta_2 & \alpha_2 & \beta_3 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \delta_{k-1} & \alpha_{k-1} & \beta_k & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \delta_{n-2} & \alpha_{n-2} & \beta_{n-1} \end{array} \right\|,$$

$$B_{n-1} = \left\| \begin{array}{ccccc} 1 & \mu_1 & \nu_1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 \quad \mu_k & \nu_k & \\ & & \ddots & \ddots & \ddots \\ & & 1 & \mu_{n-3} & \nu_{n-3} \\ & & & 1 & \mu_{n-2} \\ & & & & 1 \end{array} \right\|.$$

[1]) См., например, Ф. Р. Гантмахер [1] или Д. К. Фаддеев и В. Н. Фаддеева [1].

Элементы матриц C_{n-1} и B_{n-1} определяются рекуррентно из соотношений:

$$\beta_1 = b_1, \quad \mu_1 = \frac{a_1}{\beta_1}, \quad \alpha_1 = a_1, \quad \beta_2 = b_2 - \alpha_2\mu_1,$$

$$\nu_k = \frac{c_{k+1}}{\beta_k} \qquad (k = 1,\ 2,\ \ldots,\ n-3),$$

$$\mu_k = \frac{a_k - \alpha_{k-1}\nu_{k-1}}{\beta_k} \qquad (k = 2,\ 3,\ \ldots,\ n-2),$$

$$\alpha_{k-1} = a_{k-1} - \delta_{k-1}\mu_{k-2} \qquad (k = 3,\ 4,\ \ldots,\ n-1),$$

$$\beta_k = b_k - \delta_k\nu_{k-2} - \alpha_{k-1}\mu_{k-1} \qquad (k = 3,\ 4,\ \ldots,\ n-1),$$

$$\delta_k = c_k \qquad (k = 2,\ 3,\ \ldots,\ n-2).$$

На основании формул (74.7) и (74.12):

$$C_{n-1}B_{n-1}V_{n-1} = \Gamma_{n-1}.$$

Обозначим

$$B_{n-1}V_{n-1} = D_{n-1}, \tag{74.13}$$

тогда уравнение (74.7) можно записать в виде

$$C_{n-1}D_{n-1} = \Gamma_{n-1}.$$

Компоненты вектора $D_{n-1} = (d_1,\ d_2,\ \ldots,\ d_{n-1})$ находим последовательно по формулам

$$d_1 = \frac{\gamma_1}{\beta_1}, \qquad d_2 = \frac{\gamma_2 - \alpha_1 d_1}{\beta_2},$$

$$d_k = \frac{\gamma_k - \delta_{k-1}d_{k-2} - \alpha_{k-1}d_{k-1}}{\beta_k} \qquad (k = 3,\ 4,\ \ldots,\ n-1).$$

Теперь из соотношения (74.13) определяем компоненты вектора V_n:

$$\frac{dy_{n-1}}{d\lambda} = d_{n-1},$$

$$\frac{dy_{n-2}}{d\lambda} = d_{n-2} - \mu_{n-2}\frac{dy_{n-1}}{d\lambda},$$

. . . .

$$\frac{dy_k}{d\lambda} = d_k - \mu_k\frac{dy_{k+1}}{d\lambda} - \nu_k\frac{dy_{k+2}}{d\lambda} \qquad (k = n-3,\ n-4,\ \ldots,\ 1),$$

что соответствует определению функций $Z_l(\lambda,\ y_1,\ y_2,\ \ldots,\ y_{n-1})$ для $\lambda = \lambda_l$.

Рассмотрим тот же пример, что и в § 71.

Пусть требуется определить функцию, минимизирующую интеграл

$$\int_0^1 \left[\frac{1}{48}y'^4 + y'^2 + y^2 - 6y\right]dx$$

и удовлетворяющую граничным условиям

$$y(0)=1, \qquad y(1)=0.$$

В качестве квадратурной формулы выберем формулу левых прямоугольников, т. е. положим

$$C_k=\frac{1}{n}h \qquad (k=0,\ 1,\ \ldots,\ n-1), \qquad C_n=0.$$

Вычисляем p_k, q_k, r_k:

$$p_k=\frac{2}{n}, \qquad q_k=0, \qquad r_k=\frac{n}{4}\left[2+\frac{1}{4}(\bar{y}_k)^2\right] \qquad (k=0,\ 1,\ \ldots,\ n-1).$$

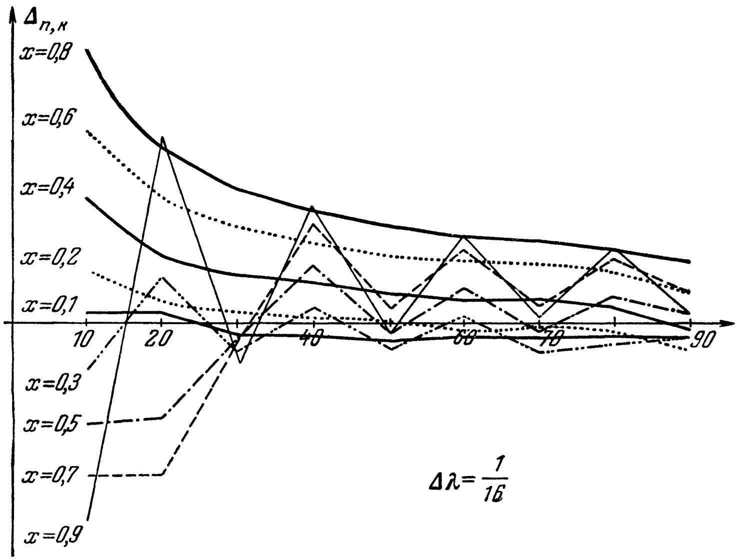

Рис. 6.

Входящую в систему уравнений (74.4) функцию Δ_k выбираем равной

$$\Delta_k=-R_ky_{k+1}+(R_k+R_{k-1}+T_k)\,y_k-R_{k-1}y_{k-1},$$

где

$$R_k=\frac{n-k}{4h^3}, \qquad T_k=\frac{1}{4h^2(1-kh)}.$$

Решением соответствующей системы (74.5) является функция

$$y_k(0)=\bar{y}_k=1-kh \qquad (k=1,\ 2,\ \ldots,\ n-1).$$

Далее осуществляем указанную выше схему.

В таблицах 74.1—74.3 и на рис. 6 мы приводим результаты счета.

Таблица 74.1

Приближенные решения $y_n(x_k)$, найденные методом дифференцирования по параметру при различных n $\left(\text{при } \Delta\lambda = \frac{1}{16}\right)$

№ п/п	x_k	Точное решение	Приближенное решение при								
			$n=10$	$n=20$	$n=30$	$n=40$	$n=50$	$n=60$	$n=70$	$n=80$	$n=90$
1	0,1	0,99	0,9925	0,9925	0,9865	0,9898	0,9864	0,9882	0,9859	0,9879	0,9865
2	0,2	0,96	0,9738	0,9664	0,9632	0,9614	0,9606	0,9587	0,9596	0,9575	0,9529
3	0,3	0,91	0,8977	0,9218	0,9034	0,9148	0,9049	0,9113	0,9033	0,9095	0,9065
4	0,4	0,84	0,8723	0,8584	0,8529	0,8497	0,8480	0,8453	0,8469	0,8430	0,8369
5	0,5	0,75	0,7240	0,7756	0,7451	0,7654	0,7489	0,7604	0,7483	0,7579	0,7527
6	0,6	0,65	0,6914	0,6730	0,6656	0,6612	0,6586	0,6558	0,6557	0,6531	0,6471
7	0,7	0,51	0,4700	0,5497	0,5063	0,5363	0,5126	0,5303	0,5138	0,5275	0,5173
8	0,8	0,36	0,4297	0,4051	0,3953	0,3895	0,3859	0,3829	0,3818	0,3797	0,3758
9	0,9	0,19	0,1374	0,2384	0,1795	0,2197	0,1877	0,2123	0,1905	0,2086	0,1928
10		—2,0667	—2,3318	—2,2034	—2,1576	—2,1345	—2,1207	—2,1115	—2,1048	—2,0998	—2,0959

Таблица 74.2

Таблица максимальных отклонений приближенного решения от точного при различных $\Delta\lambda$

$\Delta\lambda$	$n=10$	$n=20$	$n=30$	$n=40$	$n=50$	$n=60$	$n=70$	$n=80$
$\frac{1}{16}$	0,0697	0,0484	0,0353	0,0297	0,0259	0,0223	0,0218	0,0197
$\frac{1}{32}$	0,0665	0,0444	0,0337	0,0288	0,0255	0,0232	0,0204	0,0191
$\frac{1}{64}$	0,0657	0,0336	0,0310	0,0275	0,0248	0,0225	0,0226	0,0182
$\frac{1}{128}$	0,0657	0,0331	0,0272	0,0250	0,0232	0,0216	0,0222	0,0175

В таблице 74.1 указаны значения приближенных решений $y_n(x_k)$, полученных интегрированием систем (74.6) при различных n с одним и тем же шагом $\Delta\lambda = \frac{1}{16}$.

По рис. 6 можно проследить за отклонением приближенного решения $y_k(x_k)$ от точного $y(x_k)$ $\left(\Delta\lambda = \frac{1}{16}\right)$. На оси абсцисс отложено n, на оси ординат $\Delta_{n,k} = y_n(x_k) - y(x_k)$. Для каждого x_k построена своя кривая $\left(x_k = kh,\ k = 1,\ 2,\ \ldots,\ 9,\ h = \frac{1}{10}\right)$.

Из таблицы и рисунка видно, что при увеличении n наблюдается лишь незначительное уменьшение $\Delta_{n,k}$.

Уменьшение шага интегрирования $\Delta\lambda$ также не приводит к качественному улучшению результатов, что можно проследить по табл. 74.2, в которой приведены максимальные по k отклонения $|\Delta_{n,k}|$ при различных n и λ.

Для достижения большей точности в решении системы (74.3) был применен метод Ньютона — Канторовича, в котором за начальное приближение бралось вычисленное методом дифференцирования по параметру приближенное решение системы (74.3).

В таблице 74.3 приведен результат вычисления решения системы (74.2) для приведенного примера методом Ньютона — Канторовича. За начальные приближения были взяты значения из табл. 74.1 при соответствующем n.

Из таблицы 74 3 видно, что уже второе приближение по методу Ньютона мало отличается от точного решения задачи; наибольшая относительная погрешность не превосходит 0,3%.

Приближенное значение минимума интеграла мы можем подсчитать, подставляя в формулу (74.2) вычисленные значения y_k.

Таблица 74.3

Применение метода Ньютона

№ п/п	x_k	Точное решение	Приближенные значения $y_n(x_k)$, найденные по методу Ньютона — Канторовича					
			$n=10$		$n=20$		$n=30$	
			I	II	I	II	I	II
1	0,1	0,99	0,9970	0,9901	0,99802	0,99011	0,9908	0,98996
2	0,2	0,96	0,9741	0,9601	0,97594	0,96021	0,9616	0,95994
3	0,3	0,91	0,9303	0,9102	0,93339	0,91031	0,9124	0,90993
4	0,4	0,84	0,8666	0.8402	0,86974	0,84040	0,8430	0,83992
5	0,5	0,75	0,7781	0,7503	0,78418	0,75050	0,7535	0,74992
6	0,6	0,64	0,6713	0,6404	0,67573	0,64062	0,6435	0,63993
7	0,7	0,51	0,5327	0,5105	0,54335	0,51072	0,5131	0,50994
8	0,8	0,36	0,3790	0,3606	0,38547	0,36072	0,3620	0,35996
9	0,9	0,19	0,1869	0,1899	0,20105	0,19047	0,1893	0,18997
10		—2,0667	—2,3001	—2,3105		—2,2034	—2,3104	—2,3105
11			—2,0813	—2,0813		—2,0799		

№ п/п	x_k	Приближенные значения $y_n(x_k)$, найденные по методу Ньютона — Канторовича					
		$n=40$		$n=60$		$n=80$	
		I	II	I	II	I	II
1	0,1	0,99857	0,99014	0,99854	0,99014	0,99756	0,99013
2	0,2	0,97698	0,96027	0,97690	0,96027	0,97501	0,96025
3	0,3	0,93488	0,91038	0,93477	0,91038	0,93223	0,91036
4	0,4	0,87168	0,84049	0,87155	0,84049	0,86881	0,84046
5	0,5	0,78664	0,75061	0,78660	0,75061	0,78421	0,75057
6	0,6	0,67880	0,64074	0,67896	0,64073	0,67732	0,64069
7	0,7	0,54703	0,51087	0,54746	0,51085	0,54683	0,51081
8	0,8	0,39002	0,36088	0,39087	0,36089	0,39117	0,36086
9	0,9	0,20636	0,19065	0,20773	0,19067	0,20865	0,19068
10			—2,1346		—2,1115		—2,0999
11			—2,0745		—2,07213		—2,0708

В таблицах 74.1 и 74.3 в строчках 10 приведены результаты вычисления суммы (74.2) для случаев, когда приближенное решение находится методом дифференцирования по параметру и с уточнением по методу Ньютона — Канторовича.

Сравнивая полученные значения суммы (74.2) с точным значением интеграла (равным —2,0667), мы видим, что относительная погрешность велика (12,5%—1,3%). При этом указанная погрешность не уменьшилась и в случае, когда мы при вычислении суммы (74.2) использовали более точные значения y_k из табл. 74.3.

Это объясняется тем, что в сумму (74.2) входят приближенные значения производных, а точность, достаточная для нахождения функции, оказалась недостаточной для нахождения производной.

Для уточнения приближенного значения минимума интеграла можно, применяя, например, интерполяционную формулу Лагранжа, использовать при подсчете входящих в квадратурную формулу производных найденные значения функции в большем числе точек, чем мы это делали по формулам (74.1).

В примере мы использовали четырехточечную замену, что дало существенные улучшения. Например, для $n=10$ относительная погрешность при подсчете минимума интеграла уменьшилась с 12,5% до 0,75% (табл. 74.3).

§ 75. Метод Л. М. Качанова

Метод решения нелинейных систем Ритца, о котором пойдет речь в настоящем параграфе, был предложен Л. М. Качановым в его статье [4] для решения некоторых задач теории пластичности и применен в статье [6] к задаче об упруго-пластическом кручении стержня квадратного сечения. С. Н. Розе [1] уточнил вычисления статьи [6] Л. М. Качанова[1]) и показал, что его прием можно применять к общим функционалам вида (68.1). Строгого обоснования этого приема пока нет[2]); однако он заслуживает внимания благодаря своей простоте и подтверждаемой вычислениями эффективности.

Для функционала (68.1) построим систему Ритца. Она имеет вид.

$$\frac{\partial F(u_n)}{\partial a_j}=0 \qquad (j=1,\,2,\,\ldots,\,n); \tag{75.1}$$

здесь F — функционал (68.1),

$$u_n=\sum_{k=1}^{n} a_k\varphi_k \tag{75.2}$$

---

[1]) Результаты вычислений, выполненных С. Н. Розе, приведены также в книге автора и Х. Л. Смолицкого [1].

[2]) Доказательство сходимости, данное С. Н. Розе [1], к сожалению, содержит ошибку.

и φ_k — координатные элементы. По методу Л. М. Качанова поступаем следующим образом. В функционале (68.1) заменяем функции $g_j(\xi)$ постоянными $g_j(0)$ и ставим задачу о минимуме полученного таким образом квадратичного функционала

$$F_1(u)=\int\limits_{\Omega}\sum_{j=1}^{s} g_j(0)\,\tau_j(u)\,dx-(f,\ u). \tag{75.3}$$

Отыскивая приближенное решение этой задачи по методу Ритца и полагая

$$u_{n1}=\sum_{k=1}^{n} a_{k1}\varphi_k, \tag{75.4}$$

мы сможем определить коэффициенты a_{k1}, решая некоторую линейную алгебраическую систему. Совокупность коэффициентов $(a_{11}, a_{21}, \ldots, a_{n1})$ будем рассматривать как первое приближение к решению системы Ритца (75.1). Вообще, если построено r-е приближение

$$(a_{1r},\ a_{2r},\ \ldots,\ a_{nr}) \tag{75.5}$$

к решению системы (75.1), то $(r+1)$-е приближение

$$(a_{1,\,r+1},\ a_{2,\,r+1},\ \ldots,\ a_{n,\,r+1})$$

строится так: в функционале (68.1) заменяем $g_j(\xi)$ выражением

$$g_j(\tau_j(u_{nr}))=g_j\left(\tau_j\left(\sum_{k=1}^{n} a_{kr}\varphi_k\right)\right),$$

минимум полученного квадратичного функционала

$$F_k(u)=\int\limits_{\Omega}\sum_{j=1}^{s} g_j(\tau_j(u_{nr}))\,\tau_j(u)\,dx-(f,\ u)$$

ищем, применяя процесс Ритца и полагая

$$u_{n,\,r+1}=\sum_{k=1}^{n} a_{k,\,r+1}\varphi_k,$$

при этом коэффициенты $a_{k,\,r+1}$ определяются из линейной алгебраической системы.

Если производная Гато от $\operatorname{grad} F$ равномерно положительно определена, то множество приближенных решений (75.5), соответствующее значениям $r=1, 2, 3, \ldots$, ограничено и, следовательно, компактно. Если это множество — сходящееся, то его предел есть решение системы (75.1).

На практике качановский вычислительный процесс следует вести до стабилизации или до выявления его нестабильности.

§ 76. Об устойчивости процесса Ритца для нелинейных задач

В настоящем параграфе мы даем определение устойчивости процесса Ритца для нелинейных задач и приводим некоторые (далеко не полные) результаты о достаточных условиях устойчивости.

Пусть функционал F, минимум которого нужно найти, удовлетворяет условиям теоремы 70.1. Допустим еще, что производная P'_u, где $P = \operatorname{grad} F$, удовлетворяет неравенствам (66.2) и (66.15), так что

$$(P'_u h, h) \geqslant \gamma_1^2 (P'_{u_0} h, h) \geqslant \gamma^2 \| h \|^2; \qquad (76.1)$$

здесь γ и γ_1 — положительные постоянные, u_0, u, h — элементы области определенйя оператора P, причем элемент u_0 фиксирован, а элементы u и h произвольны. Как и всюду в этой главе, мы принимаем, что функционал F определен на линейном множестве, плотном в банаховом пространстве B. Тогда операторы P и P' действуют из B и B^*; пусть области их определения плотны в B.

Устойчивость процесса Ритца мы изучим при следующем довольно жестком ограничении: мы примем, что вторая производная градиента $P = \operatorname{grad} F$ ограничена в любом шаре; это означает существование такой функции D_t неотрицательной переменной t, что если $\| u \| \leqslant t$, то $\| P''_u \| \leqslant D_t$.

Примем, что координатная система $\{\varphi_n\}$ удовлетворяет условиям 2) и 3) § 70; условие 1) § 70 заменим более сильным: 1′) $\varphi_n \in D(P)$. Тогда уравнения Ритца можно записать в форме Бубнова — Галёркина (70.5):

$$\left(P \left(\sum_{k=1}^{n} a_k^{(n)} \varphi_k \right), \varphi_j \right) = 0 \qquad (j = 1, 2, \ldots, n).$$

Обозначим через $a^{(n)}$ вектор с составляющими $a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}$. Положим еще:

$$\left(P \left(\sum_{k=1}^{n} a_k^{(n)} \varphi_k \right), \varphi_j \right) = P_j^{(n)} \left(a^{(n)} \right)$$

и обозначим вектор с составляющими

$$P_1^{(n)} \left(a^{(n)} \right), \; P_2^{(n)} \left(a^{(n)} \right), \; \ldots, \; P_n^{(n)} \left(a^{(n)} \right)$$

через $P^{(n)} \left(a^{(n)} \right)$. Уравнения Ритца можно теперь записать так:

$$P^{(n)} \left(a^{(n)} \right) = 0. \qquad (76.2)$$

Составление вектора $P^{(n)}$ сопряжено с погрешностью вычислений. Пусть на самом деле мы вместо вектора $P^{(n)}$ получили в результате вычислений вектор

$$Q^{(n)} = P^{(n)} + T^{(n)}. \qquad (76.3)$$

Будем считать, что при приближенном вычислении вектора $P^{(n)}$ мы в состоянии удовлетворить следующему требованию: существуют такие три функции неотрицательного аргумента t — обозначим их через C_t, C'_t, C''_t, что в шаре $\|a^{(n)}\| \leqslant t$ верны неравенства

$$\left.\begin{array}{c} \|T^{(n)}(a^{(n)})\| \leqslant \delta C_t, \qquad \|T^{(n)\prime}(a^{(n)})\| \leqslant \delta C'_t, \\ \|T^{(n)\prime\prime}(a^{(n)})\| \leqslant \delta C''_t; \end{array}\right\} \tag{76.4}$$

здесь δ — величина, характеризующая точность вычислений вектора $P^{(n)}$; будем считать, что ее можно сделать сколь угодно малой. В качестве нормы в (76.4) будем брать евклидову норму вектора.

Применяя процесс Ритца, мы на самом деле будем решать не уравнение (76.2), а уравнение

$$Q^{(n)}(b^{(n)}) = P^{(n)}(b^{(n)}) + T^{(n)}(b^{(n)}) = 0. \tag{76.5}$$

Будем говорить, что процесс Ритца устойчив, если по данному $\varepsilon > 0$ можно найти такое $\delta > 0$, что при выполнении условий (76.4) уравнение (76.5) имеет решение $b^{(n)}$ такое, что $\|b^{(n)} - a^{(n)}\| < \varepsilon$; здесь $a^{(n)}$ — решение уравнения (76.2).

Теорема 76.1. *Если функционал F и его градиент удовлетворяют условиям, перечисленным в настоящем параграфе, а координатная система почти ортонормирована в энергетической метрике оператора P'_{u_0}, то процесс Ритца устойчив.*

Докажем прежде всего, что в условиях теоремы величины $\|a^{(n)}\|$ ограничены в совокупности. Обозначим $P(0) = f$. Уравнения (70.5) запишем в виде

$$(P(u^{(n)}) - P(0), \varphi_j) = (f, \varphi_j) \quad (j = 1, 2, \ldots, n), \qquad u^{(n)} = \sum_{k=1}^{n} a_n^{(k)} \varphi_k.$$

Но

$$P(u^{(n)}) - P(0) = \int_0^1 P'_{tu_n} u^{(n)}\, dt,$$

отсюда

$$\int_0^1 \left(P'_{tu^{(n)}} u^{(n)}, \varphi_j\right) dt = (f, \varphi_j) \qquad (j = 1, 2, \ldots, n).$$

Умножив на $a_j^{(n)}$ и сложив, получим:

$$\int_0^1 \left(P'_{tu^{(n)}} u^{(n)}, u^{(n)}\right) dt = (f, u^{(n)}).$$

Слева подынтегральную функцию заменим меньшей величиной $\gamma_1^2\left(P'_{u_0}u^{(n)},\ u^{(n)}\right)$; правую часть заменим большей величиной $\|f\|\,\|u^{(n)}\|$. Воспользуемся обозначениями $[\ ,\ ]$ и $|\ |$ для энергетического произведения и энергетической нормы оператора P'_{u_0}. Мы получим тогда

$$\gamma_1^2|u^{(n)}|^2 \leqslant \|f\|\,\|u^{(n)}\|.$$

Но по неравенству (76.1)

$$\|u^{(n)}\| \leqslant \frac{\gamma_1}{\gamma}|u^{(n)}|,$$

и мы приходим к соотношению

$$|u^{(n)}| \leqslant \frac{\|f\|}{\gamma\gamma_1}. \tag{76.6}$$

Далее,

$$|u^{(n)}|^2 = \left(R_n a^{(n)},\ a^{(n)}\right) \geqslant \lambda_1^{(n)}\|a^{(n)}\|^2,$$

где R_n — матрица элементов $[\varphi_k,\ \varphi_j]$, а $\lambda_1^{(n)}$ — ее наименьшее собственное число. В силу условий теоремы существует такая постоянная $\lambda_0 > 0$, что $\lambda_1^{(n)} \geqslant \lambda_0$. Теперь

$$\|a^{(n)}\| \leqslant \frac{1}{\sqrt{\lambda_0}}|u^{(n)}| \leqslant \frac{\|f\|}{\gamma\gamma_1\sqrt{\lambda_0}} = \sigma = \mathrm{const}, \tag{76.7}$$

что и требовалось доказать. Из неравенств (76.4) и (76.7) следует, что

$$\left.\begin{gathered}\|T^{(n)}(a^{(n)})\| \leqslant \delta C_\sigma,\quad \|T^{(n)'}(a^{(n)})\| \leqslant \delta C'_\sigma,\\ \|T^{(n)''}(a^{(n)})\| \leqslant \delta C''_\sigma.\end{gathered}\right\} \tag{76.8}$$

Будем решать систему (76.5) по методу Ньютона—Канторовича; за начальное приближение примем вектор $a^{(n)}$ — решение системы (76.2). Проверим выполнение условий применимости названного метода[1]).

Нетрудно убедиться, что производная $P^{(n)'}\left(a^{(n)}\right)$ определяется матрицей элементов $\left(P'_{u^{(n)}}\varphi_k,\ \varphi_j\right)$ $(k,\ j = 1,\ 2,\ \ldots,\ n)$. В силу первого из неравенств (76.1) наименьшее собственное число этой матрицы ограничено снизу величиной $\gamma_1^2\lambda_1^{(n)} \geqslant \gamma_1^2\lambda_0$. Но тогда

$$\left\|\left[P^{(n)'}(a^{(n)})\right]^{-1}\right\| \leqslant \frac{1}{\gamma_1^2\lambda_0}.$$

---

[1]) См. Л. В. Канторович и Г. П. Акилов [1], стр 642.

Это дает нам следующие неравенства:

$$\left.\begin{aligned} &\left\|\left[P^{(n)'}\left(a^{(n)}\right)\right]^{-1} T^{(n)}\left(a^{(n)}\right)\right\| \leqslant \frac{\delta C_\sigma}{\gamma_1^2\lambda_0}, \\ &\left\|\left[P^{(n')}\left(a^{(n)}\right)\right]^{-1} T^{(n)'}\left(a^{(n)}\right)\right\| \leqslant \frac{\delta C'_\sigma}{\gamma_1^2\lambda_0}, \\ &\left\|\left[P^{(n)'}\left(a^{(n)}\right)\right]^{-1} T^{(n)''}\left(\tilde{a}^{(n)}\right)\right\| \leqslant \frac{\delta C''_{2\sigma}}{\gamma_1^2\lambda_0}, \qquad \left\|\tilde{a}^{(n)}\right\| \leqslant 2\sigma. \end{aligned}\right\} \tag{76.9}$$

Чтобы получить последнее из неравенств Л. В. Канторовича, оценим вторую производную $P^{(n)''}\left(\tilde{a}^{(n)}\right)$ в шаре радиуса 2σ; здесь $\tilde{a}^{(n)}$—вектор

$$\tilde{a}^{(n)} = (\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_n).$$

Обозначим

$$w_n = \sum_{k=1}^{n} \tilde{a}_k \varphi_k.$$

Тогда, как легко видеть, $P^{(n)''}\left(\tilde{a}^{(n)}\right)$ определяется матрицей с тремя входами, элементы которой суть

$$\left(P''_{w_n}\varphi_l\varphi_k, \varphi_j\right) \qquad (j, k, l = 1, 2, \ldots, n).$$

Введем в рассмотрение векторы $c^{(n)}$, $d^{(n)}$, $e^{(n)}$, составляющие которых суть $c_j^{(n)}$, $d_j^{(n)}$, $e_j^{(n)}$ $(j = 1, 2, \ldots, n)$. Обозначим

$$q_n = \sum_{k=1}^{n} c_k^{(n)}\varphi_k, \qquad r_n = \sum_{k=1}^{n} d_k^{(n)}\varphi_k, \qquad s_n = \sum_{k=1}^{n} e_k^{(n)}\varphi_k.$$

Оценим модуль скалярного произведения

$$\left|\left(P^{(n)}\left(\tilde{a}^{(n)}\right) c^{(n)} d^{(n)}, e^{(n)}\right)\right| = \left|\left(P''_{w_n} q_n r_n, s_n\right)\right| \leqslant$$
$$\leqslant \left\|P''_{w_n}\right\| \|q_n\| \|r_n\| \|s_n\| \leqslant D_{2\sigma} \|q_n\| \|r_n\| \|s_n\|.$$

Имеем:

$$|q_n|^2 = (R_n c^{(n)}, c^{(n)}) \leqslant \Lambda_0 \|c^{(n)}\|^2,$$

где Λ_0 — верхняя граница собственных чисел матрицы R_n; такая граница существует в силу условий теоремы. Аналогичные неравенства верны для r_n и s_n. Теперь, используя неравенство (76.1), найдем:

$$\left|\left(P^{(n)''}\left(\tilde{a}^{(n)}\right) c^{(n)} d^{(n)}, e^{(n)}\right)\right| \leqslant \frac{\gamma^3 \Lambda_0^{\frac{3}{2}} D_{2\sigma}}{\gamma_1^3} \|c^{(n)}\| \|d^{(n)}\| \|e^{(n)}\|.$$

Положив здесь

$$e_n = P^{(n)'}\left(\tilde{a}^{(n)}\right) c^{(n)} d^{(n)},$$

найдем, что

$$\left\| P^{(n)''}\left(\tilde{a}^{(n)}\right)\right\| \leqslant \frac{\gamma^3 \Lambda_0^{\frac{3}{2}} D_{2\sigma}}{\gamma_1^3}, \qquad \left\| \tilde{a}^{(n)} \right\| \leqslant 2\sigma.$$

Последнее из неравенств Л. В. Канторовича принимает следующий вид:

$$\left\| [P^{(n)'}(a^{(n)})]^{-1} P^{(n)''}(\tilde{a}^{(n)}) \right\| \leqslant \frac{\gamma^3 \Lambda_0^{\frac{3}{2}} D_{2\sigma}}{\gamma_1^5 \lambda_0}, \qquad \| \tilde{a}^{(n)} \| \leqslant 2\sigma. \quad (76.10)$$

В силу неравенств (76.9) и (76.10) имеем (в обозначениях книги Л. В. Канторовича и Г. П. Акилова [1], стр. 642):

$$\eta = \frac{\delta C_\sigma}{\gamma_1^2 \lambda_0}, \qquad \alpha = \frac{\delta C'_\sigma}{\gamma_1^2 \lambda_0}, \qquad L = \frac{\delta C''_{2\sigma}}{\gamma_1^2 \lambda_0}, \qquad K = \frac{\gamma^3 \Lambda_0^{\frac{3}{2}} D_{2\sigma}}{\gamma_1^5 \lambda_0}.$$

Сохраняя обозначения только что процитированной книги, положим

$$h = \frac{\eta (K + L)}{(1 - \alpha)^2}$$

и выберем δ столь малым, чтобы выполнялись неравенства

$$h \leqslant \frac{1}{2}, \qquad \frac{1 - \sqrt{1 - 2h}}{h} \frac{\eta}{1 - \alpha} < \sigma, \qquad \frac{1 - \sqrt{1 - 2h}}{h} \frac{1}{1 - \alpha} < \varepsilon.$$

Тогда в шаре $\left\| b^{(n)} - a^{(n)} \right\| < \varepsilon$ система (76.5) имеет решение, и притом единственное. Теорема доказана.

Обозначим через v_n «неточное» приближение по Ритцу

$$v_n = \sum_{k=1}^{n} b_k^{(n)} \varphi_k, \qquad (76.11)$$

где $b^{(n)} = \left(b_1^{(n)}, b_2^{(n)}, \ldots, b_n^{(n)}\right)$ — решение системы (76.5). Будем говорить, что процесс вычисления приближенного решения по Ритцу устойчив, если по данному $\varepsilon > 0$ можно так подобрать число $\delta > 0$, входящее в неравенства (76.4), что $| v_n - u_n | < \varepsilon$. Нетрудно видеть, что в условиях теоремы 76.1 упомянутый процесс устойчив Действительно,

$$| v_n - u_n |^2 = (R_n (b^{(n)} - a^{(n)}), \; b^{(n)} - a^{(n)}) \leqslant \Lambda_0 \| b^{(n)} - a^{(n)} \|^2$$

и достаточно выбрать δ так, чтобы было

$$\| b^{(n)} - a^{(n)} \| < \frac{\varepsilon}{\sqrt{\Lambda_0}}.$$

Рассмотрим следующий пример. Пусть в квадрате $0 \leqslant x,\ s \leqslant 1$ ядро $K(x, s)$ представляет собой вторую итерацию ядра $L(x, s)$:

$$K(x, s) = \int_0^1 L(x, t)\, L(t, s)\, dt.$$

Относительно ядра $L(x, t)$ примем, что оно вещественное, симметричное, непрерывное, неотрицательное и положительно определенное (последнее, как известно, означает, что все его характеристические числа положительны); очевидно, этими же свойствами обладает и ядро $K(x, s)$. Ядро $L(x, s)$, очевидно, ограничено. Пусть $L(x, s) \leqslant M$, тогда одновременно $K(x, s) \leqslant M^2$.

Функционал

$$F(u) = \frac{1}{2} \int_0^1 u^2(x)\, dx + \frac{1}{4} \int_0^1 \int_0^1 K(x, s)\, u^2(x)\, u^2(s)\, dx\, ds - \\ - \int_0^1 f(x)\, u(x)\, dx, \qquad f \in L_2(0, 1), \tag{76.12}$$

рассматриваемый в вещественном пространстве $L_2(0, 1)$, имеет своей областью определения все это пространство. Действительно, если $u \in L_2(0, 1)$, то первый и третий интегралы в (76.12) имеют смысл. Второй интеграл также имеет смысл, потому что

$$\int_0^1 \int_0^1 K(x, s)\, u^2(x)\, u^2(s)\, dx\, ds \leqslant M \| u \|^4.$$

Вычислим оператор $P = \operatorname{grad} F$. Имеем:

$$\left. \frac{dF(u + \alpha h)}{d\alpha} \right|_{\alpha=0} = \int_0^1 u(x)\, h(x)\, dx + \\ + \int_0^1 h(x)\, dx \int_0^1 K(x, s)\, u(x)\, u^2(s)\, ds - \int_0^1 f(x)\, h(x)\, dx. \tag{76.13}$$

Правая часть в (76.13) есть линейный функционал над h, ограниченный в $L_2(0, 1)$. Чтобы убедиться в этом, достаточно проверить, что интеграл

$$\int_0^1 K(x, s)\, u^2(s)\, ds$$

ограничен, если $u \in L_2(0, 1)$. Но это очевидно, так как упомянутый интеграл не превосходит величины $M^2\|u\|^2$. Теперь ясно, что

$$Pu = u(x) + u(x)\int_0^1 K(x, s)\, u^2(s)\, ds - f(x). \tag{76.14}$$

Далее,

$$P'_u h = \frac{dP(u+\alpha h)}{d\alpha}\bigg|_{\alpha=0} = h(x) + h(x)\int_0^1 K(x, s)\, u^2(s)\, ds + \\ + 2u(x)\int_0^1 K(x, s)\, u(s)\, h(s)\, ds. \tag{76.15}$$

Отсюда

$$\left(P'_u h,\ h\right) = \|h\|^2 + \int_0^1\int_0^1 K(x, s)\, u^2(s)\, h^2(x)\, dx\, ds + \\ + 2\int_0^1\int_0^1 K(x, s)\, u(x)\, u(s)\, h(x)\, h(s)\, dx\, ds. \tag{76.16}$$

Оба интеграла справа имеют смысл, если u, $h \in L_2(0, 1)$. Действительно,

$$\int_0^1\int_0^1 K(x, s)\, u^2(x)\, u^2(s)\, dx\, ds \leqslant M^2\|u\|^4$$

и

$$\int_0^1\int_0^1 K(x, s)\, u(x)\, u(s)\, h(x)\, h(s)\, dx\, ds = \\ = \int_0^1\int_0^1\int_0^1 L(x, t)\, L(s, t)\, u(x)\, u(s)\, h(x)\, h(s)\, dx\, ds\, dt = \\ = \int_0^1 dt \left\{\int_0^1 L(x, t)\, u(x)\, h(x)\, dx\right\}^2;$$

внутренний интеграл имеет смысл и ограничен, так как

$$\left|\int_0^1 L(x, t)\, u(x)\, h(x)\, dx\right| \leqslant M\int_0^1 |u(x)|\,|h(x)|\, dx \leqslant M\|u\|\,\|h\|.$$

Интегралы в (76.16) также и неотрицательны, поэтому

$$\left(P'_u h,\ h\right) \geqslant \|h\|^2 \tag{76.17}$$

и оператор P'_u удовлетворяет неравенству (66.2). Та же формула (76.17) означает, что P'_u удовлетворяет и неравенству (66.15) при $u_0=0$; так как $P'_0=I$, то энергетическое пространство оператора P'_0 совпадает с $L_2(0,1)$.

Найдем еще вторую производную P''_u:

$$P''_u(h, h_1)=\frac{d}{d\alpha}P'_{u+\alpha h_1}h\Big|_{\alpha=0}=2\int_0^1 K(x, s)[u(s)h(x)h_1(s)+$$
$$+u(x)h(s)h_1(s)+u(s)h(s)\,|\,h_1(x)]\,ds. \qquad (76.18)$$

Докажем, что при любой фиксированной функции $u\in L_2(0, 1)$ билинейный оператор P''_u ограничен. Достаточно доказать это для каждого из трех слагаемых, на которые естественным образом распадается интеграл (76.18); эти слагаемые обозначим через $B_k(h, h_1)$ $(k=1, 2, 3)$.

Имеем:

$$\|B_1(h, h_1)\|\leqslant 2\|h\|\operatorname{Max}\int_0^1 K(x, s)\,|u(s)|\,|h_1(s)|\,ds\leqslant$$
$$\leqslant 2M^2\|h\|\int_0^1 |u(s)|\,|h_1(s)|\,ds\leqslant 2M^2\|u\|\,\|h\|\,\|h_1\|.$$

То же неравенство верно для B_2 и B_3. Но тогда

$$\|P''_u(h, h_1)\|\leqslant 6M^2\|u\|\,\|h\|\,\|h_1\|$$

и, следовательно, $\|P''_u\|\leqslant 6M^2\|u\|$; введенную в начале параграфа функцию D_t можно принять равной $6M^2t$.

Таким образом, выполнены все условия теоремы 76.1. Отсюда вытекает следующее.

Для задачи о минимуме функционала (76.12) или, что то же, для нелинейного интегрального уравнения

$$u(x)\left[1+\int_0^1 K(x, s)\,u^2(s)\,ds\right]=f(x) \qquad (76.19)$$

процесс Ритца устойчив, если ядро $K(x, s)$ удовлетворяет условиям настоящего параграфа и если координатная система почти ортонормирована в $L_2(0, 1)$.

ПРИЛОЖЕНИЕ

РЕАЛИЗАЦИЯ ПРОЦЕССА РИТЦА НА БЫСТРОДЕЙСТВУЮЩИХ ЭЛЕКТРОННЫХ ВЫЧИСЛИТЕЛЬНЫХ МАШИНАХ (ЭВМ)

Т. Н. СМИРНОВА

Постановка задачи

Проведение аналитических выкладок в большом числе — необходимая часть многих вычислительных процессов. Весьма трудоемкие аналитические выкладки встречаются прежде всего при использовании процесса Ритца и других процессов, ему аналогичных (Бубнова — Галёркина, наискорейшего спуска и т. п.), при приближенном решении задачи Коши по методу Пикара, Чаплыгина и другим аналитическим методам, при применении метода Ньютона — Канторовича и в ряде других случаев. Поэтому весьма актуальна задача проведения аналитических выкладок на ЭВМ. Каждый, кому приходилось применять приближенные аналитические методы решения краевых задач для дифференциальных уравнений, знает, насколько трудоемко проведение аналитических операций вручную. Во многих случаях эта работа носит стандартный характер, поэтому естественна мысль использовать машины при ее выполнении.

Это в свою очередь выдвигает задачу разработки практически удобных и универсальных методов введения такого рода заданий в машину и единых методов выполнения этих заданий.

Метод автоматизации программирования, разработанный Л. В. Канторовичем [1], [2] (см. также Л. В. Канторович и Л. Т. Петрова [1]) и его учениками, открывает один из возможных путей реализации аналитических выкладок на ЭВМ.

Этот метод можно назвать *методом крупноблочного программирования*. Л. В. Канторович предложил задавать машине вычислительный план в виде ряда строк (блоков), элементами которых могут быть 1) числовые многомерные величины, а также программы, схемы; 2) операции над этими величинами.

Расшифровывать и выполнять записанный в укрупненных элементах и операциях вычислительный план ЭВМ должна с помощью раз и навсегда составленной универсальной программы-робота «прораба».

Более детально суть метода, предложенного Л. В. Канторовичем, заключается в следующем:

1. Вводим в рассмотрение исходный класс A объектов, в котором выполняются вычисления. Такими классами, например, могут быть класс чисел, класс векторов и матриц, класс полиномов, класс «списков», класс программ и т. д.

2. Определяем допустимые операции P над объектами a из исходного класса A. Такими операциями могут быть, например, сложение чисел, скалярное умножение векторов, дифференцирование полиномов и т. д.

3. С помощью достаточно полной системы операций $\mathscr{P}=\{P\}$ любой алгорифм вычисления объекта $a \in A$ по заданным объектам $a_i \in A$ ($i=1, 2, \ldots, s$) записывается в виде схемы (вычислительного плана). Вычислительный план представляет собой последовательность строк вида

$$a^{(l)}=P\left(a_1^{(l)}, a_2^{(l)}, \ldots, a_s^{(l)}\right) \qquad (l=1, 2, \ldots, n),$$

причем $a_i^{(l)}$ ($i=1, 2, \ldots, s$) совпадает либо с одним из заданных объектов $a_i \in A$, либо с одним из объектов $a^{(k)}$ ($k=1, 2, \ldots, l-1$), построение которого определено одной из предыдущих строк.

Такой вычислительный план рассматривается как указание о последовательном построении объектов $a^{(1)}, a^{(2)}, \ldots, a^{(n)}$, что приводит к построению объекта, являющегося решением задачи.

4. Для каждого исходного класса объектов строится своя универсальная программа «прораб», которая автоматизирует выполнение вычислительных планов. Прораб должен читать строки в предписанном порядке и выполнять соответствующие операции над объектами. При построении прораба для объектов и операций разрабатывается своя система нумерации и представления (записи) в машинной памяти. Так, каждому объекту $a \in A$ сопоставляется его запись $[a]$ и справка о нем (a). Запись $[a]$ — это числовое содержание объекта. Например, пусть элементы матрицы, рассматриваемой как объект из класса многомерных величин

$$K=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ . & . & . & . \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

размещены соответственно в ячейках

$$\begin{gathered} P, \quad P+h_1, \quad P+2h_1, \ldots \\ P+h_2, \qquad P+h_2+h_1, \qquad P+h_2+2h_1, \ldots \end{gathered}$$

Это — *запись* матрицы K. *Справка* о матрице K представляет собой описание числового материала и его расположения. Она имеет следующий вид: $\begin{pmatrix} P & n & h_1 \\ & m & h_2 \end{pmatrix}$, где P — номер ячейки, в которой помещается первый элемент величины, n — число столбцов, m — число строк, h_1 — расстояние между двумя соседними элементами любой строки, h_2 — расстояние между первыми (и вообще i-ми) элементами двух соседних строк.

Записью операции является некоторая программа, практически осуществляющая эту операцию на машине. Справка об операции — это описание местоположения и размеров ее программы.

Класс многомерных величин (матриц, векторов) рассмотрен в качестве исходного в работах Л. Т. Петровой и М. А. Яковлевой [1], М. А. Яковлевой [1] и В. А. Булавского [1]. На этот класс опирается весьма широкий круг задач.

К. В. Шахбазян [1] ввела в рассмотрение класс программ и построила прораб, вычисляющий программу по заданному вычислительному плану.

Для осуществления на ЭВМ аналитических выкладок, направленного процесса преобразования исходного математического выражения в результативное Л. Т. Петровой [1] был рассмотрен в качестве исходного так называемый класс «списков». Специализированная программа прораб СП-III-59, составленная Л. Т. Петровой и И. А. Платуновой [1], позволяет выполнять на машине «Стрела» аналитические преобразования формул однородной структуры, заданных при помощи списков.

Списком, представляющим некоторое множество N, называется конечная упорядоченная последовательность конечных строк вида

$$a_{k1},\ a_{k2},\ \ldots,\ a_{kn} \qquad (k = 1, 2, \ldots, l),$$

где a_{ki} суть элементы этого множества.

Например, выражение

$$\varphi(x) \equiv \sum_{k=1}^{l} a_k x^{n_k} \psi_k(m_k x),$$

где

$$\psi_k^{(x)} = \begin{cases} \sin x, \\ \cos x, \end{cases}$$

a_k — числовой коэффициент, n_k и m_k — целые числа, можно представить в виде списка

$$\begin{Vmatrix} a_1, & n_1, & m_1, & \psi_1 \\ a_2, & n_2, & m_2, & \psi_2 \\ a_3, & n_3, & m_3, & \psi_3 \\ \cdot & \cdot & \cdot & \cdot \\ a_l, & n_l, & m_l, & \psi_l \end{Vmatrix},$$

причем одна строка списка (a_k, n_k, m_k, ψ_k) описывает один одночлен

$$a_k x^{n_k} \psi_k(m_k x).$$

При вводе в машину таких списков каждому из них ставится в соответствие его номер, кодировка — фактическая запись — и справка о типе объекта, о размещении его записи в машинной памяти или о способе вычисления его записи и т. п.

Список считается заданным, если либо задана его запись, либо задана схема его получения с помощью операций над списками из числа опорных, имеющих запись.

Во втором случае информация о списке есть схема, определяющая некоторое правило получения его записи по записям опорных списков.

Всякий алгорифм T, перерабатывающий набор a из n элементов множества N в конечную последовательность наборов $a_1^T,\ a_2^T,\ \ldots,\ a_p^T$, состоящих каждый из m элементов множества N, в записи

$$T(a) = a_1^T,\ a_2^T,\ \ldots,\ a_p^T$$

называется допустимым алгорифмом, определенным на множестве N.

Пусть, например, требуется продифференцировать по x выражение

$$\varphi(x) \equiv \sum_{k=1}^{l} a_k x^{n_k} \psi_k(m_k x),$$

где $\psi_k(x)$ было определено выше, и представить производную тоже в виде списка

$$\varphi'(x) = \sum_{k=1}^{l^*} a_k^* x^{n_k^*} \psi_k^*(m_k^* x).$$

Опишем T — допустимый алгорифм, осуществляющий это преобразование:

$$\frac{d}{dx}\left[ax^{n_k}\psi_k(m_k x)\right] = an_k x^{n_k-1}\psi_k(m_k x) \pm am_k x^{n_k}\overline{\psi}_k(m_k x),$$

где $\overline{\psi}_k(x)$ есть также либо $\sin x$, либо $\cos x$. T является объединением двух опорных алгорифмов T_1 и T_2, где T_1 — алгорифм первой результативной строки, T_2 — второй. Здесь под опорным алгорифмом понимается алгорифм, имеющий запись и переводящий каждую строку исходного списка в одну результативную строку. Записью алгорифма является программа в условных адресах.

Итак, алгорифм задается либо записью (программой) и справкой о ней, либо схемой ее получения с помощью операций объединения, соединения суперпозиции, дизъюнкции, рекурренции из опорных алгорифмов, имеющих запись.

Строками вычислительного плана являются строки-указания о построении нового списка

$$(T,\ \Sigma) = \Sigma'.$$

Здесь T — допустимый алгорифм, заданный записью или схемой, Σ — список, заданный записью или схемой.

Последовательное выполнение строк вычислительного плана в порядке их записи нарушается только схемными переходами. Среди строк вычислительного плана допускаются машинные команды и строки-указания выполнить над списком специальные операции, связанные с реализацией плана на машине (например, перевод элементов списка из одной системы в другую, печать списка и т. д.).

С помощью прораба СП III-59 на машине «Стрела» были выполнены аналитические выкладки, описанные в списочной символике.

Например, за 30 минут было получено 12 приближений при отыскании в виде ряда решения уравнения

$$y''(x) + y(x) = \alpha\,[y'(x)]^2,$$

где α — малый параметр, при начальных данных $y(0) = 1$, $y'(0) = 0$.

Т. Н. Первозванская [1] записала в списочной символике и реализовала с помощью прораба СП III-59 вычислительный план решения задачи о нахождении периодического решения уравнения Ван-дер-Поля методом малого параметра.

Для проведения аналитических выкладок в классе полиномов автором [1], [2] был построен специальный прораб. Выделение класса полиномов в качестве исходного и построение для него прораба обусловливается тем, что существует весьма широкий круг задач, алгорифмы решения которых требуют проведения трудоемких полиномиальных выкладок, а применение универсального прораба Л. Т. Петровой для полиномиальных выкладок было бы мало экономным.

В круг задач, опирающийся на полиномиальный класс, входит, в частности, нахождение приближенного решения вариационных задач по методу Ритца. В качестве системы координатных функций, при которых с ростом n метод Ритца остается устойчивым в смысле гл. II, часто можно брать системы ортонормированных (в той или иной метрике) полиномов. Нет необходимости подробно описывать возможные задачи данного круга, следует только отметить, что в него входят задачи, алгорифмы решения которых требуют проведения полиномиальных выкладок.

Выделение в качестве исходного класса полиномов позволило сопоставить объектам наиболее естественную и простую запись в памяти машины и, кроме того, завести довольно широкий набор полиномиальных операций. А это, в свою очередь, привело к тому, что запись вычислительных планов достигла наибольшей простоты и естественности.

Узость исходного класса объектов привела также к тому, что универсальная программа-прораб для этого класса получилась весьма короткой, четкой и обозримой.

Первоначальный вариант прораба был сделан на машину «Стрела-1». Программа содержала около 400 команд.

В настоящее время имеются варианты прораба на машину М-20.

Первый вариант, будем называть его прораб I-II, как и вариант на «Стрелу-1», рассчитан на полиномы алгебраические одной и двух независимых переменных и на полиномы тригонометрические одной независимой переменной.

Второй вариант, будем называть его прораб N, рассчитан на алгебраические полиномы N независимых переменных, а также на объекты более сложной структуры

ГЛАВА I

ПРОРАБ I-II НА МАШИНУ М-20

§ 1. Представление полиномов в памяти машины

В этом параграфе речь будет идти о полиномах алгебраических и тригонометрических. Для определенности будем записывать алгебраические полиномы по возрастающим степеням независимых переменных

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n, \tag{1.1}$$

$$F(x, y) = a_{00} + (a_{10}x + a_{11}y) + \dots + (a_{n0}x^n + a_{n1}x^{n-1}y + \dots + a_{nn}y^n). \tag{1.2}$$

Для тригонометрических полиномов сохраним их обычную запись

$$T_n(x) = \sum_{k=0}^{n} (a_k \cos k\omega x + b_k \sin k\omega x). \tag{1.3}$$

В машинной памяти каждый из этих полиномов (1.1), (1.2), (1.3) представляется записью коэффициентов и справкой.

Коэффициенты полинома, включая и равные нулю, записываются в подряд расположенные ячейки P, $P+1$, $P+2$, ..., $P+s-1$. Здесь s есть количество одночленов у полинома. Так, для полинома (1.1) $s = n+1$, для (1.2) $s = \frac{(n+1)(n+2)}{2}$, для (1.3) $s = 2n+2$.

Справка о полиноме записывается в одну ячейку памяти и включает в себя следующие сведения: l — признак типа полинома, P — адрес младшего коэффициента, n — степень (порядок) полинома, $\varkappa$ — адрес ω для полиномов (1.3).

Поставим в соответствие полиномам типов (1.1), (1.2), (1.3) следующие числа $l = 4$, $l = 5$, $l = 6$.

Тогда справки об (1.1), (1.2) и (1.3) соответственно будут иметь вид:

л	КОП	A_1	A_2	A_3
4	—	P	n	0000

л	КОП	A_1	A_2	A_3
5	—	P	n	0000

л	КОП	A_1	A_2	A_3
6	—	P	n	$\varkappa$

Содержание КОП в справке безразлично.

Зная степень полинома и его тип, а также адрес младшего коэффициента, мы однозначно определим массив ячеек, в которых записаны коэффициенты полинома.

Справки о полиномах располагаются в специальном массиве $R \div R_m$, который называется полем полиномиальных справок.

Коэффициенты полиномов размещаются в своем поле — поле коэффициентов $P_1 \div P_t$.

Назовем номером полинома адрес ячейки, где записана его справка. Номером числа назовем адрес этого числа. Кроме того, обозначим (R_i) полином с номером R_i, а (c) обозначим число с номером c.

§ 2. Операции над полиномами

Для проведения на машине выкладок с полиномами введем в рассмотрение опорный список полиномиальных операций. В этот список входят операции сложения, вычитания, умножения, деления полиномов (одной независимой переменной), дифференцирования, неопределенного интегрирования, вычисления значения полинома в точке и т. д.

В список входят и операции специального назначения: схемные переходы, печать полиномов, стирание (гашение) полиномов, сравнение полиномов по степени и т. д.

Ниже (см. табл. 2.1 — 2.4) приведены списки имеющихся операций. Эти списки можно дополнять новыми операциями. При их оформлении следует строго придерживаться инструкции.

Каждая операция в памяти машины представляется своей записью — подпрограммой в условных адресах — и справкой.

Справки об операциях записываются в ячейках поля $\theta + 0 \div \theta + 67_8$. Операция, справка о которой оказывается записанной в ячейке $\theta + t$, получает соответственно номер t $(t = 0, 1, 2, \ldots, 67_8)$.

Пусть условная запись некоторой полиномиальной операции находится в ячейках $s + 0 \div s + N$. Тогда ее справка в поле имеет вид:

π	КОП	A_1	A_2	A_3
r	φ	$s+0$	N	0000

Если операция такова, что в результате ее выполнения получается полином, то $r \neq 0$, в противном случае $r = 0$. Например, справки операций сложения, вычитания, умножения и т. п. имеют $r \neq 0$, а справки операций специального вида, таких, как печать, схемный переход, сравнение полиномов по степени и т. п., имеют $r = 0$.

Число φ в справке об операции является величиной сдвига входа в подпрограмму по отношению к ее началу. Например, подпрограмма вычитания полиномов занимает ячейки $s_1 + 0 \div s_1 + N_1$. Подпрограмма сложения занимает ячейки $s_1 + 1 \div s_1 + N_1$. Таким образом, справка об операции сложения имеет $\varphi = 1$.

Отметим теперь требования, которые следует выполнять при составлении подпрограмм полиномиальных операций.

1) Адрес команд подпрограмм следует записывать условно. Это требование делает подпрограммы независимыми от рабочего поля.

2) Адреса полиномиальных коэффициентов записываются условно. Значения степени n и количества одночленов s записываются в единицах адресов тоже условно.

3) В качестве рабочих используются ячейки, в которых находятся переменные команды прораба.

4) При составлении подпрограмм полиномиальных операций можно пользоваться следующими константами:

Номер ячейки	π	КОП	A_1	A_2	A_3
ρ_1	0	00	0000	n_1	0000
ρ_2	0	00	0000	n_2	0000
ρ_3	0	00	0000	s_1	0000
ρ_4	0	00	0000	s_2	0000

Эти константы формируются прорабом. Здесь n_1 и n_2, s_1 и s_2 — соответственно степени и количества одночленов первого и второго полиномов-аргументов.

Остановимся подробно на правиле вычисления условных адресов.

Имеет место следующее правило вычисления условных адресов:

$$\boxed{\text{Условный адрес} = 4000 + 20 \cdot \Delta + C(k)} \tag{2.1}$$

Здесь 4000 — признак условности, Δ — поправка к адресу ($\Delta = 0, 1, 2, 3, \ldots$ $\ldots, 140$) либо адрес поправки ($\Delta = 141, 142, \ldots, 177$), $C(k)$ — условный номер k.

Введена следующая система условных номеров:

$C(B+0) = 0015$, $B+0$ — начало рабочего поля;

$C(B+N) = 0016$, $B+N-1$ — конец рабочего поля;

$C(R_k) = 0017$, R_k — номер справки результата или адрес результата;

$C(P_1) = 0001$, P_1 — адрес первого коэффициента первого аргумента;

$C(n_1) = 0002$, n_1 — степень первого аргумента;

$C(s_1) = 0003$, s_1 — количество одночленов первого аргумента;

$C(P_2) = 0005$, P_2 — адрес первого коэффициента второго аргумента;

$C(n_2) = 0006$, n_2 — степень второго аргумента;

$C(s_2) = 0007$, s_2 — количество одночленов второго аргумента;

$C(P_{\text{св}}) = 0011$, $P_{\text{св}}$ — номер первой свободной ячейки в поле коэффициентов;

$C(0) = 0012$.

Пример 1. Условный адрес $B+4$ (ячейки рабочего поля $B+0 \div B+$ $+N-1$) равен

$$4000 + 20 \cdot 4 + 0015 = 4115.$$

Пример 2. Условный адрес 3-го коэффициента первого аргумента равен

$$4000+20\cdot 3+0001=4061.$$

Как уже было сказано выше, $\Delta=141, \ldots, 177$ в формуле (2.1) играет роль адреса справки о поправке. В этом случае поправка либо отрицательная, либо переменная, зависящая от значений n и s (степени и количества одночленов у полиномов-аргументов).

Справка о поправке записывается в одну ячейку памяти. Для справок о переменных и отрицательных поправках отводится поле $T+0 \div T+36$. Поправка, справка о которой помещается в ячейку $T+j$ этого поля, получает номер $\Delta=141+j$.

Справка включает в себя сведения о том, из каких величин составляется поправка. Как показывает опыт написания условных записей подпрограмм полиномиальных операций, наиболее часто употребляемые поправки — это поправки к адресу, равные

$$n_1,\ n_2,\ s_1,\ s_2,\ n_1+n_2,\ n_1-1,\ s_2+1,\ -1 \quad \text{и т. д.}$$

В образовании каждой поправки, имеющей номер $\Delta=141+j$ ($j=0$, 1, 2, ..., 36), участвует не более трех величин: $\pm a_1$, $\pm a_2$, $\pm a_3$.

Знак числа записывается в старшем разряде адреса A_i справки, а в прочих разрядах адреса A_i записывается адрес ячейки[1]), в единицах A_2 которой содержится абсолютная величина a_i.

Например, известно, что ρ_1, ρ_2, ρ_3, ρ_4 суть номера ячеек, в A_2 которых содержатся степени n_1 и n_2, количества одночленов s_1 и s_2 у полиномов-аргументов. Следовательно, справка о поправке s_1-n_1 имеет вид:

0	00	ρ_3	$4000+\rho_1$	0000

,

а справка о поправке n_1+n_2 имеет вид:

0	00	ρ_1	ρ_2	0000

и т. д.

Пример 3. Условный адрес последнего коэффициента первого полинома-аргумента, P_1+s_1-1, если поправка s_1-1 к адресу P_1 имеет номер 141, таков:

$$4000+20\cdot 141+0001=7021.$$

Пример 4. Условная запись команды

1	12	s_1	$B+10$	7777

такова:

1	12	4003	4215	7232

[1]) Содержание A_1 и A_2 этой ячейки не имеет значения.

Условная запись числа 7777 получается по формуле (2.1) следующим образом:

$$4000 + 20 \cdot 151 + C(0),$$

где $C(0) = 0012$, 151 — номер поправки, равной —1. Следовательно, в A_3 команды вместо 7777 записываем 7232.

Условная запись адреса $B + 10$ рабочего поля подпрограммы вычисляется по формуле (2.1)

$$4000 + 20 \cdot 10 + 0015 = 4215.$$

Ниже приводится записанный в общем виде опорный список полиномиальных операций.

Приведенные операции разбиты на четыре группы. Операции группы I суть операции над полиномами одной независимой переменной. Операции группы II суть операции над полиномами двух переменных. Операции группы III не зависят от числа переменных и типа полиномов. Операции группы IV суть операции над тригонометрическими полиномами одной переменной.

§ 3. Вычислительный план

Точным и полным описанием алгорифма решения задачи в терминах языка полиномов, чисел и операций над ними является вычислительный план (ВП).

ВП представляет собой последовательность строк — указаний о построении очередных полиномов и чисел. Строка ВП — это команда для прораба. Прораб работает по ВП, т. е. он выполняет все указания — команды ВП — в предписанном порядке. Среди строк ВП могут быть строки четырех типов:

1) Полиномиальные.
2) Машинные.
3) Схемные переходы.
4) Строки обращения к подсхемам.

1. Полиномиальные строки. 1°. Полиномиальная строка — это указание выполнить операцию, имеющую номер t над полиномами или над полиномом и числом. Результатом выполнения такой строки может быть полином или число

$$t(x_1, x_2) \to x_3.$$

В полиномиальной строке x_1 — всегда полином.

Полиномиальная строка записывается в одну ячейку памяти, причем номер операции t помещается в КОП, номер первого аргумента — в A_1, номер второго аргумента в A_2, номер результата — в A_3. В разрядах π записывается признак того, что эта строка полиномиальная, а именно, $\pi_3 = 1$:

π	КОП	A_1	A_2	A_3
1	t	s_1	s_2	s_3

Здесь $(s_i) = x_i$ $(i = 1, 2, 3)$.

Пример 1. Строку ВП

1	01	R_i	R_j	R_k

прораб воспринимает как команду вычислить полином

$$(R_k) = (R_i) + (R_j).$$

Таблица 2.1

I. Операции над полиномами одной независимой переменной

№ операции	Название	Справки					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
22	Сложение	1	01	T_0+1	0006	0000	1	22	R_1	R_2	R_3	$(R_3)=(R_1)+(R_2)$
23	Вычитание	1	00	T_0+0	0007	0000	1	23	R_1	R_2	R_3	$(R_3)=(R_1)-(R_2)$
24	Умножение	1	00	T_1+0	0021	0000	1	24	R_1	R_2	R_3	$(R_3)=(R_1)\cdot(R_2)$
04	Деление без остатка	1	00	T_2+0	0051	0000	1	04	R_1	R_2	R_3	$(R_1)=(R_2)\cdot(R_3)+(R_3+1)$
11	Деление с остатком	1	02	T_2+2	0047	0000	1	11	R_1	R_2	R_3	Справки частного и остатка поступают в подряд расположенные ячейки R_3 и R_{3+1}. При $n_1 < n_2$ (n_1 — степень делимого, n_2 — степень делителя) останов машины
07	Вычисление значения полинома в точке	0	00	T_3+0	0004	0000	1	07	R_1	A	B	$(B)=(R_1)_{x=a}$, где $a=(A)$
06	Вычисление неопределенного интеграла	1	00	T_4+0	0004	0000	1	06	R_1	0000	R_3	$(R_3)=\int(R_1)\,dx$

Продолжение

№ операции	Название	Справки					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
05	Дифференцирование	1	00	T_5+0	0004	0000	1	05	R_1	0000	R_3	$(R_3)=\frac{d}{dx}(R_1)$
10	Вычисление интеграла по промежутку	0	00	T_6+0	0013	0000	1	10	R_1	A	B	$(B)=\int\limits_a^b (R_1)\,dx$, где $(A)=(0,\ 00,\ m,\ n,\ 0000)$, $(m)=a$, $(n)=b$
12	Построение полинома Лежандра	0	00	T_7+0	0036	0000	1	12	R_1	A	R_3	$(R_3)=\frac{2n-1}{n}xP_{n-1}(x)-\frac{n-1}{n}P_{n-2}(x)$, $P_0(x)=1$, $P_1(x)=x$, $(A)=(4,\ 00,\ 0000,\ n,\ 0000)$
37	Вычисление интеграла $\int\limits_0^\infty e^{-\alpha x}P(x)\,dx$, $\alpha>0$	0	00	$T_{10}+0$	0011	0000	1	37	R_1	A	B	$(B)=\int\limits_0^\infty e^{-\alpha x}(R_1)\,dx$, где $\alpha=(A)$ — вещественное положительное число $(R_1)=P(x)$
17	Умножение полинома на число	1	00	$T_{11}+0$	0002	0000	1	17	R_1	A	R_3	$(R_3)=a(R_1)$, где $a=(A)$

Таблица 2.2

II. Операции над полиномами двух независимых переменных

№ операции	Название	Справка					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
01	Сложение	2	01	T_0+1	0006	0000	1	01	R_1	R_2	R_3	$(R_3)=(R_1)+(R_2)$
02	Вычитание	2	00	T_0+0	0007	0000	1	02	R_1	R_2	R_3	$(R_3)=(R_1)-(R_2)$
03	Умножение	2	00	$T_{12}+0$	0041	0000	1	03	R_1	R_2	R_3	$(R_3)=(R_1)\cdot(R_2)$
25	Дифференцирование по x	2	00	$T_{13}+0$	0011	0000	1	25	R_1	0000	R_3	$(R_3)=\frac{\partial}{\partial x}(R_1)$
26	Дифференцирование по y	2	00	$T_{14}+0$	0011	0000	1	26	R_1	0000	R_3	$(R_3)=\frac{\partial}{\partial y}(R_1)$
27	Вычисление неопределенного интеграла по x	2	00	$T_{15}+0$	0011	0000	1	27	R_1	0000	R_3	$(R_3)=\int(R_1)\,dx$
30	Вычисление неопределенного интеграла по y	2	00	$T_{16}+0$	0011	0000	1	30	R_1	0000	R_3	$(R_3)=\int(R_1)\,dy$

Продолжение

№ операции	Название	Справка π	Справка КОП	Справка A_1	Справка A_2	Справка A_3	Запись полиномиальной строки π	Запись полиномиальной строки КОП	Запись полиномиальной строки A_1	Запись полиномиальной строки A_2	Запись полиномиальной строки A_3	Примечание
31	Вычисление интеграла по эллипсу (кругу)	0	00	$T_{17}+0$	0035	0000	1	31	R_1	A	B	$(B)=\int\limits_{-R}^{R}\int\limits_{-\frac{r}{R}\sqrt{R^2-x^2}}^{\frac{r}{R}\sqrt{R^2-x}}(R_1)\,dx\,dy$ $(A)=(0,\ 00,\ m,\ n,\ 0000)$, $(m)=R,\ (n)=r$
32	Вычисление интеграла по половине эллипса (круга)	0	00	$T_{17}+0$	0075	0000	1	32	R_1	A	B	$(B)=\int\limits_{-R}^{R}\int\limits_{0}^{\frac{r}{R}\sqrt{R^2-x^2}}(R_1)\,dx\,dy$ $(A)=(0,\ 00,\ m,\ n,\ 0000)$, $(m)=R,\ (n)=r$
33	Вычисление интеграла по прямоугольнику	0	00	$T_{20}+0$	0040	0000	1	33	R_1	A	B	$(B)=\int\limits_{a}^{b}\int\limits_{c}^{d}(R_1)\,dx\,dy$ $(A)=(0,\ 00,\ m,\ n,\ 0000)$, $(m)=a,\ (n)=c$, $(m+1)=b,\ (n+1)=d$
34	Подстановка в полином $y=a$	1	00	$T_{21}+0$	0021	0000	1	34	R_1	A	R_3	$(R_3)=(R_1)_{y=a}$, $(A)\ =a$
35	Умножение на число	2	00	$T_{11}+0$	0002	0000	1	35	R_1	A	R_3	$(R_3)=a\cdot(R_1)$, $(A)\ =a$

Таблица 2.3

III. Операции, не зависящие от числа переменных и типа полиномов

№ операции	Название	Справка					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
13	Печать	0	00	$T_{22}+0$	0011	0000	1	13	R_1	0000	0000	Выдача на печать в восьмеричной системе справки R_1 и в десятичной системе коэффициентов (R_1)
15	Сравнение полинома с нулем	0	00	$T_{23}+0$	0001	0000	1	15	R_1	A	B	Полином считается тождественно равным нулю, если все его коэффициенты по модулю меньше некоторого положительного δ. $\delta=(A)$, $(B)=\begin{cases}\alpha \geqslant & \text{если } (R_1)\equiv 0,\\ \beta<0, & \text{если } (R_1)\not\equiv 0\end{cases}$
16	Сравнение полинома по степени	0	00	$T_{24}+0$	0001	0000	1	16	R_1	R_2	B	$(B)=\begin{cases}\alpha\geqslant 0, & \text{если } n_1\geqslant n_2,\\ \beta<0, & \text{если } n_1<n_2\end{cases}$
20	Сложение полинома с числом	0	00	$T_{25}+0$	0004	0000	1	20	R_1	A	R_3	$(R_3)=a+(R_1)$, $(A)=a$
14	Гашение (стирание) полинома	0	00	$T_{26}+0$	0000	0000	5	14	R_1	0000	0000	Стирание R_1,
							7	14	R_1	R_2	0000	Стирание R_1 и R_2,
							3	14	0000	R_2	0000	Стирание R_2
00	Схемный переход	0	00	$T_{27}+0$	0004	0000	1	00	$K+i$	A	$K+j$	Если $(A)\geqslant 0$, перейти к выполнению строки с номером $K+j$, если $(A)<0$ — к выполнению $K+i$
							1	00	$K+i$	A_ω	$K+j$	Если в результате выполнения последней машинной строки выработался признак $\omega=0$, перейти к $K+j$, если $\omega=1$ — к $K+i$

Таблица 2.4

IV. Операции с тригонометрическими полиномами

№ операции	Название	Справка					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
Q_1	Сложение	4	01	T_0+1	0006	0000	1	Q_1	R_1	R_2	R_3	$(R_3)=(R_1)+(R_2)$
Q_2	Вычитание	4	00	T_0+0	0007	0000	1	Q_2	R_1	R_2	R_3	$(R_3)=(R_1)-(R_2)$
Q_3	Умножение	4	00	$T_{30}+0$	0047	0000	1	Q_3	R_1	R_2	R_3	$(R_3)=(R_1)\cdot(R_2)$
Q_4	Дифференцирование	4	00	$T_{31}+0$	0007	0000	1	Q_4	R_1	0000	R_3	$(R_3)=\frac{d}{dx}(R_1)$
Q_5	Интегрирование по промежутку	0	00	$T_{32}+0$	0030	0000	1	Q_5	R_1	A	B	$(B)=\int_a^b (R_1)\,dx$ $(A)=(0,\ 00,\ m,\ n,\ 0000)$, $(m)=a$, $(n)=b$
Q_6	Умножение на число	4	00	$T_{11}+0$	0002	0000	1	Q_6	R_1	A	R_3	$(R_3)=a\cdot(R_1)$ $(A)=a$

Пример 2. Строку ВП

1	17	R_i	a	R_k

прораб воспринимает как приказ умножить полином (R_i) на число (a):

$$(R_k) = (a) \cdot (R_i).$$

Пример 3. Строку ВП

1	07	R_i	a	b

прораб воспринимает как приказ вычислить значение полинома (R_i) в точке $x = (a)$ и результат записать в ячейку b.

Если в полиномиальной строке ВП поместить отметки $\pi_1 = 1$ и $\pi_2 = 1$,

7	t	s_1	s_2	s_3

то прораб воспримет такую строку как приказ, во-первых, выполнить указанную операцию t над аргументами (s_1) и (s_2),

$$t\,[(s_1),\ (s_2)] = (s_3),$$

во-вторых, затереть полиномы-аргументы (s_1) и (s_2).

Аналогично, строка с отметками $\pi_1 = 1$, $\pi_2 = 0$

5	t	s_1	s_2	s_3

или с отметками $\pi_1 = 0$, $\pi_2 = 1$

3	t	s_1	s_2	s_3

будет восприниматься прорабом как приказ выполнить операцию

$$t\,[(s_1),\ (s_2)] = (s_3),$$

а затем затереть соответственно первый полином-аргумент (s_1) или второй полином-аргумент (s_2).

Пример 4. Полиномиальная строка

7	01	R_i	R_j	R_k

где R_i и R_j — номера полиномов, требует, чтобы прораб выполнил сложение полиномов

$$(R_i) + (R_j) = (R_k)$$

и затер полиномы (R_i) и (R_j).

Пример 5. Полиномиальная строка

5	03	R_i	R_j	R_i

является указанием для прораба перемножить полиномы (R_i) и (R_j), затереть полином-аргумент (R_i), справку произведения $(R_i)\cdot(R_j)$ записать в ячейку R_i. Таким образом, R_i теперь оказывается справкой полученного произведения.

После выполнения прорабом строк ВП, имеющих отметки о стирании аргументов, эти аргументы исчезают из памяти машины.

Итак, для полиномиальных строк ВП в разрядах π записываются следующие числа:

$$\pi = \begin{cases} 1 & \text{стирания нет,} \\ 3 & \text{стирание } (s_2), \\ 5 & \text{стирание } (s_1), \\ 7 & \text{стирание } (s_1) \text{ и } (s_2). \end{cases}$$

2°. Среди строк ВП могут быть строки групповые полиномиальные. Групповая строка размещается в двух ячейках. Она предписывает прорабу выполнить полиномиальную строку m раз, произведя при этом нужную переадресацию.

В общем виде групповая строка записывается следующим образом:

Номер ячейки	π	КОП	A_1	A_2	A_3
$\alpha+1$	0	77	p	$m-1$	h
$\alpha+2$	π	t	R_i	R_j	R_k

Групповая полиномиальная строка состоит из подготовительной строки $\alpha+1$ и исполнительной $\alpha+2$. В КОП-е подготовительной строки записывается условный номер 77, π всегда равно нулю. В A_1, A_2, A_3 записывается следующая информация: p — условное число, показывающее, какие адреса в исполнительной строке подлежат переадресации:

p	Подлежит переадресации в исполнительной строке			p	Подлежит переадресации в исполнительной строке		
0000	—	—	—	0004	A_1	—	—
0001	—	—	A_3	0005	A_1	—	A_3
0002	—	A_2	—	0006	A_1	A_2	—
0003	—	A_2	A_3	0007	A_1	A_2	A_3

m — число, показывающее, сколько раз должна выполняться полиномиальная строка, h — шаг переадресации, $h \geqslant 0$.

Пример 6. Групповая строка

0	77	0007	0003	0001
1	03	R_i	R_j	R_k

требует от прораба вычислить следующие произведения:

$$(R_k) = (R_i)\cdot(R_j),$$
$$(R_k+1) = (R_i+1)\cdot(R_j+1),$$
$$(R_k+2) = (R_i+2)\cdot(R_j+2),$$
$$(R_k+3) = (R_i+3)\cdot(R_j+3).$$

Если в исполнительной строке поместить $\pi = 7$, то получим групповую строку, предписывающую последовательно строить произведения и стирать аргументы. В результате выполнения такой групповой полиномиальной строки будут построены полиномы (R_k), (R_k+1), (R_k+2), (R_k+3), а полиномы (R_i), (R_j), (R_i+1), (R_j+1), (R_i+2), (R_j+2), (R_i+3), (R_j+3) окажутся затертыми, т. е. исчезнут из памяти машины.

Пример 7. Групповая строка

0	77	0000	0005	0000
5	01	R_i	R_j	R_l

является командой прорабу выполнить шесть раз сложение полиномов R_i и R_j, причем каждый раз стирать первый аргумент (R_i), справку суммы записывать в R_l. После выполнения этой строки справка полинома

$$(R_i)+6\cdot(R_j)$$

будет записана в ячейке R_l.

Пример 8. Групповая строка

0	77′	0004	0007	0002
5	13	R_i	0000	0000

является указанием прорабу выдавать на печать и сразу же после этого затирать в памяти полиномы (R_i), (R_i+2), (R_i+4), (R_i+6), (R_i+10), (R_i+12), (R_i+14), (R_i+16).

2. Машинные строки. Машинная строка — это обычная машинная команда. В ВП ее отличительным признаком является $\pi_3 = 0$.

Прораб выполняет эту команду так, как она записана, т. е. в обычном смысле М-20:

λ	φ	a	b	c

,

где $\lambda = 0, 2, 4, 6$; φ — код машинной операции; a, b, c — адреса чисел. Если в ВП встречаются большие группы машинных строк, то естественно эти группы выполнять на машинном режиме, а не на режиме прораба. Это достигается следующим путем.

Вместо необходимой группы машинных строк в ВП пишется лишь одна строка выхода на подпрограмму (группу машинных команд), которая помещается в конце ВП. Причем выход из подпрограммы на прораб осуществляется передачей управления на соответствующую команду прораба.

Следует отметить, что ограничения на π_3 в машинных строках, вынесенных таким образом, полностью снимаются.

3. Строки схемного перехода. Строки схемного перехода — это строки, позволяющие изменять порядок выполнения вычислительного плана в зависимости от получающихся результатов.

Строка схемного перехода предписывает прорабу перейти к выполнению той или иной строки (и следующих за нею) в зависимости от того, что получилось к моменту выполнения строки схемного перехода.

Различаются схемные переходы двух видов.

а) Схемный переход по знаку числа

1	00	A	p	B

.

Он предписывает прорабу перейти к выполнению строки B, если число $(p) \geqslant 0$, в противном случае перейти к выполнению строки A.

б) Схемный переход по сигналу ω

1	00	A	A_ω	B

.

Он предписывает прорабу перейти к выполнению строки B, если в результате выполнения предшествующей машинной строки, вырабатывающей сигнал ω, был выработан сигнал $\omega = 0$. Если же выработан сигнал $\omega = 1$, то прорабу следует перейти к строке A.

В ячейке A_ω прораб хранит число, поставленное в соответствие выработанному сигналу ω, до выполнения следующей машинной строки ВП, вырабатывающей сигнал ω.

4. Строки обращения к подсхемам. Часть ВП, выполняющуюся многократно при различных полиномах и числах, удобно выделить в подсхему. В некотором смысле подсхему можно отождествлять с подпрограммой.

Подсхема записывается в условных адресах. Справка о подсхеме содержит сведения о местоположении ее в памяти. Все подсхемы, на которые опирается ВП, нумеруются. Для подсхем допускаются следующие номера: 70, 71, ..., 77.

В ВП строка обращения к подсхеме является указанием прорабу о том, какую подсхему и для каких полиномов и чисел следует выполнить.

Строка обращения записывается в одну, две или три подряд расположенные ячейки (число ячеек не должно превосходить трех). Первая из ячеек называется главной, прочие — дополнительными.

Главная ячейка содержит в разрядах КОП номер подсхемы, а в разрядах π следующие числа:

$$\pi = \begin{cases} 1 \text{ — строка обращения в одной ячейке,} \\ 3 \text{ — » » в двух ячейках,} \\ 5 \text{ — » » в трех ячейках.} \end{cases}$$

Назовем адреса главной и дополнительных ячеек адресами строки обращения:

A_1^0, A_2^0, A_3^0 — адреса главной ячейки,

A_1^1, A_2^1, A_3^1 — адреса первой дополнительной ячейки,

A_1^2, A_2^2, A_3^2 — адреса второй дополнительной ячейки.

В этих адресах записываются номера полиномов и чисел, являющихся номерами аргументов и результатов.

Перенумеруем адреса строки обращения следующим образом. Пусть условный номер адреса A_k^i обозначается $C\left(A_k^i\right)$, тогда

$$\begin{aligned} &C\left(A_1^0\right)=0001, \quad &C\left(A_1^1\right)=0005, \quad &C\left(A_1^2\right)=0011,\\ &C\left(A_2^0\right)=0002, \quad &C\left(A_2^1\right)=0006, \quad &C\left(A_2^2\right)=0012,\\ &C\left(A_3^0\right)=0003, \quad &C\left(A_3^1\right)=0007, \quad &C\left(A_3^2\right)=0013. \end{aligned}$$

Полином или число, номер которого s записывается в адресе A_k^i строки обращения, получает условный номер этого адреса

$$C\left(A_k^i\right)=C(s).$$

Пример 1. Пусть строка обращения к подсхеме, имеющей номер 75, записана следующим образом:

5	75	R_i	R_j	m
—	—	R_k	a	R_p
—	—	s	R_n	t

Тогда полиномы (R_i), (R_j), (R_k), (R_p), (R_n) и числа (m), (a), (s), (t) получают следующие номера:

$$\begin{aligned} &C(R_i)=0001, \quad &C(R_k)=0005, \quad &C(s)=0011,\\ &C(R_j)=0002, \quad &C(a)=0006, \quad &C(R_n)=0012,\\ &C(m)=0003, \quad &C(R_p)=0007, \quad &C(t)=0013. \end{aligned}$$

Подсхемы выполняются на рабочем поле, отводимом прорабом. Это поле — $D+0 \div D+N$. Условный номер первой ячейки рабочего поля таков

$$C(D+0)=0017.$$

Итак, в подсхеме номера полиномов чисел, адреса ячеек рабочего поля $D+0 \div D+N$ записываются условно.

Имеет место правило вычисления условных адресов (как и в случае оформления подпрограмм полиномиальных операций, см. § 2):

$$\text{Условный адрес} = 4000+20\Delta+C(r).$$

Для подсхем $\Delta=0, 1, 2, \ldots, 140$. Отсюда следует, что длина подсхемы не должна превосходить $N \leqslant 140_8$.

Пример 2. В примере 1, в строке обращения к подсхеме среди номеров полиномов и чисел, перечислен номер t числа (t). В подсхеме ему должен соответствовать условный адрес

$$4000+20\cdot 0+0013=4013.$$

Условный адрес полинома, имеющего номер R_i+4 (в строке обращения, см. пример 1, перечислен номер R_i), имеет вид:

$$4000+20\cdot 4+0001=4101.$$

Условный адрес ячейки $D+15$ рабочего поля таков

$$4000+20\cdot 15+0017=4337.$$

Для удобства при кодировке подсхем программист может составить таблицу условных адресов ячеек рабочего поля $D+0\div D+N$:

$$\begin{aligned}&D+0-4017,\\&D+1-4037,\\&D+2-4057,\\&\dots\dots\dots\dots\\&D+k-2017+20\cdot k.\end{aligned}$$

§ 4. Прораб и автоматическое распределение памяти

Программа прораба I-II содержит 526 команд. К программе прораба прилагается набор подпрограмм полиномиальных операций. В сумме эти подпрограммы содержат 405 команд.

Остановимся подробнее на том, как оформляется задача, которую мы хотим решить с помощью полиномиального прораба. Это облегчит понимание структуры прораба и порядка его работы.

1. На языке полиномов, чисел и операций над ними записывается вычислительный план (ВП), определяется номер его начальной строки M_{cx}.

2. Отводится поле полиномиальных справок $R_1\div R_t$. Массив полиномиальных справок образуют ячейки, в которых записываются справки заданных полиномов, а также те ячейки, в которые прораб будет записывать справки вычисляемых полиномов. Поле полиномиальных справок может состоять всего лишь из одной справки. Это значит, что справка исходного промежуточного и результативного полиномов записывается последовательно в одну и ту же ячейку поля R_1.

3. Если необходимо, составляются условные записи подсхем и заводится поле $L+0\div L+s$ подсхемных справок. Кроме того, определяется максимальная длина N подсхем, на которые опирается ВП данной задачи.

4. ВП, поле полиномиальных справок, подсхемы и поле их справок вводятся на свободное место после прораба и списка полиномиальных операций. Вслед за этим определяется поле коэффициентов. Полем коэффициентов является часть оперативной памяти, начиная с ячейки, где записан первый полиномиальный коэффициент, и кончая ячейкой, за которой следуют ячейки рабочих полей.

В этом поле прораб размещает коэффициенты всех вычисляемых полиномов.

В поле коэффициентов записываются коэффициенты заданных полиномов в двоичной или в десятичной системе и заполняются справки об этих полиномах. После того как все исходные полиномы записаны, определяем $P_{св}$ — номер первой свободной ячейки в поле коэффициентов.

5. Для того чтобы прораб настроился на выполнение поставленной перед ним задачи, необходимо сообщить ему исходную информацию:

a) M_{cx} — номер начальной строки,

b) R_1 — номер первой ячейки поля полиномиальных справок и R_t — номер последней ячейки того же поля,

c) P_1 — номер первой ячейки поля коэффициентов и
$P_{\text{св}}$ — номер первой свободной ячейки в этом поле,
d) $L+0$ — номер первой ячейки поля подсхемных справок и
N — максимальную длину подсхем.

Совокупность перечисленных сведений, задаваемых прорабу для его настройки к решению поставленной задачи, будем называть «шапкой» задачи.

В «шапке» следует также отметить, подлежат ли переводу в двоичную систему коэффициенты исходных полиномов.

В памяти машины «шапка» задачи записывается в четыре ячейки памяти 0001, 0002, 0003, 0004 следующим образом:

Номер ячейки	π	КОП	A_1	A_2	A_3
0001	0	16	M_{cx}	0014	0632
0002	p	00	$P_{\text{св}}$	0000	P_1
0003	0	00	R_t	R_1	0000
0004	0	00	N	$L+0$	0000

Теперь рассмотрим порядок работы прораба.

Перед тем как начать выполнять строки ВП, прораб обрабатывает информацию, заключенную в «шапке».

1. Прораб определяет размеры рабочего поля $B+0 \div B+s$ для подпрограмм полиномиальных операций и отводит в конце оперативной памяти это рабочее поле.

2. Если в ячейке 0004 стоит не нуль, то впритык к первому полю прораб отводит рабочее поле для подсхем.

3. По отметке $p=1$ в «шапке» прораб переводит коэффициенты заданных полиномов из десятичной системы в двоичную систему. Если $p=0$, то перевода не будет.

После всего этого прораб начинает основной цикл работ: перебирает в предписанном порядке строки ВП, выполняет их, по полученным результатам выбирает направление дальнейших вычислений, размещает в памяти вычисленные полиномы, стирает полиномы, ненужные в дальнейшем.

Встречая машинную строку, прораб «выполняет» ее, т. е. передает управление на эту команду, после чего выбирает следующую строку.

Если машинная команда вырабатывает сигнал ω, то прораб «хранит» его до тех пор, пока новая машинная строка выработает сигнал ω. То же самое происходит с содержанием RA (регистра адреса).

Если встречается полиномиальная строка, прораб извлекает справки полиномов-аргументов или справку первого аргумента и число — второй аргумент и заполняет информационную таблицу:

Номер ячейки	π	КОП	A_1	A_2	A_3
0005	l	—	P_i	n_i	s_i
0006	l	—	P_j	n_j	s_j
0007	l	—	$P_{\text{св}}$	0000	0000
0010	0	00	$B+0$	$B+N_1$	R_k

В информационной таблице справки полиномов-аргументов оказываются более полными, так как в A_3 прораб записывает подсчитанные им количества одночленов s_i и s_j. В случае, когда второй аргумент есть число, в ячейку 0006 записывается это число.

В информационной таблице $P_{\text{св}}$ — номер текущей свободной ячейки поля коэффициентов, $B+0 \div B+N_1-1$ — часть рабочего поля полиномиальных операций, куда прораб перенесет подпрограмму полиномиальной операции, R_k — адрес справки результативного полинома.

По данным справок о полиномах-аргументах прораб составляет текущие константы:

Номер ячейки	π	КОП	A_1	A_2	A_3
0011	0	00	0000	n_i	0000
0012	0	00	0000	n_j	0000
0013	0	00	0000	s_i	0000
0014	0	00	0000	s_j	0000

В подпрограмме полиномиальной операции, перенесенной на рабочее поле, прораб заменяет условные адреса на истинные, опираясь на информационную таблицу.

Затем прораб передает управление подпрограмме, и она выполняется. При этом коэффициенты размещаются на свободном поле, справка записывается в R_k, управление снова передается прорабу, и прораб переходит к выполнению следующей строки ВП.

Если в строке ВП имеется отметка о том, что в дальнейшем один или оба полинома-аргумента больше не нужны, то после выполнения соответствующей операции над аргументами прораб затирает в поле коэффициенты отмеченных полиномов, коэффициенты других полиномов подтягивает на освободившееся место, а в справки полиномов, коэффициенты которых переместились в памяти, вносит соответствующие поправки. Если по ходу вычислений все поле коэффициентов оказывается занятым, прораб осуществляет «останов».

Встречая в ВП строку обращения к подсхеме, прораб по номеру подсхемы разыскивает ее справку. По справке разыскивается условная запись подсхемы и переносится на рабочее поле $D+0 \div D+N$. На основе строки обращения прораб заполняет информационную таблицу

Номер ячейки	π	КОП	A_1	A_2	A_3
0005	—	—	s_1	s_2	s_3
0006	—	—	s_4	s_5	s_6
0007	—	—	s_7	s_{10}	s_{11}
0010	—	—	0000	0000	$D+0$

Здесь $s_1 \div s_{11}$ — номера полиномов и чисел, перечисленных, как и в строке обращения.

Затем прораб заменяет условные адреса подсхемы на истинные согласно информационной таблице. Перед тем как перейти к выполнению строк подсхемы, прораб запоминает место выхода. Выполнив подсхему, прораб возвращается к той строке вычислительного плана, которая следует за строкой обращения к подсхеме.

Схемный переход заставляет прораба изменить соответствующим образом порядок вычислений. Структуру прораба и порядок его работы лучше всего проследить по его блок-схеме (рис. 7).

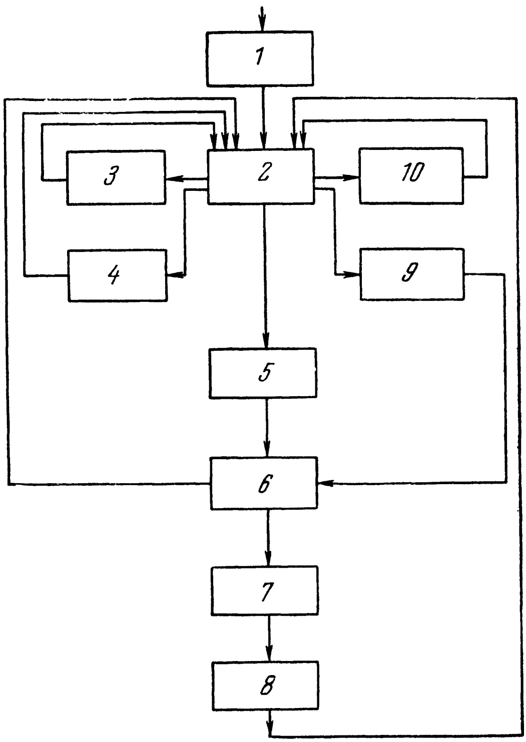

Рис. 7.

Дадим краткое описание работы каждого блока.

Блок 1 выполняет обработку данных, заключенных в «шапке», и настраивает соответственно работу прочих блоков.

Блок 2 выбирает очередную строку вычислительного плана и определяет ее тип.

Блок 3 выполняет машинную строку.

Блок 4 выполняет подпрограмму из машинных команд.

Блок 5 настраивает выполнение соответствующей полиномиальной строки (переносит на рабочее поле подпрограмму полиномиальной операции и заполняет информационную таблицу).

Блок 6 заменяет условные адреса на истинные в подпрограммах полиномиальных операций и в подсхемах, перенесенных на рабочие поля.

Блок 7 — программа полиномиальной операции, записанная в истинных адресах.

Блок 8 вносит нужные изменения в поле полиномиальных справок и в поле коэффициентов после выполнения той или иной полиномиальной операции.

Блок 9 переносит на рабочее поле соответствующую подсхему и настраивает блок 2 для последовательного перебора строк подсхемы.

Блок 10 изменяет порядок выборки строк схемы по знаку числа в соответствии со схемным переходом.

§ 5. Примеры

1. Решение граничной задачи методом наискорейшего спуска. Пусть требуется найти решение граничной задачи

$$L(y) = \frac{d}{dx}(p(x)\,y') - q(x)\,y + \varphi(x) = 0, \tag{5.1}$$

$y(x_0) = y(x_1) = 0$, при условии, что $p(x) > 0$, $q(x) \geqslant 0$.

Пусть Ω_0 — множество дважды непрерывно дифференцируемых функций, равных нулю на концах промежутка. Если $y_0 \in \Omega_0$ — начальное приближение, то по методу наискорейшего спуска следующее приближение определяется по формуле

$$y_1 = y_0 + \varepsilon_1 z_1,$$

где z_1 — решение уравнения $z'' = L(y_0)$ при краевых условиях $z(x_0) = z(x_1) = 0$,

$$\varepsilon_1 = -\frac{\int\limits_{x_0}^{x_1} z_1'^2\,dx}{\int\limits_{x_0}^{x_1} \left(qz_1^2 + pz_1'^2\right)dx}.$$

Если в уравнении (5.1) p, q, φ суть полиномы, то функция y регулярна в некотором эллипсе, содержащем промежуток (x_0, x_1). Кроме того, если p, q, φ — полиномы степени не выше m_0, и если за y_0 принять полином степени h, то $L(y_0)$ — полином степени не выше $h + m_0$, а z_1 — полином степени не выше $h + m_0 + 2$.

При выполнении каждого шага степень увеличивается не более чем на $m_0 + 2$, поэтому y_n будет полиномом степени не выше $h + n(m_0 + 2)$.

Ниже (см. табл. 5.1) приводится вычислительный план задачи $k + 0 \div k + 44$ и необходимая к нему информация $k + 45 \div k + 100$.

ВП был реализован для следующих полиномов: $p(x) = 4 + x^2$, $q(x) = 8$, $\varphi(x) = -1 - 2x - x^2$, $y_0(x) = -1 + x^2$ на промежутке $(-1, +1)$.

Предложенный вычислительный план предусматривает выдачу на печать всех последовательных приближений $y_i(x)$.

ВП не зависит от вида полиномов $p(x)$, $q(x)$, $y(x)$, $\varphi(x)$ и от промежутка интегрирования. ВП можно реализовать при произвольных заданиях a), b), d).

Для описанной в табл. 5.1 задачи «шапкой» будет

0001	0	16	$k+0$	0014	0632
0002	0	00	$k+101$	0000	$k+65$
0003	0	00	$k+56$	$k+50$	0000
0004	0	00	0000	0000	0000

В течение 3-х минут были вычислены и выданы на печать полиномы $y_i(x)$ $(i = 1, 2, 3, 4, \ldots, 17)$.

2. Решение задачи об изгибе полукруглой пластинки, жестко закрепленной по краю. Задача состоит в интегрировании бигармонического уравнения

$$\Delta^2\omega = p \tag{5.2}$$

при краевых условиях

$$\omega\Big|_S = \frac{\partial\omega}{\partial\nu}\Big|_S = 0. \tag{5.3}$$

Рассматривается полукруглая пластинка радиуса R с границей S. Часть S совпадает с осью x. Считаем, что нагрузка распределена равномерно, тогда $p = \text{const}$, и без ограничения общности можно считать, что $p = 1$. Решение задачи сводится к отысканию минимума функционала

$$F(\omega) = \int\limits_\Omega\int \{(\Delta\omega)^2 - 2p\omega\}\,dx\,dy$$

на множестве функций, удовлетворяющих краевым условиям (5.3).

Таблица 5.1

Вычислительный план задачи 1 (§ 5)

Номер ячейки	π	КОП	A_1	A_2	A_3	Содержание
$k+0$	1	05	$k+50$	0000	$k+55$	y_0'
$k+1$	5	24	$k+55$	$k+51$	$k+55$	$p \cdot y_0'$
$k+2$	5	05	$k+55$	0000	$k+55$	$\frac{d}{dx}(p \cdot y_0')$
$k+3$	1	24	$k+52$	$k+50$	$k+56$	$q \cdot y_0$
$k+4$	7	23	$k+55$	$k+56$	$k+55$	$\frac{d}{dx}(p \cdot y_0') - q \cdot y_0$
$k+5$	5	22	$k+55$	$k+53$	$k+55$	$L(y_0)$
$k+6$	5	06	$k+55$	0000	$k+55$	$z' = \int L(y_0)\,dx$
$k+7$	5	06	$k+55$	0000	$k+55$	$\bar{z} = c_2 + c_1 x + \ldots$
$k+10$	0	02	$k+46$	$k+45$	$k+62$	$x_1 - x_0$
$k+11$	0	04	0600	$k+62$	$k+62$	$\frac{1}{x_1 - x_0}$
$k+12$	1	07	$k+55$	$k+45$	$k+57$	$\bar{z}(x_0)$
$k+13$	1	07	$k+55$	$k+46$	$k+60$	$\bar{z}(x_1)$
$k+14$	0	02	$k+57$	$k+60$	$k+63$	$\bar{z}(x_0) - \bar{z}(x_1)$
$k+15$	0	05	$k+63$	$k+62$	$k+63$	c_1
$k+16$	0	05	$k+45$	$k+60$	$k+64$	$x_0 \bar{z}(x_1)$
$k+17$	0	05	$k+46$	$k+57$	$k+61$	$x_1 \bar{z}(x_0)$
$k+20$	0	02	$k+64$	$k+61$	$k+64$	$x_0 \bar{z}(x_1) - x_1 \bar{z}(x_0)$
$k+21$	0	05	$k+64$	$k+62$	$k+64$	c_2
$k+22$	1	17	$k+54$	$k+63$	$k+56$	$c_1 x$
$k+23$	5	20	$k+56$	$k+64$	$k+56$	$c_2 + c_1 x$
$k+24$	7	22	$k+55$	$k+56$	$k+55$	$z_1 = c_2 + c_1 x + \bar{z}(x)$
$k+25$	1	05	$k+55$	0000	$k+56$	z_1'
$k+26$	5	24	$k+56$	$k+56$	$k+56$	$z_1'^2$
$k+27$	1	10	$k+56$	$k+47$	$k+57$	$\int_{x_0}^{x_1} z_1'^2\,dx$
$k+30$	3	24	$k+51$	$k+56$	$k+56$	$p' \cdot z_1'^2$
$k+31$	5	10	$k+56$	$k+47$	$k+60$	$\int_{x_0}^{x_1} p z_1'^2\,dx$

Продолжение

Номер ячейки	л	КОП	A_1	A_2	A_3	Содержание
$k+32$	1	24	$k+52$	$k+55$	$k+56$	qz_1
$k+33$	3	24	$k+55$	$k+56$	$k+56$	$q \cdot z_1^2$
$k+34$	5	10	$k+56$	$k+47$	$k+61$	$\int\limits_{x_0}^{x_1} qz_1^2\,dx$
$k+35$	0	01	$k+60$	$k+61$	$k+60$	$\int\limits_{x_0}^{x_1} \left(pz'^2 + qz_1^2\right) dx$
$k+36$	0	04	$k+57$	$k+60$	$k+57$	$-\varepsilon_1$
$k+37$	5	17	$k+55$	$k+57$	$k+55$	$-\varepsilon_1 z_1$
$k+40$	5	23	$k+50$	$k+55$	$k+50$	$y_0 - \varepsilon_1 z_1 \to y_0$
$k+41$	1	13	$k+50$	0000	0000	Печать y_0
$k+42$	5	15	$k+55$	0000	$k+57$	Сравнение $\varepsilon_1 z_1$ с δ
$k+43$	1	00	0000	$k+57$	$k+0$	Схемный переход
$k+44$	0	17	0000	0000	0000	Останов
						a) Информация о промежутке интегрирования (—1, +1)
$k+45$	3	01	4000	0000	0000	«—1»
$k+46$	1	01	4000	0000	0000	«+1»
$k+47$	0	00	$k+45$	$k+46$	0000	
						b) Поле полиномиальных справок
$k+50$	4	00	$k+65$	0002	0000	$y_0(x)$
$k+51$	4	00	$k+70$	0002	0000	$p(x)$
$k+52$	4	00	$k+73$	0000	0000	$q(x)$
$k+53$	4	00	$k+74$	0002	0000	$\varphi(x)$
$k+54$	4	00	$k+77$	0001	0000	$T(x)$
$k+55$	0	00	0000	0000	0000	
$k+56$	0	00	0000	0000	0000	
						c) Поле рабочих ячеек
$k+57$	0	00	0000	0000	0000	
$k+60$	0	00	0000	0000	0000	
$k+61$	0	00	0000	0000	0000	
$k+62$	0	00	0000	0000	0000	
$k+63$	0	00	0000	0000	0000	
$k+64$	0	00	0000	0000	0000	

Продолжение

Номер ячейки	л	КОП	A_1	A_2	A_3	Содержание
d) Поле коэффициентов						
$k+65$	3	01	4000	0000	0000	$y_0(x) = -1 + x^2$
$k+66$	0	00	0000	0000	0000	
$k+67$	1	01	4000	0000	0000	
$k+70$	1	03	4000	0000	0000	$p(x) = 4 + x^2$
$k+71$	0	00	0000	0000	0000	
$k+72$	1	01	4000	0000	0000	
$k+73$	1	04	4000	0000	0000	$q(x) = 8$
$k+74$	3	01	4000	0000	0000	$\varphi(x) = -1-2x-x^2$
$k+75$	3	02	4000	0000	0000	
$k+76$	3	01	4000	0000	0000	
$k+77$	0	00	0000	0000	0000	$T(x) = x$
$k+100$	1	01	4000	0000	0000	

Задача решалась по методу Ритца, приближенное решение разыскивалось в следующем виде:

$$u(x, y) = \sum_{i=1}^{n} a_i \varphi_i (x, y).$$

В качестве координатных функций были взяты полиномы вида

$$\varphi(x, y) = x^{2k} y^l (R^2 - x^2 - y^2)^2 \qquad (k = 0, 1, 2, \ldots; \ l = 1, 2, \ldots),$$

которые удовлетворяют условиям (5.3). При данных условиях решение должно быть четным по x. Так как условия (5.3) на границе выполняются для всех φ_k, то $(\Delta^2 \varphi_i, \varphi_k) = (\Delta\varphi_i, \Delta\varphi_k)$, и система уравнений Ритца имеет вид:

$$\sum_{k=1}^{n} (\Delta\varphi_k, \Delta\varphi_i)\, a_k = p(1, \varphi_i) \qquad (i = 1, 2, \ldots, n). \tag{5.4}$$

Итак, необходимо построить систему Ритца. Матрица Ритца является симметричной, поэтому в ВП естественно предусмотреть построение лишь элементов

$$(\Delta\varphi_k, \Delta\varphi_i) = \int\limits_{\Omega}\int \Delta\varphi_k \Delta\varphi_i \, dx\, dy \qquad (k = 1, 2, \ldots, n; \ i = k, \ k+1, \ldots, n)$$

и вектора свободных членов

$$(1, \varphi_i) = \int\limits_{\Omega}\int \varphi_i \, dx\, dy \qquad (i = 1, 2, \ldots, n).$$

Ниже (см. табл. 5.2) приведен вычислительный план построения системы Ритца ($B+0 \div B+15$) и необходимая информация ($B+16 \div B+31$) к задаче.

Таблица 5.2

Вычислительный план задачи 2

Номер ячейки	π	КОП	A_1	A_2	A_3	Содержание
$B+0$	0	77	0005	$n-1$	0001	Построение вектора свободных членов
$B+1$	1	32	φ_1	$B+30$	$r+1$	
$B+2$	1	13	$B+27$	0000	0000	Печать вектора
$B+3$	0	77	0004	$n-1$	0001	Построение последовательности полиномов $\Delta\varphi_i$ $(i=1, 2, \ldots, n)$
$B+4$	1	70	φ_1	S_1	0000	
$B+5$	0	77	0003	$n-1$	0001	Построение строки матрицы Ритца
$B+6$	1	71	φ_1	φ_1	$r+1$	
$B+7$	1	13	$B+27$	0000	0000	Печать строки
$B+10$	0	33	$B+5$	(010)	$B+5$	Укорачивание строк матрицы Ритца
$B+11$	0	13	$B+6$	(110)	$B+6$	
$B+12$	0	33	$B+27$	(010)	$B+27$	
$B+13$	0	33	$B+27$	(010)	0000	Проверка на конец
$B+14$	1	00	$B+15$	$A\omega$	$B+5$	Схемный переход
$B+15$	0	17	0000	0000	0000	Останов

a) Подсхема Z построения по заданному полиному $\varphi_i(x, y)$ полинома $\Delta\varphi_i(x, y)$

$B+16$	1	25	4001	0002	4002	$\frac{\partial}{\partial x}\varphi$
$B+17$	5	25	4002	0000	4002	$\frac{\partial^2}{\partial x^2}\varphi$
$B+20$	5	26	4001	0000	4001	$\frac{\partial}{\partial y}\varphi$
$B+21$	5	26	4001	0000	4001	$\frac{\partial^2}{\partial y^2}\varphi$
$B+22$	7	01	4001	4002	4001	$\frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}$

b) Подсхема T построения элемента матрицы

$$(\Delta\varphi_k, \Delta\varphi_i) = \iint_\Omega \Delta\varphi_k \Delta\varphi_i \, dx\, dy$$

$B+23$	1	03	4001	4002	S_1	$\Delta\varphi_i \Delta\varphi_k$
$B+24$	5	32	S_1	$B+30$	4003	$\iint_\Omega \Delta\varphi_i \Delta\varphi_k \, dx\, dy$

Продолжение

Номер ячейки	π	КОП	A_1	A_2	A_3	Содержание
c) Поле справок о подсхемах Z и T						
$B+25$	0	00	$B+23$	0005	0000	
$B+26$	0	00	$B+25$	0002	0000	
d) Справка о векторе свободных членов (строке матрицы Ритца)						
$B+27$	4	00	$r+1$	$n-1$	0000	
e) Информация об области интегрирования						
$B+30$	0	00	$B+31$	$B+31$	0000	Справка
$B+31$	.	. .	. . .	. . .	. . .	Величина радиуса круга

ВП опирается на две подсхемы Z и T. Они введены с целью продемонстрировать порядок оформления подсхем. Кроме того, в ВП использованы групповые полиномиальные строки и групповые строки обращения к подсхеме.

При построении систем Ритца высокого порядка приходится заниматься построением и представлением в памяти машины большого числа координатных функций

$$\varphi(x, y) = x^{2k} y^l (R^2 - x^2 - y^2)^2 \qquad (k = 0, 1, 2, \ldots; \ l = 1, 2, \ldots). \qquad (5.5)$$

Эту часть работы также можно поручить прорабу.

В целях экономии памяти ВП построения последовательности координатных функций (5.5) по заданным полиномам

$$(R_1) = x^2,$$

$$(R_2) = y,$$

$$(S) = (R^2 - x^2 - y^2)^2$$

можно поместить в поле полиномиальных справок, причем в ячейку поля, отведенную для справки о полиноме $\varphi_l(x, y)$, поместим строку ВП, которая предписывает прорабу вычислить $\varphi_l(x, y)$, а справку о нем записать в ячейку, где находится эта строка ВП.

Ниже (см. табл. 5.3) приводится ВП построения системы координатных функций для $n = 16$.

Чтобы подготовить набор координатных функций, вычислить и выдать систему Ритца шестнадцатого порядка, прорабу требуется около 1,5 минут машинного времени.

Таблица 5.3

Номер ячейки	π	КОП	A_1	A_2	A_3	Справки полиномов
R_1	5	00	P	0002	0000	x^2
R_2	5	00	$P+6$	0001	0000	y
S	5	00	$P+11$	0004	0000	$(R^2-x^2-y^2)^2$
φ_1	1	03	S_1	R_2	φ_1	$(\varphi_1)=y(R^2-x^2-y^2)^2$
φ_2	1	03	φ_1	R_2	φ_2	$(\varphi_2)=y^2(R^2-x^2-y^2)^2$
φ_3	1	03	φ_2	R_2	φ_3	$(\varphi_3)=y^3(R^2-x^2-y^2)^2$
φ_4	1	03	φ_1	R_1	φ_4	$(\varphi_4)=x^2y(R^2-x^2-y^2)^2$
φ_5	1	03	φ_3	R_2	φ_5	$(\varphi_5)=y^4(R^2-x^2-y^2)^2$
φ_6	1	03	φ_2	R_1	φ_6	$(\varphi_6)=x^2y^2(R^2-x^2-y^2)^2$
φ_7	1	03	φ_5	R_2	φ_7	$(\varphi_7)=y^5(R^2-x^2-y^2)^2$
φ_{10}	1	03	φ_3	R_1	φ_{10}	$(\varphi_{10})=x^2y^3(R^2-x^2-y^2)^2$
φ_{11}	1	03	φ_4	R_1	φ_{11}	$(\varphi_{11})=x^4y(R^2-x^2-y^2)^2$
φ_{12}	1	03	φ_7	R_2	φ_{12}	$(\varphi_{12})=y^6(R^2-x^2-y^2)^2$
φ_{13}	1	03	φ_5	R_1	φ_{13}	$(\varphi_{13})=x^2y^4(R^2-x^2-y^2)^2$
φ_{14}	1	03	φ_6	R_1	φ_{14}	$(\varphi_{14})=x^4y^2(R^2-x^2-y^2)^2$
φ_{15}	1	03	φ_{12}	R_2	φ_{15}	$(\varphi_{15})=y^7(R^2-x^2-y^2)^2$
φ_{16}	1	03	φ_7	R_1	φ_{16}	$(\varphi_{16})=x^2y^5(R^2-x^2-y^2)^2$
φ_{17}	1	03	φ_{10}	R_1	φ_{17}	$(\varphi_{17})=x^4y^3(R^2-x^2-y^2)^2$
φ_{20}	1	03	φ_{11}	R_1	φ_{20}	$(\varphi_{20})=x^6y(R^2-x^2-y^2)^2$

ГЛАВА II

ПРОРАБ N

Незначительные изменения в структуре некоторых блоков прораба I—II привели к тому, что появился вариант, позволяющий выполнять на машине М-20 выкладки с алгебраическими полиномами N независимых переменных и с объектами более сложной структуры.

§ 6. Исходный класс объектов

Введем в рассмотрение следующий класс объектов

$$P(x_0, x_1, x_2, \ldots, x_{N-1}) = \sum_{i=1}^{n} a_i x_0^{i_0} x_1^{i_1} \ldots x_{N-1}^{i_{N-1}}. \tag{6.1}$$

Показатель степени i_k может быть числом целым (положительным и отрицательным) или нулем. В выражении (6.1) все коэффициенты $a_i \neq 0$. Кроме того, в (6.1) выполнено приведение подобных.

Объекты (6.1) представляются в памяти машины записью и справкой. Будем записывать каждый одночлен

$$a_i x_0^{i_0} x_1^{i_1} \ldots x_{N-1}^{i_{N-1}} \tag{6.2}$$

в l подряд расположенных ячейках памяти. Эту группу из l ячеек назовем «ящиком». Запишем в первые $l-1$ ячейки ящика набор показателей степени $i_0, i_1, \ldots, i_{N-1}$ одночлена (6.2), а в последнюю ячейку ящика запишем коэффициент a_i.

Порядок записи показателей i_k в ячейках ящика должен быть вполне определенным и должен соответствовать порядку записи переменных x_k в одночленах (6.2).

Каждую из $l-1$ первых ячеек ящика разобьем на 5 групп по 9 разрядов в каждой группе. Перенумеруем группы следующим образом:

0	1	2	3	4
5	6	7	8	9
10	11	12	13	14
....	......	......	......	......
$5(l-2)$	$5(l-2)+1$	$5(l-2)+2$	$5(l-2)+3$	$5(l-2)+4$

В группе разрядов с номером k всегда будем записывать показатель i_k с его знаком.

Такое разбиение ячеек приводит к следующему ограничению на размеры допустимых в (6.1) показателей

$$-256_{10} \leqslant i_k < 256_{10}.$$

В группе разрядов положительный показатель i_k представляется прямым кодом, отрицательный показатель представляется дополнительным до 1000_8 кодом.

Например, показатели $i_k = 5$, $i_k = 12$, $i_k = -4$ представляются соответственно кодами 005, 014, 774.

Длина ящика, необходимого для записи одночлена (6.2), зависит от N — числа переменных:

$$l = 1 + \left[\frac{N}{5}\right] + \begin{cases} 0, & \text{если} \quad \left\{\dfrac{N}{5}\right\} = 0, \\ 1, & \text{если} \quad \left\{\dfrac{N}{5}\right\} \neq 0. \end{cases}$$

Здесь $\left[\frac{N}{5}\right]$ — целая часть, $\left\{\frac{N}{5}\right\}$ — дробная часть числа $\frac{N}{5}$.

Весь объект (6.1) для своей записи требует n ящиков, т. е. nl ячеек, расположенных подряд.

Справка об объекте (6.1) записывается в одну ячейку. Она содержит следующие сведения:

1) P_1 — адрес первой ячейки массива, отведенного под запись,
2) n — количество одночленов,
3) p — признак справки, $p = 4$,

4	—	P_1	n	0000

Содержание КОП в справке не имеет значения.

Пример 1. Пусть требуется записать в памяти машины следующее выражение:

$$P(x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7) =$$
$$= 2\frac{x_0^2 x_2 x_4 x_5^3}{x_1 x_3^3 x_6^2 x_7} + 6\frac{x_0 x_1^4 x_2 x_3 x_4^4 x_7}{x_6^3 x_5^2} + 4x_0 x_1^5 - 3\frac{x_2 x_3^2 x_4^3 x_7^2}{x_0^2}.$$

Длина одного ящика $l = 3$, для записи всего выражения требуется 12 ячеек.

P_1	0	02	777	0 01	77	5	001	
P_1+1	0	03	776	7 77	00	0	000	$2\dfrac{x_0^2 x_2 x_4 x_5^3}{x_1 x_3^3 x_6^2 x_7}$
P_1+2	1	02	400	0 00	00	0	000	
P_1+3	0	01	004	0 01	00	1	004	
P_1+4	7	76	775	0 01	00	0	000	$6\dfrac{x_0 x_1^4 x_2 x_3 x_4^4 x_7}{x_6^3 x_5^2}$
P_1+5	1	03	600	0 00	00	0	000	
P_1+6	0	01	005	0 00	00	0	000	
P_1+7	0	00	000	0 00	00	0	000	$4x_0 x_1^5$
P_1+10	1	03	400	0 00	00	0	000	
P_1+11	7	76	000	0 01	00	2	003	
P_1+12	0	00	000	0 02	00	0	000	$-3\dfrac{x_2 x_3^2 x_4^3 x_7^2}{x_0^2}$
P_1+13	3	02	600	0 00	00	0	000	

Справка имеет вид:

4	00	P_1	0004	0000

Справки об объектах исходного класса располагаются в поле $R_1 \div R_t$. Записи объектов, т. е. показатели степеней и коэффициенты, располагаются в числовом поле $P \div P_s$.

§ 7. Операции

Для проведения на машине с помощью прораба N выкладок с объектами (6.1) введем в рассмотрение набор операций. В него входят операции сложения, вычитания, умножения, дифференцирования по переменной x_i $(i = 0, 1, \ldots, N-1)$ и т. д. В список операций входят операции специального назначения: схемные переходы, печать, гашение (стирание) и т. д.

Ниже (см. табл. 7.1) приведен список имеющихся операций. Этот список может дополняться новыми операциями. При составлении новых операций следует придерживаться тех же правил, что и в § 2.

Операции 07 и 10 в таблице 7.1 суть операции, рассчитанные на полиномы трех независимых переменных.

Операция представляется в памяти машины подпрограммой в условных адресах и справкой о местоположении подпрограммы.

При составлении подпрограмм новых операций следует пользоваться приведенной в § 2 формулой вычисления условного адреса и системой поправок Δ.

Остановимся еще раз на условных номерах, присвоенных входным данным для подпрограмм операций:

$C(P_1) = 0001$, P_1 — начало массива, отведенного для записи первого аргумента;
$C(n_1) = 0002$, n_1 — количество одночленов у первого аргумента;
$C(n_1 l) = 0003$, $n_1 l$ — размеры массива, отведенного для записи первого аргумента;
$C(P_2) = 0005$, P_2 — начало массива, отведенного для записи второго аргумента;
$C(n_2) = 0006$, n_2 — количество одночленов у второго аргумента;
$C(n_2 l) = 0007$, $n_2 l$ — размеры массива, отведенного для записи второго аргумента;
$C(P_{\text{св}}) = 0011$, $P_{\text{св}}$ — номер первой свободной ячейки в поле записей;
$C(l) = 0012$, l — длина ящика;
$C(0) = 0013$, 0 — нуль;
$C(B+0) = 0015$, $B+0$ — начало рабочего поля подпрограмм операций;
$C(B+M) = 0016$, $B+M-1$ — конец рабочего поля подпрограмм операций;
$C(R_3) = 0017$, R_3 — адрес справки результата или адрес результата.

Заметим, что все сказанное о построении ВП задачи для прораба I-II в § 3 остается в силе для прораба N.

Таблица 7.1

Список операций для прораба N

Номер операции	Название	Справки					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
01	Сложение	1	01	S_1+1	0033	0000	1	01	R_1	R_2	R_3	$(R_3)=(R_1)+(R_2)$
02	Вычитание	1	00	S_1+0	0034	0000	1	02	R_1	R_2	R_3	$(R_3)=(R_1)-(R_2)$
03	Гашение (стирание)	0	00	S_2+0	0000	0000	5 7	03 03	R_1 R_1	0000 R_2	0000 0000	Стирается (R_1) Стираются (R_1) и (R_2)
04	Дифференцирование по переменной	1	00	S_3+0	0037	0000	1	04	R_1	A	R_3	$(R_3)=\frac{\partial}{\partial x_s}(R_1)$, где $(A)=(4,\ 00,\ 0000,\ s,\ 0000)$
05	Печать	0		S_4+0	0007	0000	1	05	R_1	0000	0000	
06	Умножение	1	00	S_5+0	0110	0000	1	06	R_1	R_2	R_3	$(R_3)=(R_1)\cdot(R_2)$
11	Перенос	0	00	S_6+0	0004	0000	1	11	R_1	0000	R_3	$(R_3)=(R_1)$
12	Умножение на число	1		S_7+0	0004	0000	1	12	R_1	A	R_3	$(R_3)=a\cdot(R_1)$, где $a=(A)$.

Продолжение

Номер операции	Название	Справки					Запись полиномиальной строки					Примечание
		π	КОП	A_1	A_2	A_3	π	КОП	A_1	A_2	A_3	
13	Выделение p-го одночлена в самостоятельный полином, не зависящий от переменной x_s	0	00	$S_{10}+0$	0023	0000	1	13	R_1	A	R_3	(R_3) — p-й одночлен (R_1) $(A) = (4,\ 00,\ p,\ s,\ 0000)$
14	Подстановка $x_s = \text{const}$	0	00	$S_{11}+0$	0035	0000	1	14	R_1	A	R_3	$(R_3) = (R_1)_{x_s=\text{const}}$ $(A) = \text{const}$
07	Интегрирование по эллипсоиду	0	00	$S_{12}+0$	0037	0000	1	07	R_1	A	B	$(B) = \int\int_{\Omega}\int (R_1)\,dx\,dy\,dz$ $(A) = (0,\ 00,\ m,\ n,\ p)$ $(m) = a,\ (n) = b\ (p) = c$
10	Интегрирование по параллелепипеду	0	00	$S_{13}+0$	0033	0000	1	10	R_1	A	B	$B = \int\int_{\Omega}\int (R_1)\,dx\,dy\,dz$ $(A) = (0,\ 00,\ m,\ n,\ p)$ $(m) = a,\quad (m+1) = b,$ $(n) = c,\quad (n+1) = d,$ $(p) = e,\quad (p+1) = f$

§ 8. Прораб N и автоматическое распределение памяти

Программа прораба N содержит 470_{10} команд. К программе прораба прилагается набор подпрограмм операций В сумме эти подпрограммы содержат 320_{10} команд.

Порядок оформления задачи для прораба N таков же, как и для прораба I-II (см. § 4 пп. 1—4)

В шапку задачи, кроме перечисленной информации (§ 4 п. 5), включается также число N независимых переменных у объектов, с которыми имеет дело алгоритм задачи.

Шапка задачи имеет вид:

Номер ячейки	π	КОП	A_1	A_2	A_3
0001	0	16	M_{cx}	0120	0654
0002	p	00	P	0000	$P_{св}$
0003	0	00	R_1	R_t	N
0004	0	00	$L+0$	M	0000

Здесь M_{cx} — номер начальной строки ВП; $P \div P_{св} - 1$ — часть поля, занятая записями заданных объектов; $R_1 \div R_t$ — поле справок об объектах; N — количество независимых переменных; $L+0$ — начало поля справок о подсхемах; M — максимальная длина подсхем; $p = 1$ — отметка о необходимости перевода из десятичной системы в двоичную коэффициентов объектов, записанных в ячейках $P+l-1$, $P+2l-1$, ..., $P_{св}-1$; $p=0$ — отметка о том, что перевода коэффициентов из десятичной системы в двоичную не требуется.

Обрабатывая информацию, заключенную в шапке, прораб настраивается на работу по ВП, причем по заданному N получается число l — длина ящика и формируются константы, которые используются подпрограммами операций, а именно:

0021	0	00	0000	0000	l
0022	0	00	0000	l	0000
0023	0	00	0000	l	l
0024	0	00	l	0000	0000
0025	0	00	l	0000	l
0026	0	00	l	l	0000
0027	0	00	l	l	l

Закончив настройку, прораб приступает к выполнению ВП. Он перебирает строки ВП в предписанном порядке, выполняет их, по полученным результатам выбирает направление дальнейших вычислений, размещает в памяти вычисленные объекты, стирает ненужные в дальнейшем.

§ 9. Примеры

1. Обращение рядов. Требуется обратить ряд

$$y = a_1x + a_2x^2 + \ldots + a_nx^n + \ldots \tag{9.1}$$

Известно, что коэффициенты ряда

$$x = b_1y + b_2y^2 + \ldots + b_ny^n + \ldots \tag{9.2}$$

определяются по формуле

$$b_n = \frac{1}{n!}\left[\frac{d^{n-1}}{dx^{n-1}}\left(\frac{1}{\psi(x)}\right)^n\right]_{x=0}, \tag{9.3}$$

где

$$\psi(x) = a_1 + a_2x + a_3x^2 + \ldots + a_nx^{n-1} + \ldots \tag{9.4}$$

Отрезок ряда (9.4)

$$\psi_n(x) = \sum_{i=1}^{n} a_i x^{i-1} \tag{9.5}$$

будем рассматривать как полином $n+1$ независимых переменных

$$\psi_n(x_0, x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} 1 \cdot x_{i-1}x_n^{i-1}, \tag{9.5'}$$

где

$$x_{i-1} = a_i, \qquad x_n = x.$$

Для записи одночлена

$$1 \cdot x_{i-1}x_n^{i-1} \tag{9.6}$$

необходим ящик, состоящий из l ячеек, где

$$l = 1 + \left[\frac{n+1}{5}\right] + \begin{cases} 0, & \text{если } \left\{\frac{n+1}{5}\right\} = 0, \\ 1, & \text{если } \left\{\frac{n+1}{5}\right\} \neq 0. \end{cases}$$

Для записи отрезка ряда (9.5) требуется nl ячеек памяти.

Пусть необходимо построить выражения коэффициентов (9.3) через буквенные коэффициенты ряда (9.1) для $n = 19$. Тогда для записи одночлена (9.6) потребуется

$$l = 1 + \frac{20}{5} = 5$$

ячеек памяти. Для записи всего полинома (9.5) необходимо $nl = 95$ ячеек.

Ниже (см. табл. 9.1) приводится вычислительный план последовательного построения и выдачи на печать коэффициентов

$$b_k = \frac{R(a_1, a_2, \ldots, a_k)}{T(a_1)} \qquad (k = 1, 2, \ldots, 19), \tag{9.7}$$

где R и T есть полиномы. На печать выдаются последовательно полиномы R.

В образовании выражения (9.7) участвуют лишь k первых одночленов (9.5), поэтому естественно прорабу работать не со всем полиномом, а с его куском

$$\psi_k(x_0, x_1, \ldots, x_{k-1}, x_n) = \sum_{i=1}^{k} 1 \cdot x_{i-1}x_n^{i-1}.$$

Это достигается следующим образом. Для получения b_k справка о полиноме ψ должна содержать в A_2 число k. Переходя от получения b_k к получению b_{k+1}, следует изменить справку о ψ так, чтобы в A_2 стояло число $k+1$. Это предусмотрено в приведенном ниже ВП (см. строку $k+26$).

Таблица 9.1

Вычислительный план задачи 1

Номер ячейки	π	КОП	A_1	A_2	A_3	Содержание
$k+1$	0	00	0000	0000	n	
$k+2$	0	13	ψ	0572	ψ_1	
$k+3$	1	04	ψ	$k+33$	S_1	ψ'
$k+4$	1	11	[1]	0000	R	$(R)=1$
$k+5$	0	00	«1»	0000	$k+34$	Построение полинома R
$k+6$	0	00	n	0000	$k+35$	
$k+7$	0	00	n	0000	n_0	
$k+10$	0	01	$k+34$	«1»	$k+34$	
$k+11$	0	02	$k+35$	«1»	$k+35$	
$k+12$	1	00	$k+24$	$A\omega$	$k+13$	
$k+13$	1	04	R	$k+33$	S_2	
$k+14$	5	06	S_2	ψ	S_2	
$k+15$	5	06	R	S_1	R	
$k+16$	0	01	n_0	«1»	n_0	
$k+17$	5	12	R	n_0	R	
$k+20$	7	02	S_2	R	R	
$k+21$	0	04	«1»	$k+34$	β	
$k+22$	5	12	R	β	R	
$k+23$	1	00	$k+10$	0000	$k+10$	
$k+24$	5	14	R	$k+33$	R	
$k+25$	7	05	R	S_1	0000	Печать R
$k+26$	0	01	n	«1»	n	
$k+27$	0	02	$k+32$	n	0000	
$k+30$	1	00	$k+31$	$A\omega$	$k+2$	
$k+31$	0	17	0000	0000	0000	Останов
a) Константы						
$k+32$	1	05	4400	0000	0000	
$k+33$	4	00	0000	0023	0000	
b) Рабочие ячейки						
$k+34$	0	00	0000	0000	0000	
$k+35$	0	00	0000	0000	0000	
c) Поле полиномиальных справок						
$k+36$	4	00	$k+44$	0001	0000	[1]
$k+37$	4	00	$k+51$	0000	0000	ψ
$k+40$	0	00	0000	0000	0000	R
$k+41$	0	00	0000	0000	0000	S_1
$k+42$	0	00	0000	0000	0000	S_2
$k+43$	0	00	0000	0000	0000	

Продолжение

Номер ячейки	π	КОП	A_1	A_2	A_3	Содержание
			d) Поле записей			
$k+44$	0	00	0000	0000	0000	
$k+45$	0	00	0000	0000	0000	
$k+46$	0	00	0000	0000	0000	1
$k+47$	0	00	0000	0000	0000	
$k+50$	1	01	4000	0000	0000	
$k+51$	0	01	0000	0000	0000	
$k+52$	0	00	0000	0000	0000	
$k+53$	0	00	0000	0000	0000	a_1
$k+54$	0	00	0000	0000	0000	
$k+55$	1	01	4000	0000	0000	
$k+56$	0	00	0010	0000	0000	
$k+57$	0	00	0000	0000	0000	
$k+60$	0	00	0000	0000	0000	a_2x
$k+61$	0	00	0000	0000	0001	
$k+62$	1	01	4000	0000	0000	
$k+63$	0	00	0000	0100	0000	
$k+64$	0	00	0000	0000	0000	
$k+65$	0	00	0000	0000	0000	a_3x^2
$k+66$	0	00	0000	0000	0002	
$k+67$	1	01	4000	0000	0000	
. . . .	. . . .	. . . .	. . . .	. . . .	. . . .	
. . . .	. . . .	. . . .	. . . .	. . . .	. . . .	
. . . .	. . . .	. . . .	. . . .	. . . .	. . . .	
$k+173$	0	00	0000	0000	0000	
$k+174$	0	00	0000	0000	0000	
$k+175$	0	00	0000	0000	0000	$a_{19}x^{18}$
$k+176$	0	00	0000	0000	1022	
$k+177$	1	01	4000	0000	0000	
			«Шапка» задачи имеет вид			
0001	0	16	$k+1$	0120	0654	
0002	0	00	$k+44$	0000	$k+200$	
0003	0	00	$k+36$	$k+43$	0024	
0004	0	00	0000	0000	0000	

2. Подстановка полинома в полином. Пусть требуется в полином $R(x_0, x_1, \ldots, x_{N-1})$ вместо переменной x_i подставить полином $\bar{R}(x_0, x_1, \ldots, x_{N-1})$. В результате подстановки получится полином $\bar{\bar{R}}(x_0, x_1, \ldots, x_{N-1})$.

Ниже (см. табл. 9.2) приведен вычислительный план этой задачи, записанный условно. Таким образом, ВП подстановки полинома в полином оформлен как подсхема.

Таблица 9.2

Подсхема подстановки полинома в полином

Номер ячейки	π	КОП	A_1	A_2	A_3
$Д+0$	0	52	7212	0001	4517
$Д+1$	0	55	0005	0611	4477
$Д+2$	1	00	4277	0742	4077
$Д+3$	0	33	0657	4477	0000
$Д+4$	1	00	4137	0742	4217
$Д+5$	2	52	0000	0001	0000
$Д+6$	0	33	4477	0660	4477
$Д+7$	1	00	4077	0000	4077
$Д+10$	0	45	0733	4477	4477
$Д+11$	0	14	0130	4477	4477
$Д+12$	0	13	4337	4477	4337
$Д+13$	1	02	4001	4001	4005
$Д+14$	4	00	7211	0000	4477
$Д+15$	0	54	0100	4477	4477
$Д+16$	0	55	4477	1470	4477
$Д+17$	1	13	4001	4457	4007
$Д+20$	0	73	4477	0736	4477
$Д+21$	1	00	4577	0742	4537
$Д+22$	4	00	0001	4002	0000
$Д+23$	0	00	0000	0000	0000
$Д+24$	0	00	0000	0000	0000
$Д+25$	5	06	4007	4003	4007
$Д+26$	1	00	4417	0000	4417
$Д+27$	7	01	4005	4007	4005
$Д+30$	0	13	4317	0024	4317
$Д+31$	0	13	4457	0574	4457
$Д+32$	0	33	4517	0572	4517
$Д+33$	1	00	4717	0742	4317
$Д+34$	1	05	4005	0000	0000
В подсхеме используются следующие константы прораба:					
0611	0	00	0000	7777	0000
0660	0	00	0000	0005	0000
1470	7	77	0000	0000	0000
0736	0	01	0000	0000	0000
0024	0	00	l	0000	0000
0574	0	00	0001	0000	0000
0572	0	00	0000	0001	0000

Если алгорифм решения некоторой задачи таков, что неоднократно требуется подстановка полинома в полином, то достаточно в ВП задачи вставить строки обращения к этой подсхеме:

Номер ячейки	π	КОП	A_1	A_2	A_3
$\varkappa+1$	5	$70+j$	R	i	$\overline{R}$
$\varkappa+2$	0	00	$\overline{\overline{R}}$	0000	S
$\varkappa+3$	4	00	P	n	0000

Здесь R — номер, $\varkappa+3$ — справка исходного полинома; i — номер переменной; $\overline{R}$ — номер подставляемого полинома; $\overline{\overline{R}}$ — номер результата; S — номер рабочего полинома.

В ВП допускаются строки обращения к подсхеме подстановки полинома в полином, причем полином, в который подставляется другой полином, сам может получаться в результате выполнения ряда строк ВП. О таком полиноме, кроме номера его справки при составлении ВП, нам ничего неизвестно. Поэтому, составляя ВП, который предписывал бы прорабу, кроме всего прочего, подстановку полинома в полином, необходимо перед выполнением строки обращения предусмотреть выполнение машинной строки переноса справки исходного полинома R в ячейку $\varkappa+3$ строки обращения.

Итак, в результате выполнения следующих строк ВП:

Номер ячейки	π	КОП	A_1	A_2	A_3
$\varkappa+0$	0	00	R	0000	$\varkappa+3$
$\varkappa+1$	5	$70+j$	R	i	$\overline{R}$
$\varkappa+2$	0	00	$\overline{\overline{R}}$	0000	S
$\varkappa+3$	0	00	0000	0000	0000

будет получен полином

$$\overline{\overline{R}}(x_0, x_1, \ldots, x_{N-1}) = R(x_0, x_1, \ldots, x_{N-1})\, x_i = \overline{R}(x_0, x_1, \ldots, x_{N-1}).$$

В ячейке 0742 прораб N хранит число, поставленное в соответствие выработанному сигналу ω. Здесь имеется в виду сигнал ω, выработанный при выполнении машинной строки ВП (см. § 3).

Программа прораба I-II, программа прораба N, а также наборы подпрограмм операций к ним находятся в Ленинградском отделении Математического института АН СССР (ЛОМИ).

Автор пользуется случаем выразить глубокую благодарность В. С. Сохранской, подготовившей наборы подпрограмм операций для прорабов на машину М-20.

ЛИТЕРАТУРА

Ахиезер Н. И. 1. Лекции по вариационному исчислению, М., Гостехиздат, 1955.

Ахиезер Н. И., Глазман И. М. 1. Теория линейных операторов, М.—Л., Гостехиздат, 1950.

Бабич В. М. 1. К вопросу о распространении функций, УМН **8**, вып. 2 (54) (1953).

Бергер (Berger H. M.) 1. A new approach to the analysis of large deflexions of plates, Journ. of Appl. Mech. **22**, № 4 (1955), 465—472.

Бернштейн С. Н. 1. Об уравнениях вариационного исчисления, УМН, вып. VIII (1941), 32—74 (перевод статьи «Sur les équations du calcul des variations», Ann. Sci. École Norm. Sup. **29** (1912), 431—485).

Бирман М. Ш. 1. О методе Фридрихса расширения положительно определенного оператора до самосопряженного, Зап. Ленингр. горн. ин-та **33**, вып. 3 (1956).

Богарян О. К. 1. О сходимости невязки методов Бубнова — Галёркина и Ритца, ДАН **141**, № 2 (1961), 267—269.

Булавский В. А. 1. О символике записи вычислительных знаков при автоматизации программирования, Изв. ВУЗ, Матем. **5** (1958).

Вазов В., Форсайт Дж. 1. Разностные методы решения дифференциальных уравнений в частных производных, М., ИЛ, 1963.

Вайль, Ньюмарк. 1. Weil N. A., Newmark N. M., Large deflexions of elliptical plates, Paper Amer. Soc. Engrs № A—2 (1955).

Вайнберг М. М. 1. Вариационные методы исследования нелинейных операторов, М., Гостехиздат, 1956.

Вайникко Г. М. 1. Необходимое и достаточное условие устойчивости метода Галёркина — Петрова, Уч. зап. Тартуск. ун-та (1965).

Ватсон Г. Н. Теория бесселевых функций, ч. 1, М., ИЛ, 1949.

Налиев М. А. 1. Исследование устойчивости метода Бубнова — Галёркина для нестационарных задач, ДАН **157**, № 1 (1964), 16—19; поправка: ДАН **161**, № 2 (1965).

2. Об устойчивости метода Бубнова — Галёркина для нестационарных задач, сб. «Вопросы вычисл. матем. и вычисл. техн.», изд. АН АзербССР, вып. 3 (1964).

Вержбинская Ю. С. 1. О выборе координатных функций в методе Ритца для обыкновенных дифференциальных операторов, заданных на бесконечном промежутке, сб «Методы вычислений», изд. ЛГУ (1966).

Винокуров С. Г. 1. Применение метода Галёркина к решению задачи о больших прогибах круглой шарнирно опертой пластинки. Изв. Казанск. фил. АН СССР, сер. физ.-матем. и техн. наук, № 10 (1956), 57—61.

Вишик М. И. 1. Задача Коши для уравнений с операторными коэффициентами, смешанная краевая задача для систем дифференциальных уравнений и приближенный метод их решения, Матем. сб. **39 (81)**: 1 (1956), 51—148.

Власов В. З. 1. Общая теория оболочек, М.—Л., Гостехиздат, 1941.

Власова З. А. 1. Метод сеток для нелинейной одномерной вариационной задачи, Труды Матем. ин-та им. В. А. Стеклова **66** (1962).

Власова З. А., Рыбакова Ю. В. 1. К вопросу о численной реализации метода сеток для нелинейной одномерной вариационной задачи, Труды Матем. ин-та им. В. А. Стеклова **84** (1965).

Вольмир А. С. 1. Гибкие пластинки и оболочки, М., Гостехиздат, 1956.

Ворович И. И. 1. О некоторых прямых методах в нелинейной теории пологих оболочек, ДАН **105**, № 1 (1955), 42—45.

2. О существовании решения в нелинейной теории оболочек, Изв. АН, сер. матем. **19** (1955), 173—186.

3. О методе Бубнова — Галёркина в нелинейной теории колебаний пологих оболочек, ДАН **110**, № 5 (1956), 723—726.

4. О некоторых прямых методах в нелинейной теории оболочек, Прикл. матем. и мех. **20**, вып. 4 (1956), 449—474.

5. О существовании решений в нелинейной теории оболочек, ДАН **117**, № 2 (1957), 203—206.

6. Погрешность прямых методов в нелинейной теории оболочек, ДАН **122**, № 2 (1958), 196—199.

Галимов К. Э. 1. К вариационным принципам нелинейной теории упругости, Уч. зап. Казанск. ун-та **113**, № 10 (1953), 155—160.

2. О некоторых вариационных формулах нелинейной теории упругости, Уч. зап. Казанск. ун-та **115**, № 12 (1955), 111—118.

3. К вариационным методам решения задач нелинейной теории пластин и оболочек, Уч. зап. Казанск. ун-та **116**, № 1 (1956), 36—40.

4. Об одном методе решения краевых задач нелинейных уравнений теории пологих оболочек, Уч. зап. Казанск. ун-та **116**, № 5 (1956), 19—26.

5. О больших прогибах прямоугольной цилиндрической панели, Инж. сб. **25** (1959), 20—36.

Гантмахер Ф. Р. 1. Теория матриц, М., Гостехиздат, 1953.

Гельман И. В. 1. К задаче о минимуме нелинейного функционала, Уч. зап. Ленингр. пед. ин-та **166** (1958), 255—263.

Гельфанд И. М., Фомин С. В. 1. Вариационное исчисление, М., Физматгиз, 1961.

Глазман И. М. 1. Прямые методы качественного спектрального анализа сингулярных дифференциальных операторов, М., Физматгиз, 1963.

Горлов А. М. 1. Применение вариационного метода к решению задач упруго-пластического кручения стержней, Труды Моск. ин-та инж. ж.-д. трансп., вып. 122 (1959), 407—419.

Грёгер (Gröger Konrad). 1. Einführung und Anwendunge Sobolewscher Räume für beliebige Gebiete, Math. Nachr. (1965).

2. Nichtlineare ausgeartete elliptische Differentialgleichungen, Math. Nachr. (1965).

Гусева О. В. 1. О краевых задачах для сильно эллиптических систем, ДАН **102**, № 6 (1955), 1069—1072.

Давиденко Д. Ф. 1. Об одном новом методе численного решения систем нелинейных уравнений, ДАН **88**, № 4 (1953), 601—603.

Даугавет И. К. 1. О быстроте сходимости метода Галёркина для обыкновенных дифференциальных уравнений, Изв. ВУЗ, Матем. **5** (1958)

2. О методе моментов для обыкновенных дифференциальных уравнений. Сиб. матем. журн **6**, № 1 (1965).

Джексон Д. 1. Ряды Фурье и ортогональные полиномы, М., ИЛ, 1948

Довбыш (Гаген-Торн) Л. Н. 1. О разрешимости систем Ритца для функционалов теории пластичности, Труды Матем. ин-та им. В. А. Стеклова, № 66 (1962), 190—195.

2. Устойчивость метода Ритца для задач спектральной теории операторов. Труды Матем. ин-та им. В. А. Стеклова **84** (1965).

Довбыш (Гаген-Торн) Л. Н., Михлин С. Г. 1. О разрешимости нелинейных систем Ритца, ДАН **138**, № 2 (1961), 258—260.

Дракер (Drucher D. C.) 1. Variational principles in the mathematical theory of plasticity, Proc. Sympos. Appl. Math. **8** (1958), 7—22.

Жиро (Giraud G.) 1. Sur differentes questions relatives aux équations du type elliptique, Ann. Sci. Ecole Norm. Sup. **47** (1930).

Зигмунд А. 1. Тригонометрические ряды, т. 1, М., изд. «Мир», 1965.

Казимиров В. И. 1. О полунепрерывности интегралов вариационного исчисления, УМН **11**, вып. 3 (69) (1956), 125—130.

Календерьян Л. И. 1. Применение вариационных методов в нелинейной теории изгиба пластин, Научн. Труды Одесск. ин-та инж. морск. флота, вып. 11 (1954), (1955), 69—95

Канторович Л. В. 1. Об одной математической символике, удобной при проведении вычислений на машине, ДАН **113**, № 4 (1957).

2. О проведении аналитических выкладок на машине с программным управлением, Изв. АН Арм.ССР **10**, № 2 (1957).

Канторович Л. В., Акилов Г. П. 1. Функциональный анализ в нормированных пространствах, М, Физматгиз, 1959.

Канторович Л. В., Крылов В. И. 1. Приближенные методы высшего анализа, М.—Л., Гостехиздат, 1949.

Канторович Л. В., Петрова Л. Т. 1. О математической символике, удобной при вычислениях на машинах, Труды III Всесоюзн. матем. съезда, II, Изд. АН СССР, 1956.

Качанов Л. М. 1. Некоторые вопросы теории ползучести, М.—Л., Гостехиздат, 1949.

2. Расчет прочности лопасти водяной турбины, в сб. «Вопросы прочности лопасти водяной турбины», Изд. ЛГУ (1954), 93—173.

3. Основы теории пластичности, М., Гостехиздат, 1956.

4. О вариационных методах решения задач теории пластичности, Прикл. матем. и мех. **23**, вып. 3 (1959).

5. Теория ползучести, М., Физматгиз, 1960.

6. Пример решения вариационным методом задачи упруго-пластического кручения, Исслед. по упругости и пластичности, сб. № 1, Изд. ЛГУ (1961).

Качмаж С., Штейнгауз Г. 1. Теория ортогональных рядов, М., Физматгиз, 1958.

Кирия В. С. 1. Движение тел в сопротивляющихся средах, Труды Тбил. ун-та **44** (1951), 1—20.

Клюшников В. Д. 1. Изгиб прямоугольных пластинок с учетом больших прогибов. Инж. сб. **24** (1956), 62—72.

Корнишин М. С., Муштари Х. М. 1. Об одном алгоритме решения нелинейных задач теории пологих оболочек, Прикл. матем. и мех. **23**, № 1 (1959), 159—163.

Кошелев А. И. 1. Априорные оценки в L_p и обобщенные решения эллиптических уравнений и систем, УМН **13**, вып. 4 (82) (1958), 29—88.

Кравчук М. 1. Застосування способу моменів до розв'язання лінійних диференціальних та інтегральних рівнянь, Видавництво Укр. Акад. наук, Київ, 1936.

Красносельский М. А. 1. Топологические методы в теории нелинейных интегральных уравнений, М., Гостехиздат, 1956.

Курант Р., Гильберт Д. 1. Методы математической физики, т. I, М.—Л., ОНТИ, 1933.

2. Методы математической физики, т. II, М.—Л., Гостехиздат, 1945.

Ладыженская О. А. 1. Смешанная задача для гиперболического уравнения, М., Гостехиздат, 1953

2. Об интегральных оценках сходимости приближенных методов и решений в функционалах для линейных эллиптических операторов, Вестн. ЛГУ, № 7, сер матем., мех. и астр., вып. 2 (1958).

Лангенбах А. 1. О применении вариационного принципа к некоторым нелинейным вариационным уравнениям, ДАН **121**, № 2 (1958), 214—217.

2. Langenbach A. Variationsmethoden in der nichtlinearen Elastizität — und Plastizitätstheorie, Wiss. Z. Humbodt — Univ., Berlin, Math. — Nat. R., IX, (1959/1960).

3. О некоторых нелинейных операторах теории упругости в гильбертовом пространстве, Вестн. ЛГУ, № 1, сер. матем., мех. и астр., вып. I (1961), 38—50.

4. Elastisch-plastische Deformationen von Platten, Zeitschr. f. angew. Math. u. Mech. **41**, H. 3 (1961), 126—134.

5. Zur Lösung eines Minimum-Problems der nichtlinearen Plattentheorie, Math. Nachr. (1965).

Левин (Lewin S.) 1. Über einige mit der Konvergenz im Mittel verbundenen Eigenschaften von Funktionalfolgen, Math. Zeitschr. **32**, H. 4 (1930).

Лозинский С. М. 1. Обратные функции, неявные функции и решение уравнений, Вестн. ЛГУ, № 7, сер матем., мех. и астр., вып. 2 (1957), 131—142.

Миранда К. 1. Уравнения с частными производными эллиптического типа, М., ИЛ, 1957.

Миткевич В. М. 1. Применение метода Ритца к задаче об изгибе консольной секторной плиты, Сб. трудов Лабор. гидравл. машин АН УССР, вып. 9 (1961), 48—57.

Михлин С. Г. 1. Сингулярные интегральные уравнения, УМН **3**, вып. 3 (25) (1948), 29—112.

2. Метод наименьших квадратов в задачах математической физики, Уч. зап. ЛГУ, № 111, сер. матем. наук, вып. 16 (1949).

3. Интегральные уравнения и их приложения, М.—Л., Гостехиздат, 1949.

4. Прямые методы в математической физике, М.—Л., Гостехиздат, 1950.

5. Проблема минимума квадратичного функционала, М.—Л., Гостехиздат, 1952.

6. Некоторые вопросы теории операторов и их применение в теории упругих оболочек, ДАН **84**, № 5 (1952).

7. Оценка погрешности расчета упругой оболочки как плоской пластины, Прикл. матем. и мех. **16**, вып. 4 (1952).

8. Интегрирование уравнения Пуассона в бесконечной области, ДАН **91**, № 5 (1953).

9. Вырождающиеся эллиптические уравнения, Вест. ЛГУ, № 8 (1954).

10. Упругие оболочки, близкие к плоским пластинам, в сб. «Вопросы прочности лопасти водяной турбины», Изд. ЛГУ (1954), 5—92.

11. По поводу метода Ритца, ДАН **106**, № 3 (1956), 391—394.

12. Вариационные методы в математической физике, М., Гостехиздат, 1957.

13. Замечания о координатных функциях, Изв. ВУЗ, Матем. **5** (6) (1958), 91—94.

14. О решениях с конечной энергией у эллиптических дифференциальных уравнений, Уч. зап. Ленингр. пед. ин-та **183** (1958).

15. Об устойчивости метода Ритца, ДАН **135**, № 1 (1960).

16. О сходимости одного прямого метода, УМН **15**, вып. 1 (91) (1960).

17. Некоторые условия устойчивости метода Ритца, Вестн. ЛГУ, № 13 (1961).

18. О рациональном выборе координатных функций в методе Ритца, Журн. вычисл. матем. и матем. физики, **2**, № 3 (1962).

19. О методе Ритца в нелинейных задачах, ДАН **142**, № 4 (1962).

20. Многомерные сингулярные интегралы и интегральные уравнения, М., Физматгиз, 1962.

21. Mihlin S. G., Variational methods of solving linear and nonlinear boundary value problems, в кн. «Differential equations and their applications», Proceedings of the conference held in Prague in september 1962, Publishing House of the Czechoslovak Academy of Sciences, Prague, 1963.

22. Об устойчивости некоторых вычислительных процессов, ДАН **157**, № 2 (1964).

Михлин С. Г., Смолицкий Х. М. 1. Приближенные методы решения дифференциальных и интегральных уравнений (серия «Справочная математическая библиотека») М., «Наука», 1965.

Морозов Н. Ф., 1. К нелинейной теории тонких пластин, ДАН **114**, № 5 (1957), 968—971.

2. Единственность симметричного решения задачи о больших прогибах симметрично загруженной круглой пластины, ДАН **123**, № 3 (1958).

3. Нелинейные задачи теории тонких пластин, Вестн. ЛГУ, № 19, сер. матем., мех. и астр., вып. 4 (1958).

4. Нелинейные задачи теории тонких анизотропных пластин, Изв. ВУЗ, Матем. **6** (19) (1960).

5. К вопросу о существовании несимметричного решения в задаче о больших прогибах круглой пластины, загруженной симметричной нагрузкой, Изв. ВУЗ, Матем. **2** (1961).

Наймарк М. А. 1. Линейные дифференциальные операторы, М., Гостехиздат, 1954.

Натансон И. П. 1. К теории приближенного решения уравнений, Уч. зап. Ленингр. пед. ин-та **64** (1948), 3—8.

2. Теория функций вещественной переменной, М., Гостехиздат, 1957.

Новожилов В. В. 1. Теория тонких оболочек, М., Судпромгиз, 1951.

2. Теория упругости, М., Судпромгиз, 1958.

Первозванская Т. Н., 1. Проведение аналитических выкладок на ЭВМ при решении некоторых типов дифференциальных уравнений, Труды Матем. ин-та им. В. А. Стеклова **66** (1962), 37—44.

Петрова Л. Т. 1. Некоторые применения схемной символики, Журн. вычисл. матем. и матем. физики **1**, № 3 (1961).

Петрова Л. Т., Платунова И. А. 1. Реализация на машине вычислений в исходном классе списков, Труды Матем. ин-та им. В. А. Стеклова **66**, (1962), 16—36.

Петрова Л. Т., Яковлева М. А. 1. О крупноблочном программировании, Труды III Всесоюзн. матем. съезда, II, Изд. АН СССР, 1956.

Польский Н. И. 1. О сходимости некоторых приближенных методов анализа, Укр. матем. журн. **7**, № 1 (1955), 56—70.

Постнов В. А. Большие прогибы для одного частного случая несимметричного относительно центра закрепления кромок опорного контура, Труды Ленингр. кораблестроит. ин-та **16** (1955), 21—33.

Райснер (Reissner E. L.) 1. On a variational theorem for finite elastic deformations, Journ. Math. and Phys. **32**, № 2—3 (1953), 129—135.

Рис Ф., Секефальви-Надь Б. 1. Лекции по функциональному анализу, М., ИЛ, 1954.

Розе С. Н. 1. О сходимости метода Л. М. Качанова, Вестн. ЛГУ, № 19 (1961), 170—174.

Самокиш Б. А. 1. К устойчивости абстрактного метода Галёркина, Вестн. ЛГУ, сер. матем., мех. и астр., № 1 (1964).

Сансоне Дж. 1. Обыкновенные дифференциальные уравнения, т. I, М, ИЛ, 1953.

Слободянский М. Г. 1. Оценка погрешности приближенных решений линейных задач, Прикл. матем. и мех. **17**, вып. 2 (1953), 229—244.

Смирнов В. И. 1. Курс высшей математики, т. II, М.—Л., Гостехиздат, 1952.

2. Курс высшей математики, т. III, ч. I. М.—Л., Гостехиздат, 1949.

3. Курс высшей математики, т. V, М., Физматгиз, 1959.

Смирнова Т. Н. 1. Полиномиальный прораб на электронную вычислительную машину «Стрела», Журн. вычисл. матем. и матем. физики **1**, № 5 (1961).

2. Полиномиальный прораб и проведение аналитических выкладок на ЭВМ, Труды Матем ин-та им. В. А. Стеклова **66** (1962), 77—112.

Соболев С. Л. 1. Некоторые применения функционального анализа в математической физике, Изд. ЛГУ, 1950.

2. Об одной новой задаче математической физики, Изв. АН, сер. матем. **18**, № 1 (1954), 3—50.

Соболевский П. Е. 1. Об уравнениях с операторами, образующими острый угол, ДАН **116**, № 5 (1957).

2. Об одном неравенстве для эллиптических операторов, Труды семинара по функц. анализу, Воронежск. ун-т, вып. 6 (1958).

Спэнг Х. А. (Spang H. A.) 1. Areview of minimisation techniques for nonlinear functions, SIAM Review **4** (1962), 343—365.

Стиппс (Stippes M.) 1. Large deflections of rectangular plates, Proc. first U. S. Nat. Congr. Appl. Mech., Publ. Amer. Soc. Mech. Eng., N. Y., 1952, 339—345.

Талдыкин А. Т. 1. Системы элементов гильбертова пространства и ряды по ним, Матем. сб. **29** (**71**): 1 (1951), 79—120.

Уилкинсон Дж. Х. (Wilkinson J. H.) 1. Rounding errors in algebraic processes, National Physical Laboratory, Notes on applied science, № **32**, London (1963).

Фаддеев Д. К., Фаддеева В. Н. 1. Вычислительные методы линейной алгебры, М., Физматгиз, 1963.

Фридрихс (Friedrichs K.) 1. Spektraltheorie Halbbeschränkter Operatoren und Anwendung auf der Spektralzerlegung von Differentialoperatoren, Math. Ann. **109**, H. 4—5 (1934), 465—487.

Хайнц (Heinz E.) 1. Beiträge sur Störungstheorie der Spektralzerlegung, Math. Ann. **123**, H. 4 (1951), 415—438.

Харди Г. Г., Литтльвуд Дж. Е., Полиа Г. 1. Неравенства, М., ИЛ, 1948.

Хаусдорф Ф. 1. Теория множеств, М.—Л., ОНТИ, 1937.

Хилл Р. 1. Математическая теория пластичности, М., Гостехиздат, 1956.

Читтенден (Chittenden E. W.) 1. On the equivalence of écart and voisinage, Trans. Amer. Math. Soc. **18**, № 2 (1917), 161—166.

Шахбазян К. В. 1. Исчисление программы функциональных операций, Труды Матем. ин-та им. В. А. Стеклова **66** (1962), 45—76.

Шихобалов С. П., Краснов В. М., Максутова Т. Д., Цейтц В. В., Эдельштейн Е. И. 1. Экспериментальное исследование напряженного состояния лопасти водяной турбины, в сб. «Вопросы прочности лопасти водяной турбины», Изд. ЛГУ (1954), 174—215.

Эйдус Д. М., 1. О смешанной задаче теории упругости, ДАН **76**, № 2 (1951).

2. Контактная задача теории упругости, Матем. сб. **54** (**96**); 3 (1954), 429—440.

Эрдейи А., Магнус В., Оберхеттингер Ф., Трикоми Ф. (Erdelyi A., Magnus W., Oberhettinger F., Tricomi F. G.). 1. Higher transcendental, I—III, McGraw-Hill Book Company, INC, N. Y., 1953.

Яковлев М. Н. 1. К решению систем нелинейных уравнений методом дифференцирования по параметру, Журн. вычисл. матем. и матем. физики **4**, № 1 (1964), 146—149.

Яковлева М. А. 1. Крупноблочная система программирования, Труды Матем. ин-та им. В. А. Стеклова **66** (1962), 4—16.

Яскова Г. Н., Яковлев М. Н. 1. Некоторые условия устойчивости метода Петрова — Галёркина, Труды Матем. ин-та им. В. А. Стеклова **66** (1962), 182—189.

ИМЕННОЙ УКАЗАТЕЛЬ

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

1р. 80к.