

*С. К. Годунов, В. С. Рябенский*

# РАЗНОСТНЫЕ СХЕМЫ

## ВВЕДЕНИЕ В ТЕОРИЮ

*Допущено Министерством  
высшего и среднего специального образования СССР  
в качестве учебного пособия  
для студентов университетов и высших учебных заведений  
по специальности «Прикладная математика»*

ИЗДАНИЕ ВТОРОЕ, ПЕРЕРАБОТАННОЕ  
И ДОПОЛНЕННОЕ



ИЗДАТЕЛЬСТВО «НАУКА»  
ГЛАВНАЯ РЕДАКЦИЯ  
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ  
Москва 1977

**Разностные схемы (введение в теорию)**, С. К. Годунов, В. С. Рябенский, учебное пособие, Главная редакция физико-математической литературы изд-ва «Наука», М., 1977.

Теория разностных схем численного решения дифференциальных уравнений является одной из основных частей современной вычислительной математики.

Книга предназначена для первоначального ознакомления с теорией разностных схем и является учебным пособием для студентов университетов, высших учебных заведений с расширенной программой по математике, а также может быть использована как учебное руководство для студентов других вузов, в которых преподаются численные методы решения дифференциальных уравнений. Некоторые разделы книги представляют интерес и для специалистов в области методов вычислений.

Илл. 57, библи. 28 названий.

# ОГЛАВЛЕНИЕ

Предисловие . . . . .	9
Предисловие ко второму изданию . . . . .	10
Введение . . . . .	11

## ЧАСТЬ ПЕРВАЯ

### ОБЫКНОВЕННЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

<b>Глава 1. Разностные уравнения первого и второго порядка. Примеры разностных схем . . . . .</b>	<b>15</b>
§ 1. Простейшие разностные уравнения . . . . .	15
1. Разностные уравнения (15). 2. Порядок разностного уравнения (18). 3. Общее решение разностного уравнения (18). Задачи . . . . .	20
§ 2. Разностное уравнение первого порядка . . . . .	21
1. Фундаментальное решение (21). 2. Условие ограниченности фундаментального решения (22). 3. Частное решение (23). Задачи . . . . .	25
§ 3. Разностное уравнение второго порядка . . . . .	25
1. Общее решение однородного уравнения (6). 2. Общее решение неоднородного уравнения. Фундаментальное решение (29). 3. Оценка фундаментального решения через коэффициенты разностного уравнения (34). Задачи . . . . .	36
<b>Глава 2. Краевая задача для уравнения второго порядка . . . . .</b>	<b>38</b>
§ 4. Постановка задачи. Признаки хорошей обусловленности . . . . .	38
1. Постановка задачи (38). 2. Определение хорошей обусловленности (39). 3. Достаточный признак хорошей обусловленности (41). 4. Критерий хорошей обусловленности краевой задачи с постоянными коэффициентами (42). 5. Критерий хорошей обусловленности задачи с переменными коэффициентами (42). 6. Обоснование критерия хорошей обусловленности краевой задачи с постоянными коэффициентами (44). 7. Общие краевые задачи для систем разностных уравнений (48). Задачи . . . . .	50
§ 5. Алгоритм решения краевой задачи — прогонка . . . . .	51
1. Описание прогонки (51). 2. Пример вычислительно неустойчивого алгоритма (53). Задачи . . . . .	55

<b>Глава 3. Обоснование метода прогонки . . . . .</b>	<b>56</b>
§ 6. Свойства хорошо обусловленных краевых задач . . . . .	56
1. Оценки решений краевой задачи с возмущенными коэффициентами (56).	
2. Доказательство критерия хорошей обусловленности (61). 3. Свойства хорошо обусловленных задач (64).	
§ 7. Обоснование метода прогонки для хорошо обусловленных краевых задач . . . . .	65
1. Оценки прогоночных коэффициентов (65). 2. Оценка влияния на результат ошибок округления в процессе вычислений (67).	

## ЧАСТЬ ВТОРАЯ

### РАЗНОСТНЫЕ СХЕМЫ ДЛЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

<b>Глава 4. Элементарные примеры разностных схем . . . . .</b>	<b>71</b>
§ 8. Понятие о порядке точности и об аппроксимации . . . . .	71
1. Порядок точности разностной схемы (71). 2. Скорость сходимости решения разностного уравнения (75). 3. Порядок аппроксимации (77).	
§ 9. Неустойчивая разностная схема . . . . .	78
1. Способы аппроксимации производной (78). 2. Пример неустойчивой разностной схемы (78).	
<b>Глава 5. Сходимость решений разностных уравнений как следствии аппроксимации и устойчивости . . . . .</b>	<b>82</b>
§ 10. Сходимость разностной схемы . . . . .	82
1. Понятие о сетке и сеточной функции (82). 2. Сходящиеся разностные схемы (87). 3. Проверка сходимости разностной схемы (90).	
Задачи . . . . .	92
§ 11. Аппроксимация дифференциальной краевой задачи разностной схемой . . . . .	92
1. Невязка $\delta f(h)$ (92). 2. Вычисление невязки (94). 3. Аппроксимация порядка $h^k$ (96). 4. Примеры (97). 5. Разбиение разностной схемы на подсистемы (100). 6. Замена производных разностными отношениями (103). 7. Другие способы построения разностных схем (105).	
Задачи . . . . .	106
§ 12. Определение устойчивости разностной схемы. Сходимость как следствие аппроксимации и устойчивости . . . . .	106
1. Определение устойчивости (106). 2. Зависимость между аппроксимацией, устойчивостью и сходимостью (108). 3. Сходящаяся разностная схема для интегрального уравнения (114).	
§ 13. О выборе норм . . . . .	115
§ 14. Достаточный признак устойчивости разностных схем решения задачи Коши . . . . .	123
1. Вводный пример (124). 2. Каноническая запись разностной схемы (125).	
3. Устойчивость как ограниченность норм степеней оператора перехода (127). 4. Примеры исследования устойчивости (128). 5. Неединственность канонической записи (133).	
Задачи . . . . .	135

§ 15. Необходимый спектральный признак устойчивости . . . . .	136
1. Ограниченность норм степеней оператора перехода необходима для устойчивости (136). 2. Спектральный признак устойчивости (138). 3. Обсуждение спектрального признака устойчивости (139). Задачи . . . . .	143
§ 16. Ошибки округления . . . . .	144
1. Ошибки в коэффициентах (144). 2. Ошибки в вычислениях (147).	
§ 17. Количественная характеристика устойчивости . . . . .	149
§ 18. Прием исследования устойчивости нелинейных задач . . . . .	155
<b>Глава 6. Употребительные разностные схемы . . . . .</b>	<b>157</b>
§ 19. Схемы Рунге — Кутты и Адамса . . . . .	157
1. Схемы Рунге — Кутты (158). 2. Схемы Адамса (160). 3. Замечания об устойчивости (163). 4. Обобщение на системы уравнений (164).	
§ 20. Методы решения краевых задач . . . . .	166
1. Метод стрельбы (166). 2. Метод прогонки (169). 3. Метод Ньютона (169).	

### *ЧАСТЬ ТРЕТЬЯ*

## **РАЗНОСТНЫЕ СХЕМЫ ДЛЯ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ. ОСНОВНЫЕ ПОНЯТИЯ**

<b>Глава 7. Простейшие приемы построения и исследования разностных схем . . . . .</b>	<b>171</b>
§ 21. Напоминание и иллюстрация основных определений . . . . .	171
1. Определение сходимости (171). 2. Определение аппроксимации (172). 3. Определение устойчивости (175). Задачи . . . . .	182
§ 22. Простейшие приемы построения аппроксимирующих разностных схем . . . . .	183
1. Замена производных разностными отношениями (183). 2. Метод неопределенных коэффициентов (191). 3. Схемы с пересчетом, или схемы предиктор-корректор (200). 4. О других приемах (202). Задачи . . . . .	202
§ 23. Примеры конструирования граничных условий при построении разностных схем . . . . .	204
Задачи . . . . .	209
§ 24. Условие Куранта, Фридрикса и Леви, необходимое для сходимости . . . . .	210
1. Условие Куранта, Фридрикса и Леви (210). 2. Примеры разностных схем для задачи Коши (211). 3. Примеры разностных схем для задачи Дирихле (217). Задачи . . . . .	219
<b>Глава 8. Некоторые основные приемы исследования устойчивости</b>	<b>221</b>
§ 25. Спектральный анализ разностной задачи Коши . . . . .	221
1. Устойчивость по начальным данным (221). 2. Необходимое спектральное условие устойчивости (222). 3. Примеры (224). 4. Интегральное представление решения (231). 5. Выглаживание разностного решения как действие аппроксимационной вязкости (237). Задачи . . . . .	239

§ 26. Принцип замороженных коэффициентов . . . . .	240
1. Замораживание коэффициентов во внутренних точках (240). 2. Признак Бабенко и Гельфанда (243). Задачи . . . . .	249
§ 27. Представление решений некоторых модельных задач в виде конечных рядов Фурье . . . . .	251
1. Ряды Фурье для сеточных функций (251). 2. Представление решений разностных схем для уравнения теплопроводности на отрезке (254). 3. Представление решений разностных схем для двумерной задачи теплопроводности (257). 4. Представление решения разностной схемы для задач о колебаниях струны (260). Задачи . . . . .	262
§ 28. Принцип максимума . . . . .	262
1. Явная разностная схема (263). 2. Неявная разностная схема (265). 3. Сопоставление явной и неявной разностных схем (*66).	
<b>Глава 9. Понятие о разностных схемах для расчета обобщенных решений . . . . .</b>	<b>268</b>
§ 29. Обобщенное решение . . . . .	268
1. Механизм возникновения разрывов (269). 2. Определение обобщенного решения (270). 3. Условие на линии разрыва решения (272). 4. Распад произвольного разрыва (273). 5. Другое определение обобщенного решения (275).	
§ 30. Построение разностных схем . . . . .	275
1. Схема с искусственной вязкостью (276). 2. Метод характеристик (277). 3. Дивергентные разностные схемы (278).	

## *ЧАСТЬ ЧЕТВЕРТАЯ*

### **ЗАДАЧИ С ДВУМЯ ПРОСТРАНСТВЕННЫМИ ПЕРЕМЕННЫМИ**

<b>Глава 10. Понятие о разностных схемах расщепления . . . . .</b>	<b>284</b>
§ 31. Конструкция схем расщепления . . . . .	284
Задачи . . . . .	288
§ 32. Экономичные разностные схемы . . . . .	288
Задачи . . . . .	296
§ 33. Расщепление по физическим факторам . . . . .	296
<b>Глава 11. Эллиптические задачи . . . . .</b>	<b>298</b>
§ 34. Простейшая разностная схема для задачи Дирихле . . . . .	298
1. Аппроксимация (298). 2. Устойчивость (299). Задачи . . . . .	303
§ 35. Метод установления . . . . .	304
1. Идея метода установления (304). 2. Анализ явной схемы установления (307). 3. Схема переменных направлений (309). 4. Выбор точности (311). 5. Границы применимости методов (311). Задачи . . . . .	312

§ 36. Итерации с переменным шагом . . . . .	312
1. Идея Ричардсона (312). 2. Чебышевский набор параметров (313). 3. Нумерация итерационных параметров (316). 4. Метод Дугласа—Рэф- форда (320). Задачи . . . . .	322
§ 37. Метод Федоренко . . . . .	323
1. Идея метода (324). 2. Описание алгоритма (325).	
<b>Глава 12. Понятие о вариационно-разностных и проекционно-раз- ностных схемах . . . . .</b>	<b>327</b>
§ 38. Вариационные и проекционные методы . . . . .	327
1. Вариационная постановка краевых задач (327). 2. Сходимость миними- зирующих последовательностей (331). 3. Вариационный метод Ритца (335). 4. Проекционный метод Галеркина (341). 5. Способы решения алгебраиче- ской системы (343). 6. Вычислительная устойчивость (343). Задачи . . . . .	343
§ 39. Построение и свойства вариационно-разностных и проекционно- разностных схем . . . . .	344
1. Определение вариационно-разностных и проекционно-разностных схем (344). 2. Пример вариационно-разностной схемы для первой крае- вой задачи (346). 3. Пример вариационно-разностной схемы для третьей краевой задачи (354). 4. О методике доказательства сходимости (358). 5. Сопоставление вариационно-разностных схем с общими вариацион- ными и обычными разностными (358). Задача . . . . .	359

## ЧАСТЬ ПЯТАЯ

### УСТОЙЧИВОСТЬ ЭВОЛЮЦИОННЫХ РАЗНОСТНЫХ КРАЕВЫХ ЗАДАЧ КАК ОГРАНИЧЕННОСТЬ НОРМ СТЕПЕНЕЙ НЕКОТОРОГО ОПЕРАТОРА

<b>Глава 13. Конструкция оператора перехода . . . . .</b>	<b>362</b>
§ 40. Слоистая структура решений эволюционных задач . . . . .	362
Задачи . . . . .	365
§ 41. Запись разностных краевых задач в виде $u^{p+1} = R_h u^p + \tau r^p$ . . . . .	365
1. Канонический вид (365). 2. Устойчивость как равномерная ограниче- нность норм степеней $R_h$ (368). 3. Пример (372). Задачи . . . . .	374
§ 42. Использование частных решений при конструировании опера- тора перехода . . . . .	375
§ 43. Некоторые способы оценки норм степеней операторов . . . . .	387
1. Необходимые спектральные условия ограниченности $\ R_h^p\ $ (387). 2. Спектральный критерий ограниченности степеней самосопряженного оператора (389). 3. Признаки самосопряженности (393). 4. Оценки собст- венных значений оператора $R_h$ (391). 5. Выбор скалярного умноже- ния (393). 6. Критерии устойчивости Самарского (394). Задачи . . . . .	395

<b>Глава 14. Спектральный признак устойчивости несамосопряженных эволюционных краевых задач . . . . .</b>	<b>396</b>
§ 44. Спектр семейства операторов $\{R_h\}$ . . . . .	396
1. Необходимость усовершенствования спектрального признака устойчивости (396). 2. Определение спектра семейства операторов (398). 3. Необходимое условие устойчивости (399). 4. Обсуждение понятия спектра семейства операторов $\{R_h\}$ (400). 5. Близость необходимого признака устойчивости к достаточному (401).	
§ 45. Алгоритм вычисления спектра семейства разностных операторов над сеточными функциями на отрезке . . . . .	403
1. Характерный пример (404). 2. Алгоритм вычисления спектра в общем случае (411). Задачи . . . . .	412
§ 46. Ядра спектров семейств операторов . . . . .	412
§ 47. Об устойчивости итерационных алгоритмов решения несамосопряженных разностных уравнений . . . . .	415
<b>Дополнение. Метод внутренних граничных условий . . . . .</b>	<b>419</b>
1. Класс систем разностных уравнений (419). 2. Фундаментальное решение (420). 3. Граница сеточной области (420). 4. Разностные аналоги интегральных формул Коши и типа Коши (421). 5. Внутренние граничные условия (423). 6. Оператор граничного проектирования (423). 7. Общая краевая задача (423). 8. Основная идея метода внутренних граничных условий (424). 9. Устойчивость внутренних граничных условий (424). 10. Дополнительная идея (425). 11. Сопоставление метода внутренних граничных условий с методом сингулярных интегральных уравнений (426).	
Библиографические комментарии . . . . .	429
Литература . . . . .	434
Предметный указатель . . . . .	436



## ПРЕДИСЛОВИЕ

Многие вопросы естествознания приводят к краевым задачам для дифференциальных уравнений. С целью решения этих задач на электронных вычислительных машинах их приближенно заменяют разностными схемами.

Книга предназначена для первоначального ознакомления с теорией разностных схем и написана как учебное пособие для студентов технических вузов, Московского физико-технического и Московского инженерно-физического институтов, для студентов физических и математических факультетов университетов. Вместе с тем, вероятно, некоторые разделы книги будут интересны и специалистам в области вычислений.

Различие интересов перечисленных категорий читателей нашло отражение в структуре книги.

Книга состоит из пяти частей и небольшого Дополнения. Любое число нескольких (двух или более) первых частей составляет некоторое законченное введение в предмет. Кроме того, объем изучаемого материала можно регулировать за счет текста, напечатанного мелким шрифтом, и за счет количества решаемых задач.

В конце указана литература для углубленного изучения многих вопросов теории и приложений разностных схем и для дальнейших библиографических справок.

Более кратким введением в теорию разностных схем может служить книга [1].

Непосредственно в тексте книги ссылки на оригинальные работы даются лишь в тех немногих случаях, когда дополнительные результаты приводятся без доказательств.

Современная вычислительная техника и накопленный опыт позволяют с помощью разностных схем приближенно вычислять решения очень сложных и плохо поддающихся исследованию другими методами задач. Уверенность в том, что решение вычислено правильно, достигается применением той же вычислительной схемы для расчета немногих задач, точные решения которых заранее известны, сопоставлением результатов расчета с физическим экспериментом в том диапазоне параметров, где

этот эксперимент возможен, и с помощью других методов, которые нельзя считать математически строгими. Но понимание существа дела, необходимое для построения пригодных разностных схем, достигается путем рассмотрения серии правильно подобранных модельных задач, достаточно простых для детального изучения на принятом в математике уровне строгости, но все же улавливающих те или иные интересующие нас черты исходной задачи, недоступной для строгого изучения либо ввиду сложности, либо ввиду недостатка времени.

Делая ударение на математически строгий разбор модельных задач, мы старались в то же время дать правильное представление о соотношении теории и эксперимента на ЭВМ при создании разностных схем для практических расчетов.

Появлению этой книги способствовала предшествующая работа авторов над книгой [10], а также работа одного из них над лекционными курсами, которые он читал в течение ряда последних лет в Московском физико-техническом институте. На становление этих курсов большое влияние оказали многочисленные плодотворные дискуссии с О. М. Белоцерковским (по инициативе которого эти курсы начали читаться), В. Ф. Дьяченко, О. В. Локуциевским, Р. П. Федоренко, Л. А. Чудовым и Э. Э. Шнолем. Ряд полезных замечаний сделали Н. С. Бахвалов и Б. Л. Рождественский, прочитавшие книгу в рукописи.

Всем им мы сердечно благодарны.

*Авторы*

## **ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ**

Второе издание книги отличается от первого тем, что в него включены глава 12 о вариационно-разностных схемах, § 47 об устойчивости итерационных процессов решения несамосопряженных разностных уравнений и п. 10 Дополнения, содержащий соображения об использовании метода внутренних граничных условий для вычислений. Кроме того, устранены замеченные опечатки и неточности, а также обновлен список литературы.

*Авторы*

## ВВЕДЕНИЕ

Многие прикладные и теоретические задачи современного естествознания приводят к дифференциальным уравнениям. Исследование задачи может считаться законченным только после того, как эти уравнения решены.

В некоторых случаях удается указать формулу, выражающую решение через хорошо изученные элементарные функции. Однако, как правило, это принципиально невозможно. Поэтому получение явных формул не может считаться регулярным процессом, ведущим к решению дифференциальных уравнений. Нельзя сказать, что этот аналитический подход полностью утратил свое значение. Он остается необходимым и очень мощным инструментом изучения упрощенных, так называемых модельных задач. Изучение хорошо подобранной модельной задачи позволяет делать некоторые заключения о характере поведения решения неупрощенной, исходной задачи.

Но наряду с этим аналитическим подходом все шире используются различные методы численного решения дифференциальных уравнений. Их широкое использование стало возможно с появлением быстродействующих вычислительных машин, которые могут запоминать большие таблицы чисел и производить над ними арифметические действия по заданной программе. В соответствии с указанными возможностями машин любой численный метод состоит в переходе от искомого решения к некоторой искомой таблице чисел и к указанию последовательности арифметических действий для их вычисления. Можно, например, искать несколько первых коэффициентов разложения решения в степенной или тригонометрический ряд. Здесь излагается теория численного решения дифференциальных уравнений с помощью метода конечных разностей. Сущность этого наиболее универсального численного метода состоит в том, что за искомый набор чисел принимается таблица значений решения в точках некоторого множества, называемого обычно *сеткой*. Для вычисления искомой таблицы используются алгебраические уравнения, приближенно заменяющие дифференциальное.

Поясним это на простейшем примере разностной схемы для приближенного вычисления решения уравнения

$$u'(x) + Au(x) = 0,$$

удовлетворяющего начальному условию  $u(0) = 1$ . Зададим  $h > 0$  и вместо функции  $u(x)$  будем искать таблицу ее значений

$$u(0), u(h), u(2h), \dots, u(nh), \dots$$

Заменим производную разностным отношением

$$\frac{u(x+h) - u(x)}{h},$$

ее приближающим. Шаг  $h$  таблицы должен быть выбран достаточно малым. После такой замены вместо дифференциального уравнения мы получаем приближающее его разностное уравнение

$$\frac{u(x+h) - u(x)}{h} + Au(x) = 0,$$

которое позволяет приближенно вычислить искомую таблицу. Для этого перепишем разностное уравнение в виде рекуррентной формулы

$$u(x+h) = (1 - Ah)u(x).$$

Полагая последовательно  $x = 0, h, 2h, \dots$ , получим

$$\begin{aligned} u(h) &= (1 - Ah), \\ u(2h) &= (1 - Ah)^2, \\ &\dots \\ u(Nh) &= (1 - Ah)^N. \\ &\dots \end{aligned}$$

Выбрав  $h = 1/N$ , получим

$$u(1) = \left(1 - \frac{A}{N}\right)^N$$

вместо точного решения

$$u(1) = e^{-A}.$$

Однако, как это хорошо известно из курса математического анализа, при достаточно малом  $h$  или, что то же самое, при достаточно большом  $N$  величина  $(1 - A/N)^N$  мало отличается от  $e^{-A}$ . Тем самым показано, что приближенное решение, полученное по этой разностной схеме и зависящее от шага  $h$ , при измельчении шага сходится к точному решению дифференциального уравнения.

Другой пример разностного уравнения, приближающего то же дифференциальное уравнение

$$u'(x) + Au(x) = 0,$$

мы получим, заменяя производную разностным отношением

$$\frac{u(x+h) - u(x-h)}{2h}.$$

Это уравнение имеет вид

$$\frac{u(x+h) - u(x-h)}{2h} + Au(x) = 0.$$

Для дифференциального уравнения

$$u''(x) + Au'(x) + Bu(x) = f(x)$$

можно построить разностный аналог, заменяя  $u''(x)$ , например, следующим приближенным выражением:

$$\frac{\frac{u(x+h) - u(x)}{h} - \frac{u(x) - u(x-h)}{h}}{h} = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}.$$

Первую производную  $u'(x)$  можно заменить одним из уже употреблявшихся разностных отношений. После таких замен получим разностное уравнение

$$\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + A \frac{u(x+h) - u(x-h)}{2h} + Bu(x) = f(x).$$

В случае дифференциальных уравнений с переменными коэффициентами составление разностных схем не усложняется. Если, например, требуется вычислить решение уравнения

$$u'(x) + A(x)u(x) = 0,$$

у которого коэффициент  $A$  является функцией от  $x$ , то это можно сделать с помощью разностного уравнения

$$\frac{u(x+h) - u(x)}{h} + A(x)u(x) = 0.$$

Так же легко «справляются» разностные схемы и с нелинейными уравнениями. Например, уравнение

$$u'(x) + \sin(xu(x)) = 0$$

может быть решено приближенно по схеме

$$\frac{u(x+h) - u(x)}{h} + \sin(xu(x)) = 0.$$

Из рассмотрения примеров может сложиться впечатление, что составление разностной схемы и решение по ней дифферен-

циального уравнения не представляет трудностей. Это впечатление обманчиво.

Уже в самых простых случаях, даже при решении линейных уравнений с постоянными коэффициентами, часто бывает, что, казалось бы, разумная разностная схема имеет решение, не сходящееся при измельчении сетки к искомому решению дифференциального уравнения. Понятно, что по такой схеме нельзя вычислить искомую функцию со сколь угодно высокой точностью.

Далее, после того как сходящаяся разностная схема построена, необходимо вычислить решение возникающей системы алгебраических уравнений относительно большого числа неизвестных значений функции в узлах сетки. Это во многих важных случаях непросто. Иногда можно обойти указанное препятствие, выбрав сходящуюся разностную схему другой конструкции так, чтобы возникающую систему линейных уравнений легко было решить точно; в некоторых других случаях разработаны приемы приближенного вычисления решений разностных задач с любой наперед заданной точностью.

Каждый, кто занимается численным решением дифференциальных уравнений, должен знать трудности, связанные с построением и использованием разностных схем, и способы их преодоления.

ОБЫКНОВЕННЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

ГЛАВА I

РАЗНОСТНЫЕ УРАВНЕНИЯ ПЕРВОГО  
И ВТОРОГО ПОРЯДКА.  
ПРИМЕРЫ РАЗНОСТНЫХ СХЕМ

§ 1. Простейшие разностные уравнения

1. Разностные уравнения. Для дифференциального уравнения первого порядка

$$u'(x) + Au(x) = f(x)$$

мы построили во введении две разностные схемы:

$$\frac{u(x+h) - u(x)}{h} + Au(x) = f(x),$$

$$\frac{u(x+h) - u(x-h)}{2h} + Au(x) = f(x),$$

которые можно записать соответственно в виде

$$-\frac{1-Ah}{h}u(x) + \frac{1}{h}u(x+h) = f(x), \quad (1)$$

$$-\frac{1}{2h}u(x-h) + Au(x) + \frac{1}{2h}u(x+h) = f(x). \quad (2)$$

Для дифференциального уравнения второго порядка

$$u''(x) + Au'(x) + Bu(x) = f(x)$$

во введении было построено разностное уравнение

$$\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + A \frac{u(x+h) - u(x-h)}{2h} + Bu(x) = f(x),$$

или, в другой записи,

$$\frac{1}{h^2} \left( 1 - \frac{Ah}{2} \right) u(x-h) - \frac{1}{h^2} (2 - Bh^2) u(x) +$$

$$+ \frac{1}{h^2} \left( 1 + \frac{Ah}{2} \right) u(x+h) = f(x). \quad (3)$$

Приведенные здесь примеры разностных уравнений, приближающих простейшие дифференциальные, принадлежат к одному из следующих двух видов:

$$au(x) + bu(x+h) = f(x), \quad (1')$$

$$au(x-h) + bu(x) + cu(x+h) = f(x). \quad (2')$$

Если последовательность точек, делящих ось  $Ox$  на отрезки длины  $h$ , занумеровать слева направо так, чтобы  $x_n = x_{n-1} + h$ , и обозначить  $u(x_n)$  через  $u_n$ , а  $f(x_n)$  через  $f_n$ , то мы можем переписать наши разностные схемы в виде уравнений

$$au_n + bu_{n+1} = f_n, \quad (4)$$

$$au_{n-1} + bu_n + cu_{n+1} = f_n. \quad (5)$$

В §§ 1—4 мы займемся изучением разностных уравнений вида (4) и (5), причем не будем интересоваться, являются ли эти уравнения разностными схемами для каких-либо дифференциальных уравнений.

В уравнениях (4) и (5) неизвестные  $u_n$  образуют последовательность  $\{u_n\}$ :

$$\dots, u_{-3}, u_{-2}, u_{-1}, u_0, u_1, u_2, u_3, \dots$$

Мы будем часто сопоставлять эту последовательность с последовательностью точек, занумерованных числами

$$\dots, -3, -2, -1, 0, 1, 2, 3, \dots,$$

или, как иногда говорят, с сеткой.

Последовательность  $\{u_n\}$  можно считать функцией  $u$ , заданной в точках сетки. Тогда  $u_k$  есть значение сеточной функции  $u$  в точке, имеющей номер  $k$ .

На рис. 1 приведен график некоторой сеточной функции  $u$ . Этот график есть совокупность точек  $(x_k, u_k)$  на плоскости  $Ox_i$ .

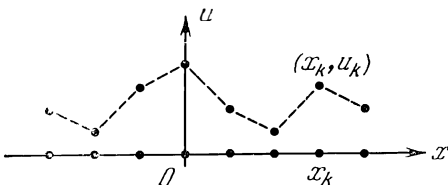


Рис. 1.

После того как мы отказались от рассмотрения связи разностных уравнений с дифференциальными, нам вовсе не обязательно считать, что расстояние между двумя соседними точками равно  $h$ . Можно выбрать его произвольным, например равным единице, а в качестве  $x_0$  взять точку нуль. Тогда сеточная функция  $u$  будет определена в точках с целыми координатами  $x_k = k$ .

Будем считать для простоты, что коэффициенты  $a, b, c$  уравнений (4), (5) постоянны. Говоря, что изучаемые уравне-



ния являются уравнениями с постоянными коэффициентами, мы имеем в виду независимость этих коэффициентов от номера  $n$ ; например, уравнение

$$u_{n-1} + 5\sqrt{n}u_n + u_{n+1} = 0$$

не является уравнением с постоянными коэффициентами.

Мы будем рассматривать только такие уравнения (4), у которых  $a$  и  $b$  отличны от нуля. В уравнении (5) отличными от нуля будем считать коэффициенты  $a$  и  $c$ .

Последовательность  $\{f_n\}$  называется *правой частью* рассматриваемых уравнений.

Если предполагать, что последовательность  $\{u_n\}$  определена во всех целых точках  $n$ ,  $-\infty < n < \infty$ , и не накладывать на эту последовательность никаких дальнейших ограничений, то легко видеть, что уравнения (4) и (5) имеют много решений. Например, уравнение  $qu_n - u_{n+1} = 0$  допускает как решение  $u_n \equiv 0$ , так и решение  $u_n = q^n$ .

Чтобы выделить единственное решение уравнения (4)

$$au_n + bu_{n+1} = f_n,$$

достаточно задать значение этого решения в какой-нибудь одной целой точке  $m$ , т. е. задать  $u_m$ . В самом деле, уравнение (4) можно записать в виде рекуррентной формулы

$$u_{n+1} = \frac{1}{b}(f_n - au_n),$$

из которой при  $n = m, m+1, \dots$  последовательно определяются  $u_{m+1}, u_{m+2}, \dots$ , т. е. все  $u_n$  при  $n > m$ . Записывая уравнение в виде другой рекуррентной формулы:

$$u_{n-1} = \frac{1}{a}(f_n - bu_n),$$

мы таким же путем определим все  $u_n$  при  $n < m$ .

Для выделения единственного решения уравнения (5)

$$au_{n-1} + bu_n + cu_{n+1} = f_n$$

достаточно задать произвольно значения  $u$  в каких-нибудь двух последовательных целых точках, например задать значения  $u_{m-1}$  и  $u_m$ . Доказательство немедленно следует из того, что рассматриваемое уравнение может быть переписано в виде следующих двух рекуррентных формул:

$$u_{n+1} = \frac{1}{c}(f_n - bu_n - au_{n-1}),$$

$$u_{n-1} = \frac{1}{a}(f_n - bu_n - cu_{n+1}).$$

**2. Порядок разностного уравнения.** Повторим еще раз полученный результат и сформулируем понятие порядка для разностных уравнений (4) и (5).

Для выделения единственного решения уравнения (4)

$$au_n + bu_{n+1} = f_n$$

достаточно задать значение  $u$  в одной точке. Такое уравнение называется *уравнением первого порядка*. Для выделения единственного решения уравнения (5)

$$au_{n-1} + bu_n + cu_{n+1} = f_n$$

достаточно задать значения решения в двух последовательных точках. В связи с этим такое уравнение называется *уравнением второго порядка*.

Можно было бы еще рассмотреть простейшее уравнение

$$au_n = f_n, \quad a \neq 0,$$

решение которого определяется единственным образом без наложения каких-либо предварительных ограничений на последовательность  $\{u_n\}$ . Такое уравнение естественно назвать *уравнением нулевого порядка*.

Простейшая разностная схема (1) для дифференциального уравнения первого порядка  $u' + Au = f$  является разностным уравнением первого порядка.

Схема (3) для дифференциального уравнения второго порядка  $u'' + Au' + Bu = f$  имеет второй порядок.

Пример схемы (2)

$$-\frac{1}{2h} u(x-h) + Au(x) + \frac{1}{2h} u(x+h) = f(x)$$

для уравнения  $u' + Au = f$  показывает, что порядок разностного уравнения может быть больше порядка дифференциального уравнения. В этом примере дифференциальное уравнение имеет первый порядок, а соответствующее ему разностное уравнение — второй.

**3. Общее решение разностного уравнения.** Опишем теперь структуру решений изучаемых разностных уравнений. Сначала рассмотрим однородное уравнение

$$a\bar{u}_n + b\bar{u}_{n+1} = 0. \quad (6)$$

Обозначим через  $Y_n$  решение уравнения (6), удовлетворяющее начальному условию  $Y_0 = 1$ . Очевидно, что  $\bar{u}_n = \alpha Y_n$  также будет решением однородного уравнения при произвольном выборе постоянной  $\alpha$ . Нетрудно показать, что любое решение однородного уравнения (6) может быть представлено в таком виде. В самом деле, каждое решение однозначно определяется своим

значением при  $n = 0$ . Но решение  $\bar{u}_n$ , принимающее заданное значение  $\bar{u}_0$ , получается по формуле  $\bar{u}_n = \alpha Y_n$ , если в качестве множителя  $\alpha$  взять число  $\bar{u}_0$ .

Рассмотрим теперь неоднородное уравнение (4)

$$au_n + bu_{n+1} = f_n.$$

Пусть  $\{\tilde{u}_n\}$  и  $\{u_n^*\}$  — два каких-нибудь его решения. Вычитая друг из друга равенства

$$a\tilde{u}_n + b\tilde{u}_{n+1} = f_n,$$

$$au_n^* + bu_{n+1}^* = f_n,$$

мы видим, что разность  $\tilde{u}_n - u_n^* = \bar{u}_n$  удовлетворяет однородному уравнению (6)  $a\bar{u}_n + b\bar{u}_{n+1} = 0$ . Поэтому любое решение  $\{\tilde{u}_n\}$  можно записать в виде

$$\tilde{u}_n = u_n^* + \bar{u}_n = u_n^* + \alpha Y_n$$

при подходящем выборе постоянной  $\alpha$ . Легко проверить, с другой стороны, что при произвольном выборе  $\alpha$  формула  $u_n = u_n^* + \alpha Y_n$  задает некоторое решение неоднородного уравнения:

$$\begin{aligned} au_n + bu_{n+1} &= a(u_n^* + \alpha Y_n) + b(u_{n+1}^* + \alpha Y_{n+1}) = \\ &= (au_n^* + bu_{n+1}^*) + \alpha(aY_n + bY_{n+1}) = f_n + \alpha \cdot 0 = f_n. \end{aligned}$$

Итак, мы показали, что общее решение однородного уравнения (6)

$$a\bar{u}_n + b\bar{u}_{n+1} = 0$$

имеет вид

$$\bar{u}_n = \alpha Y_n,$$

где  $Y_n$  — частное решение этого уравнения, удовлетворяющее начальному условию  $Y_0 = 1$ , а  $\alpha$  — произвольная постоянная. Общее решение неоднородного уравнения (4)

$$au_n + bu_{n+1} = f_n$$

может быть представлено в виде

$$u_n = u_n^* + \alpha Y_n,$$

где  $u_n^*$  — какое-нибудь частное решение этого неоднородного уравнения, а  $\alpha$  — произвольная постоянная.

Аналогичное утверждение и аналогичными рассуждениями можно доказать и для разностного уравнения второго порядка. Мы не будем эти рассуждения приводить (читатель их без труда восстановит), а только сформулируем окончательный результат.

Общее решение однородного разностного уравнения

$$a\bar{u}_{n-1} + b\bar{u}_n + c\bar{u}_{n+1} = 0 \quad (7)$$

может быть представлено в виде

$$\bar{u}_n = \alpha Y_n + \beta Z_n,$$

где  $Y_n$  и  $Z_n$  — частные решения уравнения (7), удовлетворяющие начальным условиям

$$\begin{aligned} Y_0 &= 1, Y_1 = 0, \\ Z_0 &= 0, Z_1 = 1, \end{aligned}$$

а  $\alpha$  и  $\beta$  — произвольные постоянные.

Общее решение неоднородного уравнения (5)

$$au_{n-1} + bu_n + cu_{n+1} = f_n$$

может быть представлено в виде

$$u_n = u_n^* + \alpha Y_n + \beta Z_n,$$

где  $u_n^*$  — какое-нибудь частное решение этого неоднородного уравнения.

Все результаты и рассуждения этого параграфа могли бы быть дословно повторены и для разностных уравнений с переменными коэффициентами, но мы на этом не останавливаемся, чтобы не усложнять изложение несущественными подробностями.

## ЗАДАЧИ

1. Доказать, что общее решение разностного однородного уравнения

$$a_n u_n + b_n u_{n+1} = 0$$

с переменными коэффициентами  $a_n \neq 0$ ,  $b_n \neq 0$  можно записать в виде  $u_n = \alpha y_n$ , где  $y_n$  — произвольное частное решение, не при всех  $n$  обращающееся в нуль, а  $\alpha$  — произвольная постоянная.

2. Доказать, что общее решение разностного однородного уравнения второго порядка

$$a_n u_{n-1} + b_n u_n + c_n u_{n+1} = 0$$

с переменными коэффициентами,  $a_n \neq 0$ ,  $c_n \neq 0$ , можно записать в виде

$$u_n = \alpha y_n + \beta z_n,$$

где  $y_n$  и  $z_n$  — любые два частных решения этого уравнения, для которых не равен нулю определитель

$$\begin{vmatrix} y_0 & y_1 \\ z_0 & z_1 \end{vmatrix}.$$

3. Пусть  $y_n$  и  $z_n$  — два каких-нибудь частных решения разностного уравнения второго порядка из задачи 2. Доказать, что определитель

$$\begin{vmatrix} y_n & y_{n+1} \\ z_n & z_{n+1} \end{vmatrix} = y_n z_{n+1} - z_n y_{n+1}$$

либо равен нулю при каждом  $n$ , либо отличен от нуля при всех  $n$ .

4. Во скольких последовательных точках надо задать значения решения разностного уравнения

$$au_n + bu_{n+1} + cu_{n+2} + du_{n+3} = f_n,$$

$a \neq 0$ ,  $d \neq 0$ , чтобы существовало одно и только одно решение  $\{u_n\}$ , принимающее заданные значения в этих точках? Каким следует считать порядок рассматриваемого разностного уравнения?

## § 2. Разностное уравнение первого порядка

В этом параграфе будет получена формула, выражающая общее решение разностного уравнения первого порядка с постоянными коэффициентами

$$au_n + bu_{n+1} = f_n$$

при довольно слабых ограничениях на  $f_n$ .

Как показано в § 1, общее решение может быть представлено в виде

$$u_n = u_n^* + \alpha Y_n = u_n^* + \alpha \left(-\frac{a}{b}\right)^n,$$

где  $u_n^*$  — какое-нибудь частное решение, а  $\alpha$  — произвольная постоянная.

Таким образом, задача об отыскании общего решения свелась к задаче об отыскании какого-либо одного частного решения  $u_n^*$ .

**1. Фундаментальное решение.** Сначала построим решение при некоторой специальным образом заданной правой части

$$f_n = \begin{cases} 0, & n \neq 0, \\ 1, & n = 0. \end{cases}$$

Для обозначения такой функции обычно применяется символ Кронекера

$$\delta_k^n = \begin{cases} 0, & n \neq k, \\ 1, & n = k. \end{cases}$$

Тогда  $f_n = \delta_0^n$ .

Решение уравнения

$$au_n + bu_{n+1} = \delta_0^n$$

будем обозначать через  $G_n$ :

$$aG_n + bG_{n+1} = \delta_0^n. \quad (1)$$

Решение  $G_n$  называется *фундаментальным решением* уравнения

$$au_n + bu_{n+1} = f_n,$$

потому что, как мы увидим на стр. 23, через него выражаются частные решения этого уравнения при различных, довольно произвольных, правых частях  $f_n$ .

Итак, мы хотим найти какое-нибудь решение следующих трех групп уравнений:

I.  $aG_n + bG_{n+1} = 0 \quad \text{при } n \leq -1.$

II.  $aG_0 + bG_1 = 1.$

III.  $aG_n + bG_{n+1} = 0 \quad \text{при } n \geq 1.$

Пусть  $G_n = 0$  при  $n \leq 0$ . Тогда все уравнения группы I будут выполнены. Из уравнения II найдем  $G_1 = 1/b$ . Уравнения группы III можно переписать в виде рекуррентной формулы  $G_{n+1} = -\frac{a}{b} G_n$ , из которой последовательно находим

$$\begin{aligned} G_2 &= \frac{1}{b} \left(-\frac{a}{b}\right) = -\frac{1}{a} \left(-\frac{a}{b}\right)^2, \\ G_3 &= \frac{1}{b} \left(-\frac{a}{b}\right)^2 = -\frac{1}{a} \left(-\frac{a}{b}\right)^3, \\ &\dots\dots\dots \\ G_n &= -\frac{1}{a} \left(-\frac{a}{b}\right)^n \quad \text{при } n \geq 1. \end{aligned}$$

Выпишем теперь сводку формул, выражающих  $G_n$ :

$$G_n = \begin{cases} 0 & \text{при } n \leq 0, \\ -\frac{1}{a} \left(-\frac{a}{b}\right)^n & \text{при } n \geq 1. \end{cases} \quad (2)$$

Это — одно из решений уравнения (1). Прибавляя к нему общее решение  $A\left(-\frac{a}{b}\right)^n$  соответствующего однородного уравнения  $au_n + bu_{n+1} = 0$ , получим общее решение уравнения (1):

$$G_n = \begin{cases} A \left(-\frac{a}{b}\right)^n & \text{при } n \leq 0, \\ \left(A - \frac{1}{a}\right) \left(-\frac{a}{b}\right)^n & \text{при } n \geq 1. \end{cases} \quad (3)$$

Фундаментальное решение (2) получается из общей формулы (3) при  $A = 0$ .

**2. Условие ограниченности фундаментального решения.** Если  $|a/b| = 1$ , то при любом значении постоянной  $A$  получаем фун-

даментальное решение  $G_n$ , ограниченное по абсолютной величине как при  $n \rightarrow +\infty$ , так и при  $n \rightarrow -\infty$ . Выделим из общей формулы (3) ограниченное фундаментальное решение  $G_n$  в случае  $|a/b| \neq 1$ . Если  $|a/b| < 1$ , то  $|-a/b|^n$  неограниченно возрастает при  $n \rightarrow -\infty$ . Поэтому ограниченное решение получается только при  $A = 0$  (рис. 2, а). Оно задается формулой (2).

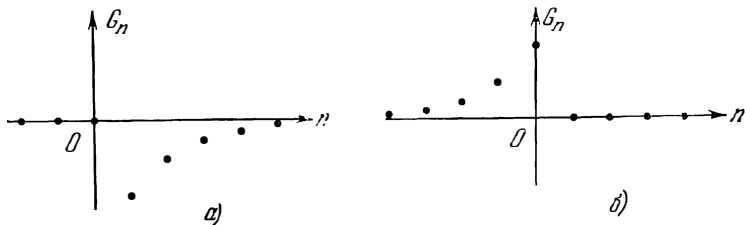


Рис. 2.

Если  $|a/b| > 1$ , то ограниченное решение получается только при  $A = 1/a$  (рис. 2, б):

$$G_n = \begin{cases} \frac{1}{a} \left(-\frac{a}{b}\right)^n, & n \leq 0, \\ 0, & n \geq 1. \end{cases} \quad (4)$$

**3. Частное решение.** Частное решение уравнения

$$au_n + bu_{n+1} = f_n \quad (5)$$

с произвольной правой частью можно записать в виде ряда

$$u_n = \sum_{k=-\infty}^{\infty} G_{n-k} f_k, \quad (6)$$

где  $G_n$  — какое-нибудь фундаментальное решение, если только этот ряд сходится.

Покажем это, воспользовавшись равенством

$$aG_{n-k} + bG_{n-k+1} = \delta_J^{n-k} (= \delta_k^n),$$

которое получается из равенства (1), если в нем всюду заменить  $n$  на  $n - k$ . Подставляя сходящийся ряд (6) в левую часть уравнения (5), получим

$$\begin{aligned} au_n + bu_{n+1} &= a \sum_{k=-\infty}^{\infty} G_{n-k} f_k + b \sum_{k=-\infty}^{\infty} G_{n-k+1} f_k = \\ &= \sum_{k=-\infty}^{\infty} (aG_{n-k} + bG_{n-k+1}) f_k = \sum_{k=-\infty}^{\infty} \delta_k^n f_k = f_n. \end{aligned}$$

Ряд (6) может оказаться расходящимся, если не делать никаких предположений о поведении правой части  $f_k$  разностного уравнения. В самом деле, если положить  $f_k = (-a/b)^k$ , то

$$G_{n-k}f_k = \begin{cases} A \left(-\frac{a}{b}\right)^n & \text{при } n \leq k, \\ \left(A - \frac{1}{a}\right) \left(-\frac{a}{b}\right)^n & \text{при } n \geq k + 1. \end{cases}$$

и ряд (6) при фиксированном  $n$  содержит бесконечное число одинаковых членов, отличных от нуля.

**Теорема.** Пусть  $|a/b| \neq 1$ ,  $G_n$  — ограниченное фундаментальное решение и  $f_k$  ограничены по модулю, т. е.  $|f_k| < F$ . Тогда ряд

$$u_n = \sum_{k=-\infty}^{\infty} G_{n-k}f_k$$

заведомо сходится.

**Доказательство.** Проведем его только для случая  $|a/b| > 1$ . Читатель после этого без труда рассмотрит противоположный случай.

При наших предположениях каждый член ряда

$$u_n = \sum_{k=-\infty}^{\infty} G_{n-k}f_k = \sum_{k=n}^{\infty} \left[ \frac{1}{a} \left(-\frac{a}{b}\right)^{n-k} \right] f_k$$

может быть по абсолютной величине оценен сверху членом сходящейся геометрической прогрессии

$$\left| -\frac{1}{a} \left(-\frac{a}{b}\right)^{n-k} f_k \right| \leq \frac{F}{|a|} \left| \frac{a}{b} \right|^{n-k}.$$

Отсюда следует сходимость ряда (6), а также оценка

$$|u_n| \leq \frac{F}{|a|} \sum_{k=n}^{\infty} \left| \frac{b}{a} \right|^{k-n} = \frac{F}{|a| - |b|}, \quad (7)$$

которая показывает, что решение (6) получилось ограниченным.

Других ограниченных решений уравнение

$$au_n + bu_{n+1} = f_n$$

не имеет, так как любое решение получается из (6) прибавлением некоторого решения  $\bar{u}_n = \alpha \left(-\frac{a}{b}\right)^n$  соответствующего однородного уравнения. Решение  $\{\bar{u}\}$  должно быть ограниченным, как разность двух ограниченных решений, что возможно лишь при  $\alpha = 0$ .



## ЗАДАЧИ

1. Найти общее решение уравнения

$$2u_n - u_{n+1} = 5^n.$$

Решение. Общее решение соответствующего однородного уравнения  $2\bar{u}_n - \bar{u}_{n+1} = 0$  имеет вид  $\bar{u}_n = \alpha 2^n$ . Частное решение  $u_n^*$  будем искать в форме  $u_n^* = C5^n$  с неопределенным коэффициентом. Подставляя  $u_n^* = C5^n$  в уравнение, получим

$$(2 \cdot 5^n - 5^{n+1})C = 5^n; \quad C = -1/3.$$

Таким образом,

$$u_n = -\frac{5^n}{3} + \alpha 2^n.$$

(Заметим, что записать частное решение  $u_n^*$  в виде ряда (6) нельзя, так как его общий член не стремится к нулю, и ряд расходится.)

2. Подобрать частное решение  $u_n^*$  уравнения

$$2u_n - u_{n+1} = 2^n.$$

У к а з а н и е. Ищите решение в виде  $u_n^* = Cn \cdot 2^n$ .

3. Подобрать частные решения  $u_n^*$  уравнения

$$2u_n - u_{n+1} = f_n$$

в случае, если правая часть имеет следующий специальный вид:

$$\text{а) } f_n \equiv 1, \text{ б) } f_n = n, \text{ в) } f_n = n^2, \text{ г) } f_n = 1 + 2n - n^2.$$

4. Подобрать частные решения  $u_n^*$  уравнения

$$u_n - u_{n+1} = f_n,$$

если правая часть  $f_n$  имеет следующий специальный вид:

$$\text{а) } f_n = 1, \text{ б) } f_n = n, \text{ в) } f_n = n^2.$$

### § 3. Разностное уравнение второго порядка

В этом параграфе будет получена формула, выражающая общее решение неоднородного разностного уравнения с постоянными коэффициентами

$$au_{n-1} + bu_n + cu_{n+1} = f_n. \quad (1)$$

В § 1 выяснено, что общее решение имеет вид

$$u_n = u_n^* + \bar{u}_n, \quad (2)$$

где  $u_n^*$  — какое-нибудь частное решение заданного неоднородного уравнения, а

$$\bar{u}_n = \alpha Y_n + \beta Z_n$$

— общее решение соответствующего однородного уравнения

$$au_{n-1} + bu_n + cu_{n+1} = 0. \quad (3)$$

Сначала найдем формулу для общего решения однородного уравнения (3), а потом фундаментальное решение и частное решение неоднородного уравнения.

**1. Общее решение однородного уравнения.** Вспоминая, что в случае разностного уравнения первого порядка существовало частное решение вида  $u_n = q^n$ , попробуем и здесь искать частное решение в виде геометрической прогрессии. Подставим выражение  $u_n = q^n$  в разностное уравнение и убедимся, что оно действительно будет решением, если  $q$  является корнем квадратного уравнения

$$a + bq + cq^2 = 0, \quad (4)$$

называемого *характеристическим уравнением*. Корни этого уравнения могут быть различными или кратными. Рассмотрим последовательно оба случая. Если корни  $q_1$  и  $q_2$  этого характеристического уравнения различны, то мы можем найти в виде геометрической прогрессии даже два независимых частных решения:

$$u_n^{(1)} = q_1^n, \quad u_n^{(2)} = q_2^n.$$

Линейная комбинация

$$\bar{u}_n = \alpha u_n^{(1)} + \beta u_n^{(2)} = \alpha q_1^n + \beta q_2^n \quad (5)$$

этих двух решений с произвольными постоянными коэффициентами  $\alpha$  и  $\beta$  тоже будет решением однородного уравнения. Покажем, что это — общее решение.

Действительно, произвольное частное решение  $\bar{u}_n$  однородного уравнения, принимающее при  $n = 0$  и  $n = 1$  любые наперед заданные значения  $\bar{u}_0$  и  $\bar{u}_1$ , может быть записано в таком виде. Достаточно определить  $\alpha$  и  $\beta$  из равенств

$$\begin{aligned} \alpha + \beta &= \bar{u}_0, \\ \alpha q_1 + \beta q_2 &= \bar{u}_1, \end{aligned}$$

т. е. положить

$$\alpha = \frac{\bar{u}_0 q_2 - \bar{u}_1}{q_2 - q_1}, \quad \beta = \frac{\bar{u}_1 - \bar{u}_0 q_1}{q_2 - q_1}.$$

В частности,  $Y_n$  и  $Z_n$ , определенные в § 1 как решения однородного уравнения, удовлетворяющие условиям

$$\begin{aligned} Y_0 &= 1, \quad Y_1 = 0, \\ Z_0 &= 0, \quad Z_1 = 1, \end{aligned}$$

имеют вид

$$\left. \begin{aligned} Y_n &= \frac{q_2}{q_2 - q_1} q_1^n - \frac{q_1}{q_2 - q_1} q_2^n, \\ Z_n &= -\frac{1}{q_2 - q_1} q_1^n + \frac{1}{q_2 - q_1} q_2^n. \end{aligned} \right\} \quad (6)$$

Из формул (6) видно, что они непригодны в случае кратного корня  $q_1 = q_2$ . Рассмотрим теперь этот случай.

При  $q_1 = q_2$  одно частное решение снова может быть записано в виде  $u_n = q_1^n$ . Чтобы найти второе, сделаем в уравнении (3) подстановку  $u_n = y_n q_1^n$ , после чего получим для  $y_n$  уравнение

$$ay_{n-1} + bq_1 y_n + cq_1^2 y_{n+1} = 0.$$

Как известно,  $a/c$  равно произведению, а  $b/c$  — сумме с обратным знаком корней характеристического уравнения (4). Так как оба эти корня равны  $q_1$ , то

$$\frac{a}{c} = q_1^2, \quad \frac{b}{c} = -2q_1,$$

вследствие чего разностное уравнение для  $y_n$  может быть переписано так:

$$cq_1^2 y_{n-1} - 2cq_1^2 y_n + cq_1^2 y_{n+1} = 0,$$

или несколько проще:

$$y_{n-1} - 2y_n + y_{n+1} = 0.$$

Переписав еще раз это уравнение в виде

$$y_{n-1} - y_n = y_n - y_{n+1},$$

мы видим, что разность  $y_{n-1} - y_n$  не меняется при изменении  $n$ . Таким образом, решением является произвольная арифметическая прогрессия. Нам достаточно найти какое-нибудь одно решение, и мы возьмем арифметическую прогрессию  $y_n = n$ . Вспоминая, что мы искали  $u_n$  в виде  $u_n = y_n q_1^n$ , получаем, что среди решений уравнения  $au_{n-1} + bu_n + cu_{n+1} = 0$  есть решение

$$u_n^{(2)} = nq_1^n.$$

Итак, в случае кратных корней  $q_1 = q_2$  в дополнение к частному решению  $u_n^{(1)} = q_1^n$  мы нашли еще одно независимое частное решение  $u_n^{(2)} = nq_1^n$ .

Линейная комбинация

$$\bar{u}_n = \alpha q_1^n + \beta nq_1^n$$

с произвольными постоянными коэффициентами тоже будет решением однородного уравнения, причем произвольное частное решение можно получить из этой формулы, соответствующим образом подбирая числа  $\alpha$  и  $\beta$ . В частности, решения  $Y_n$  и  $Z_n$  в случае кратных корней имеют вид

$$\left. \begin{aligned} Y_n &= q_1^n - nq_1^n, \\ Z_n &= \frac{1}{q_1} nq_1^n = nq_1^{n-1}. \end{aligned} \right\} \quad (7)$$

Интересно отметить, что формулы (7) могут быть получены из формул (6) для  $Y_n$  и  $Z_n$  в случае некратных корней характеристического уравнения. Тогда мы имели для  $Y_n$  и  $Z_n$  равенства

$$Y_n = \frac{q_2}{q_2 - q_1} q_1^n - \frac{q_1}{q_2 - q_1} q_2^n = -q_1 q_2 \frac{q_2^{n-1} - q_1^{n-1}}{q_2 - q_1},$$

$$Z_n = -\frac{1}{q_2 - q_1} q_1^n + \frac{1}{q_2 - q_1} q_2^n = \frac{q_2^n - q_1^n}{q_2 - q_1}.$$

Заставим корень  $q_2$  приближаться к корню  $q_1$ . При этом выражения

$$\frac{q_2^{n-1} - q_1^{n-1}}{q_2 - q_1} \quad \text{и} \quad \frac{q_2^n - q_1^n}{q_2 - q_1}$$

стремятся к некоторым пределам, а именно соответственно к  $(n-1)q_1^{n-2}$  и  $nq_1^{n-1}$ . Таким образом, мы видим, что в случае кратных корней решения  $Y_n$  и  $Z_n$  примут вид (7).

Итак, мы построили решения  $Y_n$  и  $Z_n$  во всех случаях, которые могут представиться при  $a$  и  $c$ , отличных от нуля.

Тем самым мы показали, что всегда можно выписать в явном виде любое решение интересующего нас однородного разностного уравнения второго порядка.

Интересно остановиться подробнее на случае, когда при вещественных коэффициентах  $a$ ,  $b$ ,  $c$  уравнение  $a + bq + cq^2 = 0$  имеет комплексно-сопряженные корни  $q_1$  и  $q_2$ . Покажем, что в этом случае общее решение однородного разностного уравнения (3) может быть записано в следующем виде:

$$\bar{u}_n = \gamma_1 \left( \sqrt{\frac{a}{c}} \right)^n \cos n\varphi + \gamma_2 \left( \sqrt{\frac{a}{c}} \right)^n \sin n\varphi, \quad (8)$$

где  $\varphi$  определено равенством

$$\cos \varphi = -\frac{b}{2\sqrt{ac}},$$

а  $\gamma_1$  и  $\gamma_2$  — произвольные постоянные.

Найдем явные выражения для  $q_1$  и  $q_2$ :

$$q_1 = \sqrt{\frac{a}{c}} \left[ -\frac{b}{2\sqrt{ac}} + i \sqrt{1 - \left(\frac{b}{2\sqrt{ac}}\right)^2} \right],$$

$$q_2 = \sqrt{\frac{a}{c}} \left[ -\frac{b}{2\sqrt{ac}} - i \sqrt{1 - \left(\frac{b}{2\sqrt{ac}}\right)^2} \right].$$

В нашем случае комплексных корней  $\frac{a}{c} > 0$ ,  $\left| \frac{b}{2\sqrt{ac}} \right| < 1$ . Поэтому мы можем обозначить

$$-\frac{b}{2\sqrt{ac}} = \cos \varphi, \quad \sqrt{1 - \left(\frac{b}{2\sqrt{ac}}\right)^2} = \sin \varphi,$$

после чего  $q_1$  и  $q_2$  запишутся так:

$$q_1 = \sqrt{\frac{a}{c}} (\cos \varphi + i \sin \varphi),$$

$$q_2 = \sqrt{\frac{a}{c}} (\cos \varphi - i \sin \varphi).$$

Подставим эти значения для  $q_1$  и  $q_2$  в формулу (5).

При  $\alpha = \beta = 1/2$  получим частное решение

$$u_n^{(1)} = \left( \sqrt{\frac{a}{c}} \right)^n \cos n\varphi,$$

а при  $\alpha = 1/(2i)$ ,  $\beta = -1/(2i)$  — частное решение

$$u_n^{(2)} = \left( \sqrt{\frac{a}{c}} \right)^n \sin n\varphi.$$

Линейная комбинация этих частных решений с произвольными постоянными коэффициентами  $\gamma_1$  и  $\gamma_2$  и дает общее решение (8), выписанное выше. (Возможность записать в таком виде частное решение (8), принимающее при  $n = 0$  и  $n = 1$  любые наперед заданные значения, читатель легко проверит самостоятельно.)

**2. Общее решение неоднородного уравнения. Фундаментальное решение.** Теперь займемся неоднородным разностным уравнением

$$au_{n-1} + bu_n + cu_{n+1} = f_n, \quad (9)$$

причем ограничимся важным для дальнейшего случаем, когда среди корней характеристического уравнения (4) нет равных единице по модулю:  $|q_1| \neq 1$ ,  $|q_2| \neq 1$ . Сначала будем искать

решение неоднородного уравнения (9) с правой частью  $f_n$  специального вида:

$$f_n = \delta_0^n = \begin{cases} 0, & n \neq 0, \\ 1, & n = 0. \end{cases}$$

Это решение будем обозначать через  $G_n$  и называть *фундаментальным*. Мы будем искать *ограниченное* фундаментальное решение, т. е. ограниченное решение следующих групп уравнений:

- I.  $aG_{n-1} + bG_n + cG_{n+1} = 0$  при  $n \leq -1$ .  
 II.  $aG_{-1} + bG_0 + cG_1 = 1$ .  
 III.  $aG_{n-1} + bG_n + cG_{n+1} = 0$  при  $n \geq 1$ .

Начнем со случая некрратных корней,  $q_1 \neq q_2$ .

В этом случае общее решение однородного уравнения (3) имеет вид

$$u_n = \alpha q_1^n + \beta q_2^n.$$

Поэтому каждое частное решение  $G_n$  однородного уравнения I записывается в форме

$$G_n = \alpha' q_1^n + \beta' q_2^n \quad \text{при } n \leq 0,$$

где  $\alpha'$  и  $\beta'$  — подходящим образом подобранные постоянные. Точно так же частное решение  $G_n$ ,  $n \geq 0$ , однородного уравнения III можно записать в виде

$$G_n = \alpha'' q_1^n + \beta'' q_2^n \quad \text{при } n \geq 0$$

с соответствующими постоянными  $\alpha''$  и  $\beta''$ .

В рассматриваемом нами случае  $q_1 \neq q_2$ ,  $|q_1| \neq 1$ ,  $|q_2| \neq 1$  возможны следующие варианты:

- а)  $|q_1| < 1$ ,  $|q_2| > 1$ ;  
 б)  $|q_1| < 1$ ,  $|q_2| < 1$ ;  
 в)  $|q_1| > 1$ ,  $|q_2| < 1$ ;  
 г)  $|q_1| > 1$ ,  $|q_2| > 1$ .

Построим ограниченное фундаментальное решение  $G_n$  в случае а). Из условия ограниченности  $G_n$  при  $n \rightarrow -\infty$  видно, что  $\alpha' = 0$ , а из условия ограниченности  $G_n$  при  $n \rightarrow +\infty$  следует  $\beta'' = 0$ . Поэтому

$$G_n = \begin{cases} \beta' q_2^n & \text{при } n \leq 0, \\ \alpha'' q_1^n & \text{при } n \geq 0. \end{cases}$$

При  $n = 0$  обе последние формулы должны давать одно и то же значение  $G_0$ . Отсюда  $\beta' = \alpha''$ . Подберем  $\beta'$  из условия выполнения уравнения II:

$$\alpha\beta'q_2^{-1} + b\beta' + c\beta'q_1 = 1,$$

$$\beta' = \frac{1}{aq_2^{-1} + b + cq_1}.$$

Знаменатель этой дроби отличен от нуля:

$$aq_2^{-1} + b + cq_1 = (aq_2^{-1} + b + cq_2) + c(q_1 - q_2) = c(q_1 - q_2) \neq 0.$$

Итак,

$$G_n = \begin{cases} \frac{1}{aq_2^{-1} + b + cq_1} q_2^n, & n \leq 0, \\ \frac{1}{aq_2^{-1} + b + cq_1} q_1^n, & n \geq 0. \end{cases}$$

Мы построили ограниченное фундаментальное решение в случае а) (рис. 3, а).

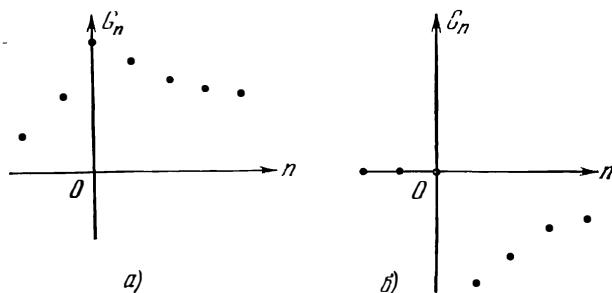


Рис. 3.

Заметим для дальнейшего, что при условиях

$$\left. \begin{aligned} \max(|a|, |b|, |c|) &\geq B > 0, \\ |q_1| &< 1 - \frac{\theta}{2}, \quad |q_2^{-1}| < 1 - \frac{\theta}{2}, \end{aligned} \right\} \quad (10)$$

где  $B > 0$  и  $\theta > 0$  — какие-нибудь числа, имеет место оценка

$$|G_n| \leq \frac{4}{B\theta} \left(1 - \frac{\theta}{2}\right)^{|n|}. \quad (11)$$

Для вывода оценки (11) отметим, что в силу первого условия (10) обязательно либо  $|a| > B/4$ , либо  $|c| > B/4$ , либо  $\sqrt{b^2 - 4ac} \geq \sqrt{B^2 - B^2/4} > B/2$ .

Очевидны также соотношения

$$aq_2^{-1} + b + cq_1 = c(q_1 - q_2) = a(q_2^{-1} - q_1^{-1}) = \sqrt{b^2 - 4ac},$$

$$|q_1 - q_2| \geq |q_2| - |q_1| \geq \frac{1}{1 - \theta/2} - (1 - \theta/2) = \theta \frac{2 - \theta/2}{2 - \theta} > \theta,$$

$$|q_2^{-1} - q_1^{-1}| > \theta.$$

Из этих соотношений следует оценка

$$|aq_2^{-1} + b + cq_1| \geq \frac{B\theta}{4}$$

и неравенство (11).

В случае б) из условия ограниченности  $G_n$  при  $n \rightarrow -\infty$  следует  $\alpha' = \beta' = 0$ , так что

$$G_n = \begin{cases} 0 & \text{при } n \leq 0, \\ \alpha'' q_1^n + \beta'' q_2^n & \text{при } n \geq 0. \end{cases}$$

Из условия  $G_0 = 0$  вытекает  $\alpha'' = -\beta''$ . Коэффициент  $\alpha''$  подбираем так, чтобы удовлетворить уравнению II:

$$\alpha'' = -\frac{1}{c(q_1 - q_2)}.$$

Ограниченное фундаментальное решение (рис. 3, б) в случае б) имеет вид

$$G_n = \begin{cases} 0 & \text{при } n \leq 0, \\ \frac{1}{c(q_1 - q_2)} (q_1^n - q_2^n) & \text{при } n \geq 0. \end{cases}$$

В случае в) по аналогии со случаем а) ограниченное фундаментальное решение имеет вид

$$G_n = \begin{cases} \frac{1}{aq_1^{-1} + b + cq_2} q_1^n & \text{при } n \leq 0, \\ \frac{1}{aq_1^{-1} + b + cq_2} q_2^n & \text{при } n \geq 0. \end{cases}$$

Случай г) аналогичен случаю б).

Если корни кратные,  $q_1 = q_2$ , то при построении ограниченного фундаментального решения вместо формулы

$$u_n = \alpha q_1^n + \beta q_2^n$$

используется формула

$$u_n = \alpha q_1^n + \beta n q_1^n.$$



В случае  $|q_1| < 1$  для  $G_n$  получим

$$G_n = \begin{cases} 0 & \text{при } n \leq 0, \\ \frac{1}{c} n q_1^{n-1} & \text{при } n \geq 1, \end{cases}$$

а в случае  $|q_1| > 1$  получим

$$G_n = \begin{cases} -\frac{1}{a} n q_1^{n+1} & \text{при } n \leq 0, \\ 0 & \text{при } n \geq 1. \end{cases}$$

Итак, мы разобрали все варианты, которые могут представиться в случае  $|q_1| \neq 1$ ,  $|q_2| \neq 1$ ,  $a \neq 0$ ,  $c \neq 0$ , и увидели, что ограниченное фундаментальное решение существует. Из выписанных формул следует, что оно экспоненциально убывает при  $n \rightarrow \infty$ :

$$|G_n| < G \rho^{n+1}, \quad (12)$$

где  $G > 0$  и  $0 < \rho < 1$  — некоторые постоянные.

При этом в качестве  $\rho$  может служить любое число, удовлетворяющее неравенству

$$\rho > \max \left[ \min \left( |q_1|, \frac{1}{|q_1|} \right), \min \left( |q_2|, \frac{1}{|q_2|} \right) \right].$$

Мы выяснили вопрос о существовании и виде фундаментального решения, т. е. решения неоднородного уравнения (9). В случае произвольной правой части  $\{f_n\}$  частное решение  $u_n^*$  можно записать в виде суммы ряда

$$u_n^* = \sum_{k=-\infty}^{\infty} G_{n-k} f_k, \quad (13)$$

если только этот ряд сходится. Это проверяется совершенно так же, как аналогичный факт для разностного уравнения первого порядка в § 2. Из оценки (12) следует, что ряд (13) заведомо сходится, если правая часть  $\{f_k\}$  ограничена,  $|f_k| < F$ . В этом случае

$$\begin{aligned} |u_n^*| &= \left| \sum_{k=-\infty}^{\infty} G_{n-k} f_k \right| \leq \sum_{k=-\infty}^n |G_{n-k} f_k| + \sum_{k=n+1}^{\infty} |G_{n-k} f_k| \leq \\ &\leq GF \left[ \sum_{k=-\infty}^n \rho^{n-k} + \sum_{k=n+1}^{\infty} \rho^{k-n} \right] \leq \frac{2G}{1-\rho} F. \quad (14) \end{aligned}$$

Для уравнения (9), для которого  $|q_1| \neq 1$  и  $|q_2| \neq 1$ , решение  $\{u_n^*\}$ , задаваемое формулой (13), есть единственное ограничен-

нос решение при заданной правой части. В противном случае второе ограниченное решение получалось бы из построенного прибавлением некоторого ограниченного решения  $\{\bar{u}_n\}$  однородного уравнения (3). Но из формул для общего решения этого уравнения видно, что при  $|q_1| \neq 1$ ,  $|q_2| \neq 1$  единственным ограниченным при  $-\infty < n < \infty$  решением будет  $\bar{u}_n \equiv 0$ .

В частности, ограниченное фундаментальное решение  $G_n$  в случае  $|q_1| \neq 1$ ,  $|q_2| \neq 1$  тоже единственно.

Заметим, что при выполнении условий (10), воспользовавшись оценкой (11), из (13) легко вывести

$$|u_n^*| \leq \frac{16}{B\theta^2} \sup_m |f_m|. \quad (15)$$

**3. Оценка фундаментального решения через коэффициенты разностного уравнения.** В п. 2 мы видели, что характер поведения фундаментального решения  $G_n$  уравнения (9) существенно зависит от расположения корней  $q_1$  и  $q_2$  характеристического уравнения

$$P(q) \equiv a + bq + cq^2 = 0 \quad (4)$$

на комплексной плоскости. Для приложений особенно важен случай, когда  $a, b, c$  вещественны, а корни  $q_1$  и  $q_2$  один меньше, а другой больше единицы по модулю:

$$|q_1| < \rho, \quad |q_2^{-1}| < \rho, \quad 0 < \rho < 1. \quad (16)$$

Здесь мы укажем удобный необходимый и достаточный признак такого расположения корней, не требующий их вычисления.

*Теорема. Корни  $q_1$  и  $q_2$  уравнения (4) с вещественными коэффициентами один больше, а другой меньше единицы по модулю в том и только том случае, если выполняется оценка вида*

$$\frac{|b| - |a + c|}{|b| + |a| + |c|} \geq \theta > 0, \quad (17)$$

где  $\theta$  — некоторое число, причем в случае выполнения (17)

$$|q_1| < 1 - \frac{\theta}{2}, \quad |q_2^{-1}| < 1 - \frac{\theta}{2}. \quad (18)$$

*Доказательство.* Заметим, что

$$\begin{aligned} P(1) \cdot P(-1) &= (a + c + b)(a + c - b) = |a + c|^2 - b^2 = \\ &= (|a + c| - |b|)(|a + c| + |b|). \end{aligned}$$

Если (17) не выполнено ни при каком  $\theta > 0$ , то числитель дроби (17) равен нулю или отрицателен.

В первом случае  $P(1) \cdot P(-1) = 0$ , т. е. 1 или  $-1$  является корнем уравнения (4), и (16) не выполнено.

Во втором случае  $P(1)P(-1) > 0$ , т. е. в точках  $q = -1$  и  $q = 1$  многочлен  $P(q)$  принимает значения одного знака. Поэтому многочлен  $P(q)$  не может иметь на отрезке  $-1 \leq q \leq 1$  ровно один корень. Этих корней либо два, либо ни одного.

Если корней два, то оба они меньше единицы по модулю, и (16) не выполнено. Если на отрезке  $[-1, 1]$  корней нет, то либо вещественных корней вообще нет, они комплексно-сопряженные и равные по модулю, либо оба вещественных корня больше единицы по модулю, и (16) снова не выполнено.

Если при некотором  $\theta > 0$  условие (17) выполнено, то  $P(-1)P(1) < 0$ , значения  $P(q)$  на концах отрезка  $[-1, 1]$  имеют разные знаки, так что на этом отрезке лежит ровно один корень. Тогда второй корень, тоже вещественный, лежит вне этого отрезка, так что (16) при некотором  $\rho < 1$  выполнено. Уточним последний результат, а именно получим оценку (18).

Из (17) следует

$$|b| - |a + c| \geq \theta(|b| + |a| + |c|) > \\ > \frac{\theta}{2}|b| + 0 \cdot |a| + \left[\theta - \left(\frac{\theta}{2}\right)^2\right] \cdot |c|.$$

Поэтому

$$|b| \cdot \left(1 - \frac{\theta}{2}\right) > |a + c| + \left[\theta - \left(\frac{\theta}{2}\right)^2\right] \cdot |c| \geq \\ \geq \left|a + c \left\{1 - \theta + \left(\frac{\theta}{2}\right)^2\right\}\right| = \left|a + c \left(1 - \frac{\theta}{2}\right)^2\right|.$$

Отсюда видно, что выражения

$$P\left(1 - \frac{\theta}{2}\right) = a + c \left(1 - \frac{\theta}{2}\right)^2 + b \left(1 - \frac{\theta}{2}\right), \\ P\left[-\left(1 - \frac{\theta}{2}\right)\right] = a + c \left(1 - \frac{\theta}{2}\right)^2 - b \left(1 - \frac{\theta}{2}\right)$$

имеют разные знаки, так что многочлен  $P(q)$  на отрезке  $-(1 - \theta/2) \leq q \leq 1 - \theta/2$  имеет корень  $q_1$ ,  $|q_1| < 1 - \theta/2$ . Очевидно, что числа

$$q'_1 = \frac{1}{q_1}, \quad q'_2 = \frac{1}{q_2},$$

обратные корням уравнения (4), удовлетворяют уравнению

$$a' + b'q' + c'(q')^2 = 0$$

с коэффициентами  $a' = c$ ,  $b' = b$ ,  $c' = a$ , удовлетворяющими тому же условию (17):

$$\frac{|b'| - |a' + c'|}{|b'| + |a'| + |c'|} = \frac{|b| - |a + c|}{|b| + |a| + |c|} \geq \theta > 0.$$

Поэтому один из корней  $q'_1, q'_2$  удовлетворяет неравенству  $|q'| < 1 - \theta/2$ . Этим корнем может быть только  $q'_2 = 1/q_2$ , что и завершает доказательство оценок (18).

Для уравнения с вещественными коэффициентами при условии (17) автоматически выполнены условия (10), а значит, и оценка (15) для ограниченного частного решения  $u_n^*$  неоднородного разностного уравнения (9).

### ЗАДАЧИ

1. Написать общие решения уравнений

$$u_{n-1} - 5u_n + 6u_{n+1} = 0, \quad n = 0, \pm 1, \dots,$$

$$u_{n-1} - \frac{5}{2}u_n + u_{n+1} = 0, \quad n = 0, \pm 1, \dots,$$

$$9u_{n-1} + 3u_n + u_{n+1} = 0, \quad n = 0, \pm 1, \dots$$

2. Найти ограниченное при  $n \rightarrow +\infty$  решение уравнения

$$u_{n-1} - \frac{5}{2}u_n + u_{n+1} = 0,$$

принимая значение  $u_0 = 1$ .

3. Выписать тысячный член последовательности  $u_0, u_1, u_2, \dots$ , первые два члена которой равны единице,  $u_0 = 1, u_1 = 1$ , а последующие определяются рекуррентным соотношением

$$u_{n+1} = u_{n-1} + u_n, \quad n = 1, 2, \dots$$

4. Найти условие, накладываемое на корни характеристического уравнения, необходимое и достаточное для того, чтобы разностное уравнение

$$au_{n-1} + bu_n + cu_{n+1} = 0, \quad n = 0, \pm 1, \pm 2, \dots,$$

имело хотя бы одно нетривиальное ограниченное решение (решение  $u_n \equiv 0$  называется *тривиальным*).

5. Найти условия, которым должны удовлетворять корни характеристического уравнения, необходимые и достаточные для того, чтобы все решения уравнения

$$au_{n-1} + bu_n + cu_{n+1} = 0, \quad n = 0, \pm 1, \dots,$$

были ограничены.

6. Каковы должны быть корни характеристического уравнения, чтобы при  $n \rightarrow +\infty$  все решения уравнения  $au_{n-1} + bu_n + cu_{n+1} = 0$  стремились к нулю?

7. Найти какое-нибудь частное решение неоднородного разностного уравнения

$$u_{n-1} - \frac{5}{2}u_n + u_{n+1} = f_n, \quad n = 0, \pm 1, \dots,$$

если правая часть имеет следующий специальный вид:

а)  $f_n = 1$ . У к а з а н и е. Искать решение вида  $u_n^* = A$ .

б)  $f_n = n$ . У к а з а н и е. Искать решение вида  $u_n^* = A + Bn$ .

в)  $f_n = 3^n$ . У к а з а н и е. Искать решение вида  $u_n^* = A \cdot 3^n$ .

г)  $f_n = \cos n$ . У к а з а н и е. Искать решение вида  $u_n^* = A \sin n + B \cos n$ .

8. Построить какое-нибудь ограниченное фундаментальное решение уравнения

$$u_{n-1} + u_n + u_{n+1} = f_n.$$

Существуют ли у этого уравнения неограниченные фундаментальные решения?

9. Построить какое-нибудь фундаментальное решение уравнения

$$u_{n-1} - 2u_n + u_{n+1} = f_n.$$

Существует ли ограниченное фундаментальное решение?

10. При каком условии на корни характеристического уравнения разностное уравнение

$$au_{n-1} + bu_n + cu_{n+1} = f_n$$

не имеет ограниченных фундаментальных решений?

11. Пользуясь ограниченным фундаментальным решением, выписать то решение  $(u_0, u_1, \dots, u_N)$  уравнения

$$u_{n-1} - \frac{5}{2}u_n + u_{n+1} = f_n, \quad n = 1, 2, \dots, N-1,$$

которое удовлетворяет условиям  $u_0 = \varphi$ ,  $u_N = \psi$ , где  $\varphi$  и  $\psi$  — заданные числа.

12. Найти все собственные числа  $\rho$  и соответствующие собственные функции  $\psi = \{\psi_m\}$ ,  $m = 0, 1, \dots, M$ , оператора  $\Lambda_{xx}$ ,

$$\Lambda_{xx}\psi = \rho\psi,$$

где  $\Lambda_{xx}$  — оператор, который каждой сеточной функции  $u = \{u_m\}$  ставит в соответствие сеточную функцию  $v = \{v_m\}$  по формулам

$$v_m = \frac{1}{h^2}(u_{m+1} - 2u_m + u_{m-1}) \quad 0 < m < M,$$

$$v_0 = v_M = 0, \quad Mh = 1.$$

От в е т:

$$\rho_k = -\frac{4}{h^2} \sin^2 \frac{\pi k}{2M}, \quad \psi_m^{(k)} = \sin \frac{k\pi m}{M}, \quad k = 1, 2, \dots, M-1.$$

## КРАЕВАЯ ЗАДАЧА ДЛЯ УРАВНЕНИЯ ВТОРОГО ПОРЯДКА

Краевые задачи рассматриваемого в этой главе вида возникают при использовании разностных схем для численного решения обыкновенных дифференциальных уравнений и уравнений с частными производными.

### § 4. Постановка задачи. Признаки хорошей обусловленности

**1. Постановка задачи.** Простейшая краевая задача состоит в отыскании сеточной функции  $\{u_n\}$ ,  $n = 0, 1, \dots, N$ , удовлетворяющей разностному уравнению

$$a_n u_{n-1} + b_n u_n + c_n u_{n+1} = f_n, \quad n = 1, 2, \dots, N-1, \quad (1)$$

во внутренних точках  $0 < n < N$  сеточного отрезка  $0 \leq n \leq N$  и принимающей заданные значения

$$u_0 = \varphi, \quad u_N = \psi \quad (2)$$

на его краях. Краевая задача для систем разностных уравнений будет сформулирована в п. 7.

Изучая уравнение  $a_n u_{n-1} + b_n u_n + c_n u_{n+1} = f_n$ ,  $a_n \neq 0$ ,  $c_n \neq 0$ , мы отметили, что при произвольном задании значений  $\{u_n\}$  в каких-нибудь двух последовательных точках, например при произвольном задании  $u_0$  и  $u_1$ , определяется, и притом только одно, решение  $\{u_n\}$ .

Интересно выяснить, можно ли однозначно определить решение, если задаться его значениями в двух не обязательно соседних точках, как это сделано в краевой задаче (1), (2). Следующий пример показывает, что задача (1), (2) может оказаться неразрешимой.

Рассмотрим краевую задачу

$$u_{n-1} - u_n + u_{n+1} = 0, \quad n = 1, 2, \dots, 299, \quad (3)$$

$$u_0 = 0, \quad u_{300} = 1 \quad (4)$$

Общее решение уравнения (3), как показано в § 3, может быть записано в виде

$$u_n = \gamma_1 \cos \frac{n\pi}{3} + \gamma_2 \sin \frac{n\pi}{3}.$$

Из условия  $u_0 = 0$  следует, что  $\gamma_1 = 0$ . Для выполнения условия  $u_{300} = 1$  нужно подобрать  $\gamma_2$  из уравнения

$$u_{300} = \gamma_2 \sin \frac{300\pi}{3} = 1.$$

Но это уравнение неразрешимо, так как при любом  $\gamma_2$  левая часть его равна нулю, но не единице.

Если бы вместо условия  $u_{300} = 1$  мы задали  $u_{300} = 0$  (оставляя по-прежнему  $u_0 = 0$ ), то  $\gamma_1$  снова пришлось бы взять равным нулю, тогда как  $\gamma_2$  в этом случае может быть любым:

$$\gamma_2 \sin \frac{300\pi}{3} = \gamma_2 \cdot 0 = 0.$$

Мы видим, что краевая задача (1), (2) может, вообще говоря, вовсе не иметь решения либо решение ее может оказаться неединственным. Однако с краевыми задачами приходится часто встречаться.

Оказывается, что существуют довольно широкие классы разностных уравнений, для которых краевая задача (1), (2) не только всегда однозначно разрешима, но и обладает «слабой» чувствительностью к ошибкам округления при задании правых частей  $\varphi$ ,  $\psi$  и  $\{f_n\}$ , т. е. «хорошо обусловлена».

**2. Определение хорошей обусловленности.** Обычно при изучении разностных схем для приближенного решения дифференциальных краевых задач рассматривают не одну задачу, а целое семейство таких задач, возникающих при все более мелких шагах сетки. Тогда число  $N$  можно считать параметром, от которого зависит это семейство. Измельчению сетки соответствует возрастание  $N$ .

Будем говорить, что разностная краевая задача (1), (2) с коэффициентами  $a_n$ ,  $b_n$ ,  $c_n$ , ограниченными в совокупности,  $|a_n|$ ,  $|b_n|$ ,  $|c_n| < K$ , *хорошо обусловлена*, если при всех достаточно больших  $N$  она имеет одно и только одно решение  $\{u_n\}$  при произвольных правых частях  $\varphi$ ,  $\psi$  и  $\{f_n\}$  и если числа  $u_0$ ,  $u_1$ , ...,  $u_N$ , образующие решение, удовлетворяют оценке

$$|u_n| \leq M \max \{ |\varphi|, |\psi|, \max_m |f_m| \}, \quad (5)$$

где  $M$  — некоторое число, не зависящее от  $N$ .

Иногда к числу хорошо обусловленных относят и те задачи, для которых  $M$  нельзя выбрать постоянным, но можно выбрать растущим не быстрее заданной степени  $N$ , например  $M = CN$  или  $M = CN^2$ .

Приведенное определение хорошей обусловленности равносильно одному из принятых в теории систем линейных уравнений, когда мерой обусловленности системы уравнений  $Ax = g$  с матрицей  $A$  считают число  $\|A\| \cdot \|A^{-1}\|$  — произведение норм матриц  $A$  и  $A^{-1}$ .

Выполнение неравенства (5) означает, что чувствительность решения  $\{u_n\}$  к ошибкам (например, ошибкам измерения или округления), допущенным при задании правых частей  $\varphi$ ,  $\psi$  и  $\{f_n\}$ , не возрастает с ростом числа  $N$ . Действительно, если вместо  $\varphi$ ,  $\psi$  и  $\{f_n\}$  задать соответственно  $\varphi + \Delta\varphi$ ,  $\psi + \Delta\psi$ ,  $\{f_n + \Delta f_n\}$ , то решение  $\{u_n\}$  получит приращение  $\{\Delta u_n\}$ . Это приращение ввиду линейности задачи (1), (2) является решением задачи

$$\left. \begin{aligned} a_n \Delta u_{n-1} + b_n \Delta u_n + c_n \Delta u_{n+1} &= \Delta f_n, & 0 < n < N, \\ \Delta u_0 &= \Delta\varphi, & \Delta u_N &= \Delta\psi \end{aligned} \right\}$$

и в силу (5) удовлетворяет оценке

$$|\Delta u_n| \leq M \max \{ |\Delta\varphi|, |\Delta\psi|, \max_m |\Delta f_m| \}.$$

Далеко не всякая однозначно разрешимая краевая задача (1), (2) является хорошо обусловленной. Например, если правым частям задачи

$$\left. \begin{aligned} u_{n+1} - 5u_n + 6u_{n-1} &= f_n, & 0 < n < N, \\ u_0 &= \varphi, & u_N &= \psi \end{aligned} \right\}$$

придать приращения

$$\Delta f_n \equiv 0, \quad \Delta\psi = 0, \quad \Delta\varphi = \varepsilon,$$

то решение  $\{u_n\}$  получит приращение

$$\Delta u_n = 2^n \frac{1 - (2/3)^{N-n}}{1 - (2/3)^N} \Delta\varphi, \quad n = 0, 1, \dots, N.$$

Отсюда

$$\Delta u_{N-1} \geq 2^{N-1} \cdot \frac{1}{3} \varepsilon.$$

Возмущение  $\varepsilon$  при задании  $\varphi$  вызвало быстро возрастающее с ростом  $N$  возмущение решения. Число  $M$  в неравенстве (5) заведомо нельзя взять растущим медленнее экспоненты  $\frac{1}{3} \cdot 2^{N-1}$ .

### 3. Достаточный признак хорошей обусловленности.

**Теорема.** Если коэффициенты  $a_n$ ,  $b_n$ ,  $c_n$  удовлетворяют условию

$$|b_n| \geq |a_n| + |c_n| + \delta, \quad \delta > 0, \quad (6)$$



то задача (1), (2) хорошо обусловлена, причем решение  $\{u_n\}$  удовлетворяет оценке

$$|u_n| \leq \max \left\{ |\varphi|, |\psi|, \frac{1}{\delta} \max_m |f_m| \right\}. \quad (7)$$

**Доказательство.** Сначала предположим, что при заданных фиксированных  $\varphi$ ,  $\psi$  и  $\{f_n\}$  задача (1), (2) имеет решение  $\{u_n\}$ , и установим для него оценку (7). Пусть наибольшее среди чисел  $|u_n|$ ,  $n = 0, 1, \dots, N$ , есть число  $|u_k|$ . Если  $k = 0$  или  $k = N$ , то неравенство (7) очевидно, так как  $u_0 = \varphi$ ,  $u_N = \psi$ . Остается рассмотреть случай  $0 < k < N$ ,  $|u_k| \geq |u_n|$ . В этом случае, с учетом (6), можно написать

$$\begin{aligned} |b_k| \cdot |u_k| &= | -a_k u_{k-1} - c_k u_{k+1} + f_k | \leq \\ &\leq |a_k| \cdot |u_{k-1}| + |c_k| \cdot |u_{k+1}| + |f_k| \leq (|a_k| + |c_k|) |u_k| + |f_k|, \\ |u_n| &\leq |u_k| \leq \frac{|f_k|}{|b_k| - |a_k| - |c_k|} \leq \frac{|f_k|}{\delta}, \end{aligned}$$

и неравенство (6) также выполнено.

Осталось доказать, что задача (1), (2) имеет, и притом только одно, решение  $\{u_n\}$  при произвольных правых частях  $\varphi$ ,  $\psi$  и  $\{f_n\}$ .

Задачу (1), (2) можно рассматривать как систему  $N + 1$  линейных уравнений относительно такого же числа неизвестных  $u_0, u_1, \dots, u_N$ . Поэтому нужно установить, что определитель этой системы отличен от нуля. Как известно из алгебры, определитель системы отличен от нуля в том и только том случае, если соответствующая однородная система имеет лишь нулевое решение. Но для системы (1), (2) однородная система получается при  $\varphi = \psi = f_m \equiv 0$ . Из оценки (7), доказанной для каждого решения  $\{u_n\}$ , видно, что в этом случае имеется только тривиальное решение  $u_n \equiv 0$ .

Достаточным условием хорошей обусловленности задачи (1), (2) является также следующее условие:

$$\frac{|b_n| - |a_n| - |c_n|}{|b_n| + |a_n| + |c_n|} \geq \theta > 0, \quad \max \{ |a_n|, |b_n|, |c_n| \} \geq B > 0, \quad (8)$$

где  $\theta$  и  $B$  — некоторые постоянные, не зависящие от  $N$  и  $n$ .

Действительно, из (8) следует (6) с постоянной

$$\delta = \theta (|b_n| + |a_n| + |c_n|) \geq \theta B > 0.$$

Поэтому (7) примет вид

$$|u_n| \leq \max \left\{ |\varphi|, |\psi|, \frac{1}{\theta B} \max_m |f_m| \right\}. \quad (9)$$

#### 4. Критерий хорошей обусловленности краевой задачи с постоянными коэффициентами.

**Теорема.** Для хорошей обусловленности краевой задачи

$$\left. \begin{aligned} au_{n-1} + bu_n + cu_{n+1} = f_n, \quad 0 < n < N, \\ u_0 = \varphi, \quad u_N = \psi \end{aligned} \right\} \quad (10)$$

с постоянными коэффициентами необходимо и достаточно, чтобы корни  $q_1$  и  $q_2$  характеристического уравнения

$$a + bq + cq^2 = 0 \quad (11)$$

были по модулю один больше, а другой меньше единицы, т. е. чтобы удовлетворялись неравенства вида

$$|q_1| \leq 1 - \frac{\theta}{2}, \quad |q_2^{-1}| \leq 1 - \frac{\theta}{2}, \quad (12)$$

где  $\theta$  — некоторая положительная постоянная.

В случае, если коэффициенты  $a$ ,  $b$ ,  $c$  вещественны, критерию хорошей обусловленности (12) в силу доказанного в п. 3 § 3 можно придать удобную форму:

$$\frac{|b| - |a + c|}{|b| + |a| + |c|} \geq \theta > 0. \quad (13)$$

Удобство критерия (13) состоит в том, что его выполнение проверяется непосредственно, без вычисления корней  $q_1$  и  $q_2$ .

Доказательство критерия (12) будет проведено в п. 6 этого параграфа.

**5. Критерий хорошей обусловленности задачи с переменными коэффициентами.** Критерий (12) хорошей обусловленности краевой задачи для разностного уравнения с постоянными коэффициентами, сформулированный в предыдущем пункте, обобщается на случай задачи

$$a_n u_{n-1} + b_n u_n + c_n u_{n+1} = f_n, \quad 0 < n < N, \quad (1)$$

$$u_0 = \varphi, \quad u_N = \psi \quad (2)$$

с переменными коэффициентами, если только эти коэффициенты изменяются достаточно «плавно». Сформулируем это обобщение точно, причем относительно уравнения (1) будем предполагать, что его коэффициенты ограничены в совокупности,  $|a_n| < M$ ,  $|b_n| < M$ ,  $|c_n| < M$ , и что все три коэффициента  $a_n$ ,  $b_n$ ,  $c_n$  ни при каком  $n$  одновременно не становятся малыми:

$$d_n = \max\{|a_n|, |b_n|, |c_n|\} \geq B > 0.$$

Предполагается, что  $M$  и  $B$  не зависят от  $N$  и  $n$ .

**Теорема.** Пусть коэффициенты задачи (1), (2) удовлетворяют условиям

$$\left. \begin{aligned} |a_k - a_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, & |b_k - b_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, \\ |c_k - c_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, & D > 0, & \omega > 0. \end{aligned} \right\} \quad (14)$$

Тогда для хорошей обусловленности задачи (1), (2) необходимо и достаточно, чтобы корни  $q_1$  и  $q_2$  квадратного уравнения

$$a_n + b_n q + c_n q^2 = 0, \quad 0 < n < N, \quad (15)$$

удовлетворяли условию вида

$$|q_1| < 1 - \frac{\theta}{2}, \quad |q_2^{-1}| < 1 - \frac{\theta}{2}, \quad (16)$$

где  $\theta > 0$  — некоторое число, не зависящее от  $N$  и  $n$ .

Условия (14) выражают требование гладкости коэффициентов. Они выполнены, например, если

$$a_n = a(n/N), \quad b_n = b(n/N), \quad c_n = c(n/N),$$

где  $a(x)$ ,  $b(x)$ ,  $c(x)$  — некоторые функции, определенные на отрезке  $0 \leq x \leq 1$  и удовлетворяющие условию Гёльдера:

$$|a(x) - a(x')| \leq D |x - x'|^\omega,$$

$$|b(x) - b(x')| \leq D |x - x'|^\omega,$$

$$|c(x) - c(x')| \leq D |x - x'|^\omega.$$

Уравнение (15) является характеристическим уравнением, построенным для разностного уравнения

$$au_{s-1} + bu_s + cu_{s+1} = 0$$

с постоянными коэффициентами  $a$ ,  $b$ ,  $c$ , совпадающими со значениями переменных коэффициентов  $a_n$ ,  $b_n$ ,  $c_n$  при зафиксированном  $n$ , т. е.  $a = a_n$ ,  $b = b_n$ ,  $c = c_n$ .

Если  $a_n$ ,  $b_n$ ,  $c_n$  — вещественные коэффициенты, то в силу п. 3 § 3 условие (16) можно заменить легко проверяемым условием

$$\frac{|b_n| - |a_n + c_n|}{|b_n| + |a_n| + |c_n|} \geq \theta > 0, \quad (17)$$

где  $\theta$  не зависит от  $N$  и  $n$ .

Сформулированный критерий (14), (16) или (14), (17) будет доказан в § 6. Там же будет показано, что условия гладкости (14) игнорировать нельзя.

Заметим, что если  $|a_n + c_n| = |a_n| + |c_n|$ , то условие (17) совпадает с условием (8) и обеспечивает хорошую обусловленность и без предположений о гладкости и вещественности коэффициентов.

**6. Обоснование критерия хорошей обусловленности краевой задачи с постоянными коэффициентами.** Докажем сформулированный в п. 4 критерий хорошей обусловленности краевой задачи

$$\left. \begin{aligned} au_{n-1} + bu_n + cu_{n+1} = f_n, \quad 0 < n < N, \\ u_0 = \varphi, \quad u_N = \psi, \end{aligned} \right\} \quad (10)$$

а именно следующее утверждение. Для хорошей обусловленности задачи (10) необходимо и достаточно, чтобы корни характеристического уравнения

$$a + bq + cq^2 = 0 \quad (11)$$

удовлетворяли неравенствам вида

$$|q_1| \leq 1 - \frac{\theta}{2}, \quad |q_2^{-1}| \leq 1 - \frac{\theta}{2}, \quad (12)$$

где  $\theta$  — некоторая положительная постоянная.

**Достаточность.** Решение задачи (10) представим в виде суммы двух сеточных функций, положив

$$u_n = \bar{u}_n + \tilde{u}_n, \quad (18)$$

где  $\{\bar{u}_n\}$  — решение задачи

$$\left. \begin{aligned} a\bar{u}_{n-1} + b\bar{u}_n + c\bar{u}_{n+1} = 0, \quad 0 < n < N, \\ \bar{u}_0 = \varphi, \quad \bar{u}_N = \psi, \end{aligned} \right\} \quad (19)$$

а  $\{\tilde{u}_n\}$  — решение задачи

$$\left. \begin{aligned} a\tilde{u}_{n-1} + b\tilde{u}_n + c\tilde{u}_{n+1} = f_n, \quad 0 < n < N, \\ \tilde{u}_0 = 0, \quad \tilde{u}_N = 0. \end{aligned} \right\} \quad (20)$$

Решение задачи (19) имеет вид  $\bar{u}_n = Aq_1^n + Bq_2^n$ , где  $A$  и  $B$  определяются из условий  $\bar{u}_0 = \varphi$ ,  $\bar{u}_N = \psi$ :

$$\bar{u}_n = \frac{\varphi - \psi q_2^{-N}}{1 - (q_1 q_2^{-1})^N} q_1^n + \frac{\psi - \varphi q_1^N}{1 - (q_1 q_2^{-1})^N} q_2^{n-N}. \quad (21)$$

Обозначив  $1 - \theta/2 = \rho$ , из (21) получаем

$$|\bar{u}_n| \leq 2 \frac{\max(\rho^n, \rho^{N-n})}{1 - \rho^{2N}} \max(|\varphi|, |\psi|). \quad (22)$$

Поэтому при всех  $N \geq 2$  и  $n = 0, 1, \dots, N$

$$|\bar{u}_n| \leq \frac{2}{1-\rho} \max(|\varphi|, |\psi|) = \frac{4}{\theta} \max(|\varphi|, |\psi|). \quad (23)$$

Если  $n$  и  $N - n$  — достаточно большие числа, то коэффициент в неравенстве (22) сколь угодно мал. Например, при  $n > 6/\theta$ ,  $N - n > 6/\theta$

$$\rho^{6/\theta} = \left[ \left(1 - \frac{\theta}{2}\right)^{2/\theta} \right]^3 < \left(\frac{4}{9}\right)^3.$$

Здесь использовано известное неравенство

$$\left(1 - \frac{1}{a}\right)^a \left(1 + \frac{1}{b}\right)^b \leq \left[ \frac{a(1 - a^{-1}) + b(1 + b^{-1})}{a + b} \right]^{a+b} = 1$$

при  $a = 2/\theta$ ,  $b = 2$ . Таким образом,

$$\frac{\max(\rho^n, \rho^{N-n})}{1 - \rho^{2N}} \leq \left(\frac{4}{9}\right)^3 \frac{1}{1 - (4/9)^6} < \frac{1}{10},$$

так что из (22) при  $n > 6/\theta$ ,  $N - n > 6/\theta$  получим

$$|\bar{u}_n| < \frac{1}{5} \max(|\varphi|, |\psi|). \quad (24)$$

Оценим решение  $\{\bar{u}_n\}$  задачи (20). Представим  $\bar{u}_n$  в виде суммы

$$\bar{u}_n = u_n^* + u_n', \quad 0 \leq n \leq N,$$

решений двух задач — задачи

$$au_{n-1}^* + bu_n^* + cu_{n+1}^* = \begin{cases} f_n, & 0 < n < N, \\ 0, & n \leq 0 \text{ или } n \geq N, \end{cases} \quad (25)$$

и задачи

$$\left. \begin{aligned} au'_{n-1} + bu'_n + cu'_{n+1} &= 0, & 0 < n < N, \\ u'_0 &= -u_0^*, & u'_N &= -u_N^*. \end{aligned} \right\} \quad (26)$$

Ограниченное решение  $\{u_n^*\}$  задачи (25) существует, единственно и удовлетворяет оценке (15) § 3:

$$|u_n^*| \leq \frac{16}{B\theta^2} \max_m |f_m|, \quad (27)$$

где  $B = \max(|a|, |b|, |c|)$ .

В частности,

$$\left. \begin{aligned} |u_0^*| &\leq \frac{16}{B\theta^2} \max_m |f_m|, \\ |u_N^*| &\leq \frac{16}{B\theta^2} \max_m |f_m|. \end{aligned} \right\} \quad (27')$$

Для оценки решения  $\{u'_n\}$  задачи (26), совпадающей по своему виду с задачей (19), воспользуемся формулой (21) и оценкой (23), заменив только  $\varphi$  и  $\psi$  на  $-u'_0$  и  $-u'_N$ :

$$|u'_n| \leq \frac{4}{\theta} \max(|u'_0|, |u'_N|).$$

Теперь примем еще во внимание (27'):

$$|u'_n| \leq \frac{64}{B\theta^3} \max_m |f_m|. \quad (28)$$

Объединяя оценки (27) и (28) с учетом  $\theta < 2$ , получим

$$|\tilde{u}_n| \leq \frac{128}{B\theta^3} \max_m |f_m|. \quad (29)$$

Следовательно, для решения  $\{u_n\}$  исходной задачи, объединяя оценки (23) и (29), получим

$$|u_n| \leq |\bar{u}_n| + |\tilde{u}_n| \leq \frac{128}{B\theta^3} \max_m |f_m| + \frac{4}{\theta} \max(|\varphi|, |\psi|). \quad (30)$$

Оценка (30) обеспечивает хорошую обусловленность  $|u_n| \leq M \max(|\varphi|, |\psi|, \max_m |f_m|)$ , причем за  $M$  можно принять

$$M = \frac{128}{B\theta^3} + \frac{4}{\theta}.$$

В случае  $n > 6/\theta$ ,  $N - n > 6/\theta$  можно уточнить оценку (30), воспользовавшись вместо неравенства (23) неравенством (24):

$$|u_n| \leq \frac{128}{B\theta^3} \max_m |f_m| + \frac{1}{5} \max(|\varphi|, |\psi|) \quad (31)$$

или

$$|u_n| < M_1 \max_m |f_m| + \frac{1}{5} \max(|\varphi|, |\psi|), \quad (31')$$

где  $M_1$  зависит только от  $\theta$  и  $B$ , но не от  $N$ . Оценкой (31) мы будем пользоваться в § 6.

Необходимость. Заметим сначала, что если условия (12) не выполнены ни при каком положительном  $\theta$ , то корни характеристического уравнения

$$P(q) \equiv a + bq + cq^2 = 0$$

по модулю либо оба меньше единицы, либо оба больше единицы, либо хотя бы один из них равен единице:

$$1) \quad |q_1| < \rho < 1, \quad |q_2| < \rho < 1, \quad (32)$$

$$2) \quad |q_1| > \rho > 1, \quad |q_2| > \rho > 1, \quad (33)$$

$$3) \quad |q_1| = 1. \quad (34)$$

Покажем, что во всех трех случаях хорошей обусловленности нет.

Для этого во всех трех случаях построим некоторые функции  $\{u_n\}$  так, чтобы они были решениями задачи вида

$$\left. \begin{aligned} au_{n-1} + bu_n + cu_{n+1} &= f_n, & 0 < n < N, \\ u_0 &= u_N = 0 \end{aligned} \right\} \quad (35)$$

и чтобы выполнялись неравенства

$$\max_n |u_n| > M_N \max_m |f_m|, \quad (36)$$

где  $M_N$  — некоторая неограниченно возрастающая при  $N \rightarrow \infty$  величина. В случае (32), считая для определенности, что  $q_1 \neq q_2$ , положим

$$u_n = \begin{cases} q_1^n - q_2^n, & 0 \leq n \leq N-1, \\ 0, & n = N. \end{cases}$$

Тогда

$$\max_n |u_n| \geq |u_1| = |q_1 - q_2| > 0. \quad (37)$$

Правая часть  $\{f_n\}$  в задаче (35) есть

$$f_n = au_{n-1} + bu_n + cu_{n+1} = \begin{cases} 0, & \text{если } n \neq N-1, \\ -c(q_1^N - q_2^N), & \text{если } n = N-1. \end{cases}$$

Отсюда

$$\max_n |f_n| = |f_{N-1}| \leq 2|c|\rho^N. \quad (38)$$

Сопоставляя (37) и (38), видим, что в неравенстве (36) надо положить

$$M_N = \frac{|q_1 - q_2|}{2|c|\rho^N} = O\left(\frac{1}{\rho^N}\right),$$

так что  $M_N$  экспоненциально растет с ростом  $N$ .

Случай (33) аналогичен случаю (32).

Если выполнено (34), то положим

$$u_n = q_1^n \sin \frac{n\pi}{N}, \quad 0 \leq n \leq N$$

Тогда, очевидно,

$$\max_n |u_n| \geq \frac{1}{2}. \quad (39)$$

Для  $|f_n|$  получаем оценку

$$\begin{aligned} |f_n| &= |au_{n-1} + bu_n + cu_{n+1}| = \\ &= \left| (aq_1^{n-1} + bq_1^n + cq_1^{n+1}) \sin \frac{n\pi}{N} + aq_1^{n-1} \left( \sin \frac{(n-1)\pi}{N} - \sin \frac{n\pi}{N} \right) + \right. \\ &+ \left. cq_1^{n+1} \left( \sin \frac{(n+1)\pi}{N} - \sin \frac{n\pi}{N} \right) \right| = \left| aq_1^{n-1} \left( \sin \frac{(n-1)\pi}{N} - \sin \frac{n\pi}{N} \right) + \right. \\ &+ \left. cq_1^{n+1} \left( \sin \frac{(n+1)\pi}{N} - \sin \frac{n\pi}{N} \right) \right| \leq (|a| + |c|) \frac{\pi}{N}. \quad (40) \end{aligned}$$

Из (39) и (40) следует, что неравенство (36) выполнено, если

$$M_N = \frac{N}{2(|a| + |c|)}.$$

Таким образом, здесь нет хорошей обусловленности, если понимать под ней требование независимости  $M$  от  $N$  в неравенстве (5).

**7. Общие краевые задачи для систем разностных уравнений.** Задача (1), (2) является лишь простейшей краевой задачей для уравнения второго порядка.

Сформулируем без доказательства необходимые и достаточные условия хорошей обусловленности общих краевых задач для систем разностных уравнений на сеточном отрезке (В. С. Рябенский, ЖВМ и МФ 4, 2 (1964)).

Краевая задача состоит в отыскании вектор-функции  $\{u_n\}$ ,  $n = 0, 1, 2, 3, \dots, N$ , удовлетворяющей условиям

$$\sum_{k=-k_0}^{k_0} A_{k,n} u_{n+k} = f_n, \quad k_0 \leq n \leq N - k_0, \quad (1')$$

$$\sum_{i=0}^{2k_0} \alpha_i u_i = \varphi, \quad \sum_{i=0}^{2k_0} \beta_i u_{N-i} = \psi. \quad (2')$$

Здесь  $A_{k,n}$  — квадратные матрицы некоторого порядка  $m \geq 1$ ;  $u_n, f_n$  — векторы той же размерности;  $\alpha_i$  — матрицы, имеющие по  $m$  столбцов и  $r \geq 0$  строк;  $\beta_i$  — матрицы, имеющие по  $m$  столбцов и  $s \geq 0$  строк;  $\varphi$  — заданный  $r$ -мерный вектор;  $\psi$  — заданный  $s$ -мерный вектор.

Задача (1'), (2') хорошо обусловлена, если она имеет решение  $\{u_n\}$  при произвольных  $\{f_n\}$ ,  $\varphi, \psi$ , причем

$$\max_n \|u_n\| \leq M \max \{ \|\varphi\|, \|\psi\|, \max_j \|f_j\| \},$$

где  $M$  не зависит от  $N$ .

Относительно коэффициентов  $A_{k,n}$  будем предполагать, что

$$A_{k,n} = A_k \left( \frac{n}{N} \right),$$

где  $A_k(x)$  — матрица, определенная на отрезке  $0 \leq x \leq 1$ , удовлетворяющая на этом отрезке условию гладкости

$$\|A_k(x) - A_k(x')\| \leq D |x - x'|^\omega, \quad D > 0, \quad \omega > 0. \quad (14')$$

Далее, предположим, что

$$d(x) = \max_k \|A_k(x)\| \geq B > 0.$$

При этих ограничениях для хорошей обусловленности задачи (1'), (2') необходимо и достаточно, чтобы выполнялось каждое из следующих условий 1°—3°:

1° Среди корней  $\mu$  и  $\nu$  уравнений

$$\det \sum_{k=-k_0}^{k_0} A_k(x) \mu^{k_0+k} = 0$$

$$\det \sum_{k=-k_0}^{k_0} A_k(x) \nu^{k_0-k} = 0$$



нет равных единице по модулю, причем корни  $\mu$  и  $\nu$  этих уравнений удовлетворяют каждому одному из следующих четырех неравенств:

$$\begin{aligned} |\mu| < 1 - \frac{\theta}{2}, \quad |\nu| < 1 - \frac{\theta}{2}, \\ |\mu^{-1}| < 1 - \frac{\theta}{2}, \quad |\nu^{-1}| < 1 - \frac{\theta}{2}, \end{aligned}$$

где  $\theta > 0$  не зависит от  $x$ .

2° Размерность  $r$  матриц  $\alpha_i$  равна числу тех корней  $\mu$ , модуль которых меньше единицы, а размерность  $s$  матриц  $\beta_i$  равна числу тех корней  $\nu$ , модуль которых меньше единицы.

3° Среди решений  $\{u_n\}$ ,  $n \geq 0$ , задачи

$$\left. \begin{aligned} \sum_{k=-k_0}^{k_0} A_k(0) u_{n+k} = 0, \quad k_0 \leq n < \infty, \\ \sum_{i=0}^{2k_0} \alpha_i u_i = 0 \end{aligned} \right\}$$

и среди решений  $\{u_n\}$ ,  $n \leq N$ , задачи

$$\left. \begin{aligned} \sum_{k=-k_0}^{k_0} A_k(1) u_{n+k} = 0, \quad -\infty < n \leq N - k_0, \\ \sum_{i=0}^{2k_0} \beta_i u_{N-i} = 0 \end{aligned} \right\}$$

нет ограниченных, отличных от тождественного нуля.

Последнему условию, 3°, можно придать вид необращения в нуль некоторых определителей с элементами, не зависящими от  $N$ .

Проиллюстрируем сформулированный критерий, исследовав условия хорошей обусловленности задачи

$$\left. \begin{aligned} 0 \cdot u_{n-1} - 2u_n + u_{n+1} = f_n, \quad 0 < n < N, \\ u_1 - \alpha u_0 = \varphi, \quad u_N - \beta u_{N-1} = \psi, \end{aligned} \right\}$$

где  $\alpha$  и  $\beta$  — некоторые числа;  $m = 1$ ,  $r = 1$ ,  $s = 1$ ,  $k_0 = 1$ . Корни уравнений

$$0 - 2\mu + \mu^2 = 0 \quad \text{и} \quad 0 \cdot \nu^2 - 2\nu + 1 = 0$$

равны

$$\mu_1 = 0, \quad \mu_2 = 2, \quad \nu_1 = 1/2 \quad (\nu_2 = \infty).$$

Среди них нет равных единице по модулю, и условие 1° выполнено.

Условие 2° тоже выполнено, так как количество скалярных граничных условий на левой и правой границах равно  $r = s = 1$  и равно числу тех корней  $\mu$  и  $\nu$ , которые меньше единицы по модулю.

Выясним, при каких значениях  $\alpha$  задача

$$\left. \begin{aligned} 0 \cdot u_{n-1} - 2u_n + u_{n+1} = 0, \quad n \geq 1, \\ \alpha u_0 - u_1 = 0, \end{aligned} \right\}$$

не имеет нетривиальных ограниченных решений. Общий вид решения задачи

$$0 \cdot u_{n-1} - 2u_n + u_{n+1} = 0, \quad n > 0.$$

есть

$$u_n = c_1 \mu_1^n + c_2 \mu_2^n, \quad n \geq 0, \quad \mu_1^0 = 1.$$

Из условия ограниченности находим  $c_2 = 0$ . Поэтому

$$u_n = c_1 \mu_1^n = \begin{cases} c_1, & \text{если } n = 0, \\ 0, & \text{если } n > 0. \end{cases}$$

Учитывая условие  $\alpha u_0 - u_1 = 0$ , видим, что при  $\alpha \neq 0$  нетривиальных решений нет, а при  $\alpha = 0$  они есть.

Выясним, при каких  $\beta$  задача

$$\left. \begin{aligned} 0 \cdot u_{n-1} - 2u_n + u_{n+1} &= 0, \quad n < N, \\ u_N - \beta u_{N-1} &= 0 \end{aligned} \right\}$$

не имеет ограниченных при  $n \rightarrow -\infty$  нетривиальных решений. Общее решение задачи  $0 \cdot u_{n-1} - 2u_n + u_{n+1} = 0$ ,  $n < N$ , есть  $u_n = c_1 \nu_1^{-n} = c_1 (1/2)^{-n} = c_1 2^n$ . Оно ограничено при  $n \rightarrow -\infty$ . Из граничного условия  $u_N - \beta u_{N-1} = 0$  видно, что

$$c_1 2^N - \beta c_1 2^{N-1} = c_1 2^{N-1} (2 - \beta) = 0$$

и нетривиальное решение,  $c_1 \neq 0$ , существует только при  $\beta = 2$ .

Итак, рассматриваемая краевая задача хорошо обусловлена при любых  $\alpha \neq 0$  и  $\beta \neq 2$ . Если  $\alpha = 0$  или  $\beta = 2$ , задача не является хорошо обусловленной.

## ЗАДАЧИ

Разностную краевую задачу

$$\left. \begin{aligned} au_{n-1} + bu_n + cu_{n+1} &= f_n, \quad 0 < n < N, \\ u_0 - \alpha u_1 &= \varphi, \quad u_N - \beta u_{N-1} = \psi \end{aligned} \right\} \quad (*)$$

будем называть *хорошо обусловленной*, если она имеет одно и только одно решение при каждом  $N$ , и если числа  $u_0, u_1, \dots, u_N$ , образующие решение  $\{u_n\}$ , удовлетворяют неравенству  $|u_n| < M \max \{ |\varphi|, |\psi|, \max_m |f_m| \}$ , где  $M$  от  $N$  не зависит.

1. Если оба корня  $q_1$  и  $q_2$  характеристического уравнения  $a + bq + cq^2 = 0$  по модулю меньше (больше) единицы, то разностная краевая задача (\*) не может быть хорошо обусловлена. Для простоты считать  $q_1 \neq q_2$ . Доказать.

2. Если хотя бы один из корней  $q_1, q_2$  характеристического уравнения по модулю равен единице, то разностная краевая задача (\*) не может быть хорошо обусловлена. Доказать.

3. Если  $|q_1| < 1$ ,  $|q_2| > 1$ , но

$$1 - \alpha q_1 = 0 \quad \text{или} \quad 1 - \beta q_2 = 0,$$

то задача (\*) не может быть хорошо обусловлена. Доказать.

4. Для хорошей обусловленности разностной краевой задачи (\*) необходимо и достаточно, чтобы один корень характеристического уравнения по модулю был меньше единицы,  $|q_1| < 1$ , а второй больше единицы,  $|q_2| > 1$ , и чтобы  $1 - \alpha q_1 \neq 0$ ,  $1 - \beta q_2 \neq 0$ . Доказать.

5. Задача с постоянными (комплексными) коэффициентами

$$au_{n-1} + bu_n + cu_{n+1} = f_n, \quad n = 0, \pm 1, \dots$$

с произвольной периодической правой частью

$$f_{n+N} = f_n$$

имеет при всех достаточно больших  $N$  периодическое решение  $\{u_n\}$ ,  $u_{n+N} = u_n$ , удовлетворяющее оценке

$$|u_n| \leq M \max_m |f_m|.$$

где  $M$  от  $N$  и от  $\{f_n\}$  не зависит, в том случае, если среди корней характеристического уравнения  $a + bq + cq^2 = 0$  нет равных единице по модулю. Доказать.

## § 5. Алгоритм решения краевой задачи — прогонка

**1. Описание прогонки.** Опишем теперь простой и удобный метод решения разностной краевой задачи рассмотренного нами в § 4 вида:

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, & 0 < n < N, \\ u_0 &= \varphi, & u_N &= \psi. \end{aligned} \right\} \quad (1)$$

Он представляет собою один из вариантов метода исключения неизвестных и носит название *метода прогонки*.

Запишем уравнение  $u_0 = \varphi$  системы (1) в виде

$$u_0 = L_{1/2} u_1 + K_{1/2},$$

где  $L_{1/2} = 0$  и  $K_{1/2} = \varphi$ . Из уравнения

$$a_1 u_0 + b_1 u_1 + c_1 u_2 = f_1,$$

отвечающего в системе (1) номеру  $n = 1$ , исключим  $u_0$  с помощью равенства  $u_0 = L_{1/2} u_1 + K_{1/2}$ . Результат запишем в разрешенном относительно  $u_1$  виде

$$u_1 = L_{3/2} u_2 + K_{3/2},$$

введя обозначения

$$L_{3/2} = \frac{-c_1}{b_1}, \quad K_{3/2} = \frac{a_1 \varphi - f_1}{-b_1}.$$

Соотношением  $u_1 = L_{3/2} u_2 + K_{3/2}$  можно воспользоваться, чтобы исключить  $u_1$  из уравнения

$$a_2 u_1 + b_2 u_2 + c_2 u_3 = f_2,$$

отвечающего номеру  $n = 2$ . Результат исключения опять запишем в явном относительно  $u_2$  виде

$$u_2 = L_{5/2} u_3 + K_{5/2}.$$

Описанный процесс исключения можно продолжить для  $n = 3, 4, \dots$

Подставляя

$$u_{n-1} = L_{n-1/2}u_n + K_{n-1/2}$$

в уравнение

$$a_n u_{n-1} + b_n u_n + c_n u_{n+1} = f_n,$$

получим

$$u_n = \frac{-c_n}{b_n + a_n L_{n-1/2}} u_{n+1} + \frac{f_n - a_n K_{n-1/2}}{b_n + a_n L_{n-1/2}}.$$

Отсюда видно, что коэффициенты получаемых в процессе исключения соотношений

$$u_n = L_{n+1/2}u_{n+1} + K_{n+1/2}$$

вычисляются по рекуррентным формулам

$$\left. \begin{aligned} L_{n+1/2} &= \frac{-c_n}{b_n + a_n L_{n-1/2}}, \\ K_{n+1/2} &= \frac{f_n - a_n K_{n-1/2}}{b_n + a_n L_{n-1/2}}. \end{aligned} \right\} \quad (2)$$

Последнее из получаемых таким образом соотношений имеет вид

$$u_{N-1} = L_{N-1/2}u_N + K_{N-1/2}.$$

Так как  $u_N = \psi$ , то можно вычислить  $u_{N-1}$ :

$$u_{N-1} = L_{N-1/2}\psi + K_{N-1/2}.$$

После этого  $u_{N-2}$ ,  $u_{N-3}$  и т. д. определяются соответственно из равенств

$$u_{N-2} = L_{N-3/2}u_{N-1} + K_{N-3/2},$$

$$u_{N-3} = L_{N-5/2}u_{N-2} + K_{N-5/2}$$

и т. д., пока не будет определено  $u_1$ .

Повторим кратко, в чем состоит описанный сейчас вычислительный процесс.

Сначала проводится вычисление коэффициентов  $L_{n+1/2}$ ,  $K_{n+1/2}$  в порядке возрастания номеров (прямая прогонка) по рекуррентным формулам (2), причем  $L_{1/2} = 0$  и  $K_{1/2} = \varphi$  заданы.

Затем вычисление неизвестных  $u_n$  производится также рекуррентно в порядке убывания номеров (обратная прогонка) по формулам

$$\left. \begin{aligned} u_N &= \psi, \\ u_n &= L_{n+1/2}u_{n+1} + K_{n+1/2}, \quad n = N-1, N-2, \dots, 1. \end{aligned} \right\} \quad (3)$$

Отметим, что для вычисления методом прогонки решения  $u_0, u_1, \dots, u_N$  системы (1), состоящей из  $N+1$  уравнений, нужно проделать арифметические операции в количестве только в конечное число раз больше, чем число неизвестных. На решение произвольной линейной системы  $N$  уравнений с  $N$  неизвестными методом исключения приходится обычно затрачивать арифметические действия в количестве порядка  $N^3$ . Такого сокращения числа арифметических действий при решении системы (1) методом прогонки удалось достигнуть, удачно используя специфику этой системы.

В § 7 будет показано, что при решении описанным здесь методом прогонки краевой задачи (1), удовлетворяющей одному из указанных в § 4 условий хорошей обусловленности

$$|b_n| \geq |a_n| + |c_n| + \delta, \quad \delta > 0, \quad (4)$$

или

$$\frac{|b_n| - |a_n| - |c_n|}{|b_n| + |a_n| + |c_n|} \geq \theta > 0, \quad d_n = \max(|a_n|, |b_n|, |c_n|) \geq B > 0,$$

или

$$\begin{aligned} \frac{|b_n| - |a_n + c_n|}{|b_n| + |a_n| + |c_n|} &\geq \theta > 0, \quad d_n \geq B > 0, \\ |a_k - a_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, \quad |b_k - b_l| \leq D \left| \frac{k-l}{N} \right|^\omega, \\ |c_k - c_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, \quad D > 0, \quad \omega > 0, \end{aligned}$$

выражения  $b_n + a_n L_{n-1/2}$ , на которые приходится делить, не обращаются в нуль, а погрешности, допускаемые в процессе вычислений, не накапливаются и не приводят к возрастающим с ростом  $N$  ошибкам в вычисляемых значениях решения.

Эти два замечательных свойства прогонки — малое число арифметических действий для ее реализации и слабая чувствительность к вычислительным погрешностям — делают прогонку очень удобным вычислительным алгоритмом.

**2. Пример вычислительно неустойчивого алгоритма.** Для решения хорошо обусловленной разностной краевой задачи (1) возможны разные алгоритмы. Мы описали алгоритм прогонки, обладающий достоинствами малого числа необходимых арифметических действий и вычислительной устойчивости. Укажем другой, еще более простой алгоритм, однако вычислительно неустойчивый и практически непригодный при больших значениях  $N$ .

Задав  $U_0^{(1)} = \varphi$ ,  $U_1^{(1)} = 0$ , найдем решение  $U^{(1)} = \{U_n^{(1)}\}$ ,  $n = 0, 1, \dots, N$ , разностного уравнения (1). Понятно, что, вообще говоря,  $U_N^{(1)} \neq \psi$ . Задав  $U_0^{(2)} = \varphi$ ,  $U_1^{(2)} = 1$ , вычислим решение

$U^{(2)} = \{U_n^{(2)}\}$ . Это решение также не удовлетворяет условию на правой границе. Положим теперь

$$u_n = \sigma U_n^{(1)} + (1 - \sigma) U_n^{(2)}, \quad n = 0, 1, \dots, N. \quad (5)$$

Очевидно, что при любом  $\sigma$  выполнено условие  $u_0 = \varphi$  и удовлетворяется уравнение (1). Выберем  $\sigma$  так, чтобы выполнялось условие

$$u_N = \sigma U_N^{(1)} + (1 - \sigma) U_N^{(2)} = \psi,$$

т. е. положим

$$\sigma = \frac{\psi - U_N^{(2)}}{U_N^{(1)} - U_N^{(2)}},$$

и по формуле (5) получим искомое решение задачи (1).

Если бы вычисления велись на идеальной, лишь умозрительной возможной машине точно, то этот алгоритм был бы хорош. Покажем теперь, что чувствительность его к погрешностям округления для хорошо обусловленной задачи (1) быстро возрастает при  $N \rightarrow \infty$ . Сделаем это на примере, когда  $a_n \equiv 1$ ,  $b_n \equiv -26/5$ ,  $c_n \equiv 1$ ,  $f_n \equiv 0$ .

Условие (4) хорошей обусловленности выполнено. В этом случае точное решение разностной краевой задачи выражается формулой

$$u_n = \frac{5^{N-n} - 5^{n-N}}{5^N - 5^{-N}} \varphi + \frac{5^n - 5^{-n}}{5^N - 5^{-N}} \psi. \quad (6)$$

Для  $U_n^{(1)}$ ,  $U_n^{(2)}$  в силу (5) § 3 получим

$$U_n^{(1)} = -\frac{\varphi}{24} 5^n + \frac{\varphi}{24} 5^{2-n},$$

$$U_n^{(2)} = \frac{5 - \varphi}{24} 5^n + \left[ 5 - \frac{25}{24} (5 - \varphi) \right] 5^{-n}.$$

Заметим, что значения  $\max_n |U_n^{(1)}|$  и  $\max_n |U_n^{(2)}|$  растут, как  $5^N$ .

Поэтому при больших  $N$  при вычислении  $U_n^{(1)}$  и  $U_n^{(2)}$  произойдет выход чисел за допустимые границы. Но допустим, что этого не произошло и что абсолютно точно найдены  $\{U_n^{(1)}\}$ ,  $\{U_n^{(2)}\}$  и  $\sigma$ . Допустим, что единственная ошибка округления  $\varepsilon$  допущена при вычислении  $1 - \sigma$ . Тогда по формуле (5) получим вместо  $\{u_n\}$

$$\{u_n + \Delta u_n\},$$

где  $\Delta u_n = \varepsilon U_n^{(2)}$ .

Погрешность  $\{\Delta u_n\}$  при  $n \sim N$  будет иметь вид

$$\Delta u_n \sim 5^N \varepsilon$$

и при фиксированной относительной погрешности  $\varepsilon$ , допущенной при вычислении  $1 - \sigma$ , будет быстро возрастать, «забывая» точное решение  $\{u_n\}$ , которое в силу формулы (6) остается ограниченным.

Описанный алгоритм называют *методом стрельбы*. В других ситуациях (см. § 20) он может оказаться устойчивым и вполне эффективным.

### ЗАДАЧИ

1. Как надо видоизменить алгоритм прогонки, чтобы воспользоваться им для вычисления решения  $\{u_n\}$ ,  $0 \leq n \leq N$ , разностного уравнения

$$a_n u_{n-1} + b_n u_n + c_n u_{n+1} = f_n, \quad 0 < n < N,$$

при краевых условиях вида

$$u_0 - \alpha u_1 = \varphi, \quad u_N - \beta u_{N-1} = \psi,$$

если числа  $\alpha$  и  $\beta$  отличны от нуля?

2. При вычислении решения задачи

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, & 0 < n < N \\ u_0 &= \varphi, & u_N &= \psi \end{aligned} \right\}$$

можно было бы вести исключение неизвестных  $u_n$  в порядке убывания номеров  $n$

Выписать рекуррентные соотношения для вычисления коэффициентов  $\tilde{K}_{n+1/2}$ ,  $\tilde{K}_{n+1/2}$  получаемых при этом прогоночных соотношений

$$u_{n+1} = \tilde{L}_{n+1/2} u_n + \tilde{K}_{n+1/2}, \quad n = N-1, N-2, \dots, 0.$$

3. Наложив на коэффициенты  $a_n$ ,  $b_n$ ,  $c_n$  разностного уравнения ограничения  $a_n > 0$ ,  $c_n > 0$ ,  $-b_n > a_n + c_n + \delta$ , показать, что прогоночные коэффициенты  $L_{n-1/2}$ , возникающие при решении задачи

$$\left. \begin{aligned} u_0 &= \alpha u_1 + \varphi, & 0 < \alpha < 1, \\ a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, & 0 < n < N, \\ u_{N-1} &= \beta u_N + \psi, \end{aligned} \right\}$$

удовлетворяют неравенствам  $0 \leq L_{n-1/2} \leq 1$ . Как этот факт сказывается на накоплении погрешностей при обратной прогонке? Может ли здесь обратиться в нуль знаменатель прогоночных формул прямой прогонки?

4. Какой вариант прогонки избрать для вычисления решения предыдущей задачи, если  $\alpha = 10$ ,  $\beta = -0,5$ ? Учесть опасность необходимости делить на нуль при вычислении коэффициентов прогоночных соотношений по рекуррентным формулам.

## ОБОСНОВАНИЕ МЕТОДА ПРОГОНКИ \*)

## § 6. Свойства хорошо обусловленных краевых задач

Здесь мы докажем сформулированный в п. 5 § 4 признак хорошей обусловленности разностной краевой задачи вида

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, \quad 0 < n < N, \\ u_0 &= \varphi, \quad u_N = \psi \end{aligned} \right\} \quad (1)$$

и установим некоторые свойства хорошо обусловленных разностных краевых задач с тем, чтобы воспользоваться этими свойствами в § 7 для обоснования алгоритма прогонки.

**1. Оценки решений краевой задачи с возмущенными коэффициентами.** Рассмотрим задачу вида (1):

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, \quad p < n < q, \\ u_p &= \varphi, \quad u_q = \psi, \end{aligned} \right\} \quad (1')$$

где  $p$  и  $q \geq p + 2$  — какие-нибудь целые числа. То обстоятельство, что мы нумеруем компоненты решения номерами от  $p$  до  $q$ , а не от 0 до  $N$ , непринципиально, но окажется удобным в дальнейшем. Относительно коэффициентов будем предполагать, что они ограничены в совокупности:  $|a_n|, |b_n|, |c_n| < M_1$ ,  $M_1$  не зависит от  $N$  и  $n$ .

Пусть задача (1') разрешима при произвольных  $\varphi, \psi$  и  $\{f_n\}$ , причем числа  $u_p, u_{p+1}, \dots, u_q$ , образующие решение, удовлетворяют неравенствам

$$|u_n| \leq M_1 \max_m |f_m| + M_2 \max(|\varphi|, |\psi|), \quad (2)$$

где  $M_1$  и  $M_2$  — некоторые положительные постоянные,  $M_1 \geq M_2$ ,  $M_1 \geq 1$ .

---

\*) Материал гл. 3 в последующих главах не используется и при первом чтении может быть пропущен.



Рассмотрим задачу

$$\left. \begin{aligned} \tilde{a}_n \tilde{u}_{n-1} + \tilde{b}_n \tilde{u}_n + \tilde{c}_n \tilde{u}_{n+1} &= f_n, \quad p < n < q, \\ \tilde{u}_p &= \varphi, \quad \tilde{u}_q = \psi. \end{aligned} \right\} \quad (3)$$

Если предполагать, что возмущения коэффициентов  $\tilde{a}_n - a_n$ ,  $\tilde{b}_n - b_n$ ,  $\tilde{c}_n - c_n$  не слишком сильные, а именно:

$$\left. \begin{aligned} |\tilde{a}_n - a_n| &< \varepsilon < \frac{1}{6M_1}, \\ |\tilde{b}_n - b_n| &< \varepsilon < \frac{1}{6M_1}, \\ |\tilde{c}_n - c_n| &< \varepsilon < \frac{1}{6M_1}, \end{aligned} \right\} \quad (4)$$

то возмущенная система (3) будет обладать следующими четырьмя свойствами:

1° Задача (3) будет иметь решение  $\{\tilde{u}_n\}$  при любых правых частях.

2° Решение  $\{\tilde{u}_n\}$  будет удовлетворять оценке вида (2), но с заменой  $M_1$  и  $M_2$  соответственно на  $2M_1$  и  $2M_2$ :

$$|\tilde{u}_n| \leq 2M_1 \max_m |f_m| + 2M_2 \max(|\varphi|, |\psi|). \quad (5)$$

3° Коэффициенты  $\tilde{a}_n$ ,  $\tilde{b}_n$ ,  $\tilde{c}_n$  будут удовлетворять оценкам

$$|\tilde{a}_n| < M_1 + \frac{1}{6M_1}, \quad |\tilde{b}_n| < M_1 + \frac{1}{6M_1}, \quad |\tilde{c}_n| < M_1 + \frac{1}{6M_1}.$$

4° Решения  $\{u_n\}$  и  $\{\tilde{u}_n\}$  будут мало отличаться друг от друга, а именно:

$$|\tilde{u}_n - u_n| \leq \varepsilon [6M_1^2 \max_m |f_m| + 6M_1 M_2 \max(|\varphi|, |\psi|)]. \quad (6)$$

Свойство 3° очевидно. Докажем свойство 2°, а из него выведем свойство 1°. Предположим, что система (3) разрешима при некоторых правых частях. Фиксировав эти правые части, обозначим

$$\mu = \max_k |\tilde{u}_k|$$

и получим для  $\mu$  неравенство

$$\mu \leq 2M_1 \max_m |f_m| + 2M_2 \max(|\varphi|, |\psi|). \quad (7)$$

Для этого перепишем (3) следующим образом:

$$\left. \begin{aligned} a_n \tilde{u}_{n-1} + b_n \tilde{u}_n + c_n \tilde{u}_{n+1} &= f_n + (a_n - \tilde{a}_n) \tilde{u}_{n-1} + \\ &+ (b_n - \tilde{b}_n) \tilde{u}_n + (c_n - \tilde{c}_n) \tilde{u}_{n+1}, \quad 0 < n < N, \\ \tilde{u}_0 &= \varphi, \quad \tilde{u}_N = \psi. \end{aligned} \right\} \quad (8)$$

Из этой записи и из оценок (2) и (4) вытекает неравенство

$$\begin{aligned} \mu &\leq M_1 \left( \max_m |f_m| + \frac{3}{6M_1} \mu \right) + M_2 \max(|\varphi|, |\psi|) \leq \\ &\leq \frac{1}{2} \mu + M_1 \max_m |f_m| + M_2 \max(|\varphi|, |\psi|). \end{aligned}$$

Решая последнее неравенство относительно  $\mu$ , получим (7), из которого следует (5).

Из неравенства (5) следует, что однородная система, соответствующая задаче (3) и возникающая при  $\varphi = \psi = f_n \equiv 0$ , имеет только нулевое решение  $\tilde{u}_n \equiv 0$ . Поэтому определитель системы (3) отличен от нуля, и задача (3) однозначно разрешима при произвольных правых частях. Свойства 1° и 2° доказаны. Осталось доказать свойство 4°, т. е. неравенство (6).

Вычитая почленно из равенств (8) равенства (1), получим

$$\left. \begin{aligned} a_n(\tilde{u}_{n-1} - u_{n-1}) + b_n(\tilde{u}_n - u_n) + c_n(\tilde{u}_{n+1} - u_{n+1}) = \\ = (a_n - \tilde{a}_n)\tilde{u}_{n-1} + (b_n - \tilde{b}_n)\tilde{u}_n + (c_n - \tilde{c}_n)\tilde{u}_{n+1}, \quad 0 < n < N, \\ \tilde{u}_0 - u_0 = 0, \quad \tilde{u}_N - u_N = 0. \end{aligned} \right\}$$

Применим (2):

$$|\tilde{u}_n - u_n| \leq M_1 \max_m |(\tilde{a}_m - a_m)\tilde{u}_{m-1} + (\tilde{b}_m - b_m)\tilde{u}_m + (\tilde{c}_m - c_m)\tilde{u}_{m+1}|.$$

Воспользовавшись (4) и (5), отсюда выводим

$$|\tilde{u}_n - u_n| \leq M_1 \varepsilon [3 \cdot 2M_1 \max_m |f_m| + 3 \cdot 2M_2 \max(|\varphi|, |\psi|)],$$

т. е. неравенство (6).

Рассмотрим теперь задачу, которая получена из (1') возмущением не только коэффициентов, но и правых частей:

$$\left. \begin{aligned} \tilde{a}_n \tilde{u}_{n-1} + \tilde{b}_n \tilde{u}_n + \tilde{c}_n \tilde{u}_{n+1} = \tilde{f}_n, \quad p < n < q, \\ \tilde{u}_p = \tilde{\varphi}, \quad \tilde{u}_r = \tilde{\psi}. \end{aligned} \right\} \quad (9)$$

Можно показать, что

$$\begin{aligned} |\tilde{u}_n - u_n| &\leq \varepsilon [6M_1^2 \max_m |\tilde{f}_m| + 6M_1 M_2 \max(|\tilde{\varphi}|, |\tilde{\psi}|)] + \\ &+ M_2 \max(|\tilde{\varphi} - \varphi|, |\tilde{\psi} - \psi|) + M_1 \max_m |\tilde{f}_m - f_m|. \quad (10) \end{aligned}$$

Наметим только схему доказательства, которое легко провести по этой схеме.

Изменив сначала только правые части и оставив старые коэффициенты, с помощью (2) увидим, что каждое  $u_n$  изменится не более чем на

$$M_1 \max_m |\tilde{f}_m - f_m| + M_2 \max(|\tilde{\varphi} - \varphi|, |\tilde{\psi} - \psi|).$$

Изменив затем в системе с измененными правыми частями коэффициенты, убедимся, что в силу свойства 4<sup>о</sup> компоненты  $u_n$  дополнительно изменятся на величины, не превосходящие

$$\varepsilon [6M_1^2 \max_m |\tilde{f}_m| + 6M_1M_2 \max(|\tilde{\varphi}|, |\tilde{\psi}|)],$$

что и приведет к оценке (10).

Выведем из описанных нами следствий неравенства (2) еще одно. А именно, пусть для решений системы (1') имеет место при некотором  $\lambda > 0$ ,  $\rho + \lambda < n < q - \lambda$ , оценка

$$|u_n| < M_1 \max_m |f_m| + M_2' \max(|\varphi|, |\psi|).$$

Тогда для решения возмущенной системы

$$\left. \begin{aligned} \tilde{a}_n \tilde{u}_{n-1} + \tilde{b}_n \tilde{u}_n + \tilde{c}_n \tilde{u}_{n+1} &= \tilde{f}_n, & \rho < n < q, \\ \tilde{u}_\rho &= \varphi, & \tilde{u}_q &= \psi, \end{aligned} \right\}$$

удовлетворяющей условиям

$$|\tilde{a}_n - a_n|, |\tilde{b}_n - b_n|, |\tilde{c}_n - c| < \varepsilon < \frac{1}{24M_1^2} < \frac{1}{6M_1}, \quad (11)$$

верно при тех же условиях  $\rho + \lambda < n < q - \lambda$  неравенство

$$|\tilde{u}_n| \leq 2M_1 \max_m |f_m| + \left(M_2' + \frac{1}{4}\right) \max(|\varphi|, |\psi|). \quad (12)$$

Чтобы убедиться в этом, определим вспомогательную сеточную функцию  $\{v_n\}$  как решение системы

$$\left. \begin{aligned} a_n v_{n-1} + b_n v_n + c_n v_{n+1} &= 0, & \rho < n < q, \\ v_\rho &= \varphi, & v_q &= \psi. \end{aligned} \right\}$$

При  $\rho + \lambda < n < q - \lambda$  будет

$$|v_n| < M_2' \max(|\varphi|, |\psi|). \quad (13)$$

Затем применим для оценки  $|\tilde{u}_n - v_n|$  неравенство (10), из которого следует, с учетом (11), что

$$\begin{aligned} |\tilde{u}_n - v_n| &\leq \\ &\leq \varepsilon [6M_1^2 \max_m |f_m| + 6M_1M_2 \max(|\varphi|, |\psi|)] + M_1 \max_m |f_m| \leq \\ &\leq \frac{1}{4} \max(|\varphi|, |\psi|) + 2M_1 \max_m |f_m|. \end{aligned}$$

Принимая во внимание оценку (13), отсюда сразу получаем неравенство (12).

**Замечание.** Важно подчеркнуть, что величина  $\varepsilon$  в оценках (4), в пределах которой можно возмущать коэффициенты исходной задачи, не нарушая разрешимости, а также коэффициенты в оценке (5) решения возмущенной задачи и в оценках (6) и (10) отклонения решения возмущенной задачи от решения невозмущенной задачи — все эти числа зависят только от коэффициентов  $M_1$  и  $M_2$  в оценке (2). Конкретные значения коэффициентов разностного уравнения и число точек  $q - p + 1$  сами по себе роли не играют: их влияние сказывается только через константы  $M_1$  и  $M_2$ , при которых справедлива оценка (2).

**2. Доказательство критерия хорошей обусловленности.** В п. 5 § 4 сформулирован критерий хорошей обусловленности задачи (1) при условиях гладкости коэффициентов

$$\left. \begin{aligned} |a_k - a_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, & |b_k - b_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, \\ |c_k - c_l| &\leq D \left| \frac{k-l}{N} \right|^\omega, & D > 0, \quad \omega > 0, \end{aligned} \right\} \quad (14)$$

и условиях

$$\left. \begin{aligned} d_n = \max(|a_n|, |b_n|, |c_n|) &\geq B > 0, \\ |a_n| &\leq M_1, \quad |b_n| \leq M_1, \quad |c_n| \leq M_1. \end{aligned} \right\} \quad (14')$$

Для хорошей обусловленности задачи (1) при условиях (14), (14') необходимо и достаточно, чтобы корни квадратного уравнения

$$a_n + b_n q + c_n q^2 = 0$$

удовлетворяли неравенствам

$$|q_1| \leq 1 - \frac{\theta}{2}, \quad |q_2^{-1}| \leq 1 - \frac{\theta}{2}. \quad (15)$$

где  $\theta > 0$  не зависит от  $N$  и  $n$ .

Необходимость доказывается примерно таким же способом, как это сделано в п. 4 § 4' при рассмотрении случая постоянных коэффициентов, и мы не будем на этом останавливаться.

При доказательстве достаточности мы будем пользоваться указанным в п. 6 § 4 критерием хорошей обусловленности (15) разностной краевой задачи

$$\left. \begin{aligned} au_{n-1} + bu_n + cu_{n+1} &= f_n, \quad p < n < q, \\ u_p &= \varphi, \quad u_q = \psi \end{aligned} \right\} \quad (16)$$

с постоянными коэффициентами, где  $p$  и  $q$ ,  $q \geq p + 2$ , — произвольные целые числа. В отличие от § 4 мы нумеруем компоненты решения  $\{u_n\}$  не номерами  $n = 0, 1, \dots, N$ , а номерами

$n = p, p + 1, \dots, q$ , что не меняет дела. Задача (16) всегда имеет решение, причем при всех  $n, p \leq n \leq q$ , справедлива оценка (30) § 4:

$$|u_n| \leq M_1 \max_m |f_m| + M_2 \max(|\varphi|, |\psi|), \quad p \leq n \leq q, \quad (17)$$

а при  $n, p + \frac{6}{\theta} < n < q - \frac{6}{\theta}$ , оценка (31) § 4:

$$|u_n| \leq M_1 \max_m |f_m| + M'_2 \max(|\varphi|, |\psi|), \quad (18)$$

где

$$M_1 = \frac{128}{B\theta^3}, \quad M_2 = \frac{4}{\theta}, \quad M'_2 = \frac{1}{5}.$$

Выберем  $\varepsilon$ , положив

$$\varepsilon = \frac{1}{24M_1^2}. \quad (19)$$

Будем считать  $N$  настолько большим, чтобы выполнялось неравенство

$$D \left| \frac{24}{N\theta} \right|^\omega < \varepsilon, \quad \text{т. е.} \quad N > \frac{24}{\theta} \left| \frac{D}{\varepsilon} \right|^{1/\omega}. \quad (20)$$

Переходим к доказательству хорошей обусловленности задачи (1). Рассмотрим краевую задачу вида

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, & p < n < q, \\ u_p &= \varphi, & u_q &= \psi, \end{aligned} \right\} \quad (21)$$

где  $p$  и  $q$  — произвольные фиксированные числа,  $0 \leq p, q \leq N$ ,  $q \geq p + 2$ . В частном случае  $p = 0, q = N$  эта задача совпадает с задачей (1), а вообще получается из задачи (1) некоторым «урезанием» — отбрасыванием уравнений при  $n \leq p$  и  $n \geq q$  и заданием  $u_p$  и  $u_q$ . Мы покажем, что при произвольном  $N$ , удовлетворяющем условию (20), задача (21) однозначно разрешима при произвольных правых частях, причем числа  $\{u_n\}$ ,  $p \leq n \leq q$ , удовлетворяют оценке вида

$$|u_n| \leq M \max(|\varphi|, |\psi|, \max_m |f_m|), \quad (22)$$

где  $M$  — некоторая постоянная, зависящая от  $B, \theta$ , но не от  $N, p, q$ .

Рассмотрим отдельно случай  $q - p \leq 24/\theta$  и случай  $q - p > 24/\theta$ .

Если  $q - p \leq 24/\theta$ , то коэффициенты задачи (21) при любых  $k$  и  $l, p \leq k, l \leq q$ , удовлетворяют в силу условий гладкости

(14) и благодаря тому, что  $N$  в соответствии с (20) достаточно велико, следующим оценкам:

$$|a_k - a_l| \leq D \left| \frac{k-l}{N} \right|^\omega \leq D \left| \frac{q-p}{N} \right|^\omega \leq D \left| \frac{24}{\theta N} \right|^\omega < \varepsilon, \\ |b_k - b_l| < \varepsilon, \quad |c_k - c_l| < \varepsilon.$$

Эти коэффициенты «почти» постоянны и не более чем на  $\varepsilon$  отличаются от коэффициентов задачи (16), где в качестве  $a, b, c$  выбраны  $a_{p+1}, b_{p+1}, c_{p+1}$ . Решение задачи (16) удовлетворяет оценке (17). Число  $\varepsilon$  выбрано по формуле (19) в соответствии с требованием (4). Поэтому для оценки решения задачи (21) можно воспользоваться неравенством (5):

$$|u_n| \leq \frac{256}{B\theta^3} \max_n |f_m| + \frac{8}{\theta} \max(|\varphi|, |\psi|). \quad (23)$$

Рассмотрим теперь случай  $q - p > 24/\theta$ , в частности  $p = 0, q = N$ . Предположим, что при некоторых фиксированных  $\varphi, \psi$  и  $\{f_m\}$  существует решение  $\{u_n\}$ ,  $p \leq n \leq q$ . Выберем последовательность целых чисел  $p = N_0 < N_1 < \dots < N_r = q$  так, чтобы выполнялись неравенства

$$\frac{6}{\theta} < N_{k+1} - N_k < \frac{12}{\theta}. \quad (24)$$

Решение задачи с постоянными коэффициентами

$$\left. \begin{aligned} av_{n-1} + bv_n + cv_{n+1} &= f_n, \quad N_{k-1} < n < N_{k+1}, \\ v_{N_{k-1}} &= \varphi, \quad v_{N_{k+1}} = \psi, \end{aligned} \right\} \quad (25)$$

где

$$a = a_{N_k}, \quad b = b_{N_k}, \quad c = c_{N_k},$$

при  $n = N_k$  в силу неравенств

$$N_{k-1} + \frac{6}{\theta} < N_k < N_{k+1} - \frac{6}{\theta}$$

удовлетворяет оценке (18):

$$|v_{N_k}| \leq M_1 \max_m |f_m| + M_2' \max(|\varphi|, |\psi|),$$

где

$$M_1 = \frac{128}{B\theta^3}, \quad M_2' = \frac{1}{5}.$$

Задача

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, \quad N_{k-1} < n < N_{k+1}, \\ u_{N_{k-1}} &= \varphi_k, \quad u_{N_{k+1}} = \psi_k \end{aligned} \right\} \quad (26)$$

можно рассматривать как возмущение задачи (25), причем коэффициенты задачи (26) в силу неравенства  $N_{k+1} - N_{k-1} \leq 24/\theta$  не более чем на  $\varepsilon$  отличаются от коэффициентов задачи (25). Можно воспользоваться оценкой (12) для решения возмущенной задачи. При  $n = N_k$  получим

$$\begin{aligned} |u_{N_k}| &\leq 2M_1 \max_m |f_m| + \left(M'_2 + \frac{1}{4}\right) \max(|u_{N_{k-1}}|, |u_{N_{k+1}}|) \leq \\ &\leq 2M_1 \max_m |f_m| + \frac{1}{2} \max(|u_{N_{k-1}}|, |u_{N_{k+1}}|). \end{aligned}$$

Следовательно,

$$\begin{aligned} \max_{0 < k < r} |u_{N_k}| &\leq 2M_1 \max_m |f_m| + \frac{1}{2} \max[|\varphi|, |\psi|, \max_{0 < k < r} |u_{N_k}|] \leq \\ &\leq 2M_1 \max_m |f_m| + \frac{1}{2} \max_{0 < k < r} |u_{N_k}| + \frac{1}{2} \max(|\varphi|, |\psi|). \end{aligned}$$

Отсюда

$$\max_{0 < k < r} |u_{N_k}| \leq 4M_1 \max_m |f_m| + \max(|\varphi|, |\psi|).$$

Теперь для произвольного  $n$  найдем  $N_{k-1}$  и  $N_{k+1}$ , между которыми оно заключено, и воспользуемся оценкой (23):

$$\begin{aligned} |u_n| &\leq 2M_1 \max_m |f_m| + 2M_2 \max(|u_{N_{k-1}}|, |u_{N_{k+1}}|) \leq \\ &\leq 2M_1 \max_m |f_m| + 2M_2 [4M_1 \max_m |f_m| + \max(|\varphi|, |\psi|)] \leq \\ &\leq (2M_1 + 8M_1M_2) \max_m |f_m| + 2M_2 \max(|\varphi|, |\psi|). \quad (27) \end{aligned}$$

Оценка (27), полученная при  $q - p > 24/\theta$ , в силу (23) остается справедливой и для  $q - p \leq 24/\theta$ . Задача (21) разрешима при произвольных правых частях, так как из оценки (27) видно, что при  $\varphi = \psi = f_m \equiv 0$  существует только нулевое решение.

Мы завершили доказательство того, что при условиях гладкости (14) и при условиях (14') условие (15) является критерием хорошей обусловленности задачи (1).

Следующий пример показывает, что условия гладкости (14) нельзя игнорировать.

Легко проверить, что разностная краевая задача

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= 0, & 0 < n < N, \\ u_0 &= 0, & u_N &= 0, \end{aligned} \right\}$$

где  $a_n \equiv 1$ ,  $b_n = (-1)^n$ ,  $c_n \equiv 1$  и  $N = 6N_1$ , имеет при любом натуральном  $N_1$  нетривиальное решение

$$u_n = \begin{cases} \sin \frac{n\pi}{6}, & \text{если } n \text{ четное,} \\ -\cos \frac{n\pi}{6}, & \text{если } n \text{ нечетное.} \end{cases}$$

Следовательно, эта краевая задача не является хорошо обусловленной, несмотря на то что

$$\frac{|b_n| - |a_n + c_n|}{|b_n| + |a_n| + |c_n|} = \frac{1}{3}, \quad |a_n| = |b_n| = |c_n| = 1,$$

т. е.

$$|q_1| < 1 - \frac{1}{6}, \quad |q_2^{-1}| < 1 - \frac{1}{6}.$$

**3. Свойства хорошо обусловленных задач.** Сформулируем полученные в § 4 и в п. 2 настоящего параграфа результаты о хорошей обусловленности задачи (1) в форме, удобной для использования при исследовании прогонки в § 7.

Для хорошей обусловленности разностной краевой задачи (1) достаточно, чтобы выполнялся один из следующих трех признаков:

*первый признак:*

$$|b_n| \geq |a_n| + |c_n| + \delta, \quad \delta > 0;$$

*второй признак:*

$$\frac{|b_n| - |a_n| - |c_n|}{|b_n| + |a_n| + |c_n|} \geq \theta > 0, \quad d_n = \max(|a_n|, |b_n|, |c_n|) \geq B > 0;$$

*третий признак:*

$$\frac{|b_n| + |a_n + c_n|}{|b_n| + |a_n| + |c_n|} \geq \theta > 0, \quad M \geq d_n \geq B > 0,$$

причем предполагается, что коэффициенты вещественны и удовлетворяют условиям гладкости (14)

$$|a_k - a_l| \leq D \left| \frac{k-l}{N} \right|^\omega, \quad |b_k - b_l| \leq D \left| \frac{k-l}{N} \right|^\omega,$$

$$|c_k - c_l| \leq D \left| \frac{k-l}{N} \right|^\omega, \quad D > 0, \quad \omega > 0.$$

В случае выполнения любого из первых двух признаков задача (1) разрешима при  $N \geq 2$  и при произвольных правых частях, а в случае выполнения третьего признака задача (1) разрешима при всех достаточно больших  $N$  и произвольных правых частях. При тех же  $N$  наряду с задачей (1) разрешимы все «урезанные» краевые задачи вида (21).



Решение  $\{u_n\}$  исходной задачи и решения  $\{u_n\}$  всех урезанных задач удовлетворяют оценке

$$|u_n| \leq M \max(|\varphi|, |\psi|, \max_m |f_m|), \quad p \leq n \leq q,$$

где  $M$  не зависит от  $N, p, q$ .

### § 7. Обоснование метода прогонки для хорошо обусловленных краевых задач

Теперь все подготовлено для исследования прогонки, которая была описана в § 5. Пусть требуется вычислить решение разностной краевой задачи

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, & 0 < n < N, \\ u_0 &= \varphi, & u_N &= \psi, \\ |a_n|, |b_n|, |c_n| &< M. \end{aligned} \right\} \quad (1)$$

Относительно этой задачи будем предполагать, что сама она и все задачи, полученные ее урезаниями:

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n + c_n u_{n+1} &= f_n, & p < n < q, \\ u_p &= \varphi, & u_q &= \psi, \end{aligned} \right\}$$

имеют решения  $\{u_n\}$  при произвольных правых частях, причем

$$|u_n| \leq M \max(|\varphi|, |\psi|, \max_m |f_m|). \quad (2)$$

В процессе исследования прогонки мы будем пользоваться тем, что в силу оценок (4) и (5) п. 1 § 6 разностная задача с возмущенными коэффициентами

$$\left. \begin{aligned} \tilde{a}_n \tilde{u}_{n-1} + \tilde{b}_n \tilde{u}_n + \tilde{c}_n \tilde{u}_{n+1} &= f_n, & 0 < n < N, \\ \tilde{u}_0 &= \varphi, & \tilde{u}_N &= \psi, \\ |\tilde{a}_n - a_n|, |\tilde{b}_n - b_n|, |\tilde{c}_n - c_n| &\leq \varepsilon < \frac{1}{6M}, \end{aligned} \right\} \quad (3)$$

а также все задачи, полученные урезанием задачи (3), имеют решения  $\{\tilde{u}_n\}$  при произвольных правых частях, причем

$$|\tilde{u}_n| \leq 2M \max(|\varphi|, |\psi|, \max_m |f_m|). \quad (4)$$

**1. Оценки прогоночных коэффициентов.** Здесь мы покажем, что при вычислении прогоночных коэффициентов никогда не придется делить на нуль, и получим оценки прогоночных коэффициентов, пригодные как для исходной задачи (1), так и для возмущенной задачи (3). Для этого достаточно рассматривать возмущенную задачу, так как исходная задача является частным случаем возмущенной (при  $\varepsilon = 0$ ).

Рассмотрим следующую урезанную систему:

$$\left. \begin{aligned} \tilde{a}_n \tilde{u}_{n-1} + \tilde{b}_n \tilde{u}_n + \tilde{c}_n \tilde{u}_{n+1} &= f_n, & 0 < n < l, \\ \tilde{u}_0 &= \varphi, & \tilde{u}_l &= \psi. \end{aligned} \right\}$$

Она разрешима. Найдем из нее  $\tilde{u}_{l-1}$ . Из формул Крамера для решений систем линейных алгебраических уравнений вытекает, что  $\tilde{u}_{l-1}$  представимо в виде

$$\tilde{u}_{l-1} = L\psi + \sum_{i=1}^{l-1} \Lambda_i f_i + \Lambda_0 \varphi = L\tilde{u}_l + K, \quad (5)$$

где  $L$  и  $\Lambda_i$  зависят только от  $\tilde{a}_n, \tilde{b}_n, \tilde{c}_n$ . Вследствие оценки (4), справедливой при произвольных  $\varphi, \psi, \{f_m\}$ , отсюда при  $\varphi = 0, f_i \equiv 0, \psi = 1$  следует

$$|L| = |\tilde{u}_{l-1}| \leq 2M,$$

а при  $\tilde{u}_l = \psi = 0$  следует

$$|K| = |\tilde{u}_{l-1}| \leq 2M \max(|\varphi|, \max_m |f_m|).$$

Величинам  $L$  и  $K$  удобно присвоить индекс  $l - 1/2$  и полученные соотношения и неравенства записывать так:

$$\left. \begin{aligned} \tilde{u}_{l-1} &= L_{l-1/2} \tilde{u}_l + K_{l-1/2}, \\ |L_{l-1/2}| &\leq 2M, \quad |K_{l-1/2}| \leq 2M \max(|\varphi|, \max_m |f_m|). \end{aligned} \right\} \quad (6)$$

Соотношение такого же вида было получено при описании прогонки в § 5. Из формул Крамера (5) видно, что  $L_{l-1/2}$  однозначно определяется через коэффициенты  $\tilde{a}_n, \tilde{b}_n, \tilde{c}_n$ , а  $K_{l-1/2}$  однозначно определяется по  $\varphi, f_n, \tilde{a}_n, \tilde{b}_n, \tilde{c}_n$  ( $0 < n < l$ ). Отсюда следует, что коэффициенты  $L_{l-1/2}, K_{l-1/2}$  совпадают с полученными в § 5 прогоночными коэффициентами, для которых там были выписаны рекуррентные формулы

$$\left. \begin{aligned} L_{l/2} &= 0, & K_{l/2} &= \varphi, \\ \tilde{u}_{l+1/2} &= \frac{-\tilde{c}_l}{\tilde{b}_l + \tilde{a}_l L_{l-1/2}}, & K_{l+1/2} &= \frac{f_l - \tilde{a}_l K_{l-1/2}}{\tilde{b}_l + \tilde{a}_l L_{l-1/2}}. \end{aligned} \right\} \quad (7)$$

Понятно, что это последнее утверждение справедливо, только если рекуррентные формулы имеют смысл, т. е. если ни один из знаменателей в этих формулах не обращается в нуль.

Докажем, что знаменатели действительно не обращаются в нуль.

Пусть мы уже показали, что по формулам (7) можно вычислить

$$\left. \begin{array}{l} L_{1/2}, L_{3/2}, \dots, L_{l-1/2} \\ K_{1/2}, K_{3/2}, \dots, K_{l-1/2}, \end{array} \right\} l \geq 1;$$

проверим их применимость для  $L_{l+1/2}$ ,  $K_{l+1/2}$ . Для этого достаточно показать, что

$$|\tilde{b}_l + \tilde{a}_l L_{l-1/2}| \geq \frac{1}{2M}. \quad (8)$$

Рассмотрим систему уравнений

$$\left. \begin{array}{l} \tilde{u}_0 = 0, \\ \tilde{a}_i \tilde{u}_{i-1} + \tilde{b}_i \tilde{u}_i + \tilde{c}_i \tilde{u}_{i+1} = 0, \quad i = 1, 2, \dots, l-1, \\ \tilde{a}_l \tilde{u}_{l-1} + \tilde{b}_l \tilde{u}_l + \tilde{c}_l \tilde{u}_{l+1} = 1, \\ \tilde{u}_{l+1} = 0. \end{array} \right\} \quad (9)$$

О решении такой системы мы знаем, что оно существует. Из первых  $l$  (однородных) уравнений следует  $\tilde{u}_{l-1} = L_{l-1/2} \tilde{u}_l$ . Из условия (4) следует, что  $|\tilde{u}_l| \leq 2M$ . Из единственного неоднородного уравнения, входящего в систему (9), следует, что

$$(\tilde{a}_l L_{l-1/2} + \tilde{b}_l) \tilde{u}_l = 1.$$

Поэтому

$$\frac{1}{|\tilde{b}_l + \tilde{a}_l L_{l-1/2}|} \leq 2M,$$

что и доказывает оценку (8), а вместе с тем осмысленность рекуррентных формул (7), а также оценки (6).

**2. Оценка влияния на результат ошибок округления в процессе вычислений.** Будем решать задачу (1) прогонкой. При реальных вычислениях на каждом шаге вычислительного процесса допускаются вычислительные погрешности, связанные с ошибками округления. Поэтому реальный вычислительный процесс ведется по формулам

$$\left. \begin{array}{l} L_{1/2} = 0, \quad K_{1/2} = \varphi + \kappa_{1/2}, \\ L_{l+1/2} = \frac{-c_l}{a_l L_{l-1/2} + b_l} + \lambda_{l+1/2}, \quad l = 1, 2, \dots, N-1, \\ K_{l+1/2} = \frac{i_l - a_l K_{l-1/2}}{a_l L_{l-1/2} + b_l} + \kappa_{l+1/2}, \quad l = 1, 2, \dots, N-1, \\ u_N = \psi + v_N, \\ u_l = L_{l+1/2} u_{l+1} + K_{l+1/2} + v_l, \quad l = N-1, N-2, \dots, 1. \end{array} \right\} \quad (10)$$

Предположим, что для всех вычислительных погрешностей справедливы оценки

$$|\kappa_{l+1/2}| < \delta, \quad |\lambda_{l+1/2}| < \delta, \quad |v_l| < \delta$$

с достаточно малым  $\delta$ ,

$$\delta < \frac{1}{6M^2(2M+1)}.$$

Покажем, что в этом случае в прогоночных формулах (10) ни один знаменатель не обращается в нуль, и оценим, насколько допускаемые погрешности могут исказить результат вычислений.

Обозначим

$$K_{1/2} = \tilde{K}_{1/2}; \quad K_{l+1/2} + v_l = \tilde{K}_{l+1/2}, \quad l > 0.$$

Очевидно, что сводка формул (10) может быть переписана так:

$$\left. \begin{aligned} L_{1/2} &= 0, \quad \tilde{K}_{1/2} = \varphi + \kappa_{1/2}, \\ L_{l+1/2} &= -\frac{c_l - (a_l L_{l-1/2} + b_l) \lambda_{l+1/2}}{a_l L_{l-1/2} + b_l}, \quad l = 1, 2, \dots, N-1, \\ \tilde{K}_{l+1/2} &= \frac{f_l - a_l (\tilde{K}_{l-1/2} - v_{l-1})}{a_l L_{l-1/2} + b_l} + \kappa_{l+1/2} + v_l = \\ &= \frac{f_l + a_l v_{l-1} + (a_l L_{l-1/2} + b_l) (\kappa_{l+1/2} + v_l) - a_l \tilde{K}_{l-1/2}}{a_l L_{l-1/2} + b_l}, \\ & \quad l = 1, 2, \dots, N-1, \\ u_N &= \psi + v_N, \\ u_l &= L_{l+1/2} u_{l+1} + \tilde{K}_{l+1/2}, \quad l = N-1, N-2, \dots, 1, \end{aligned} \right\} \quad (10')$$

и рассматриваться как схема вычислительного процесса для решения разностной краевой задачи

$$\left. \begin{aligned} \bar{a}_n \bar{u}_{n-1} + \bar{b}_n \bar{u}_n + \bar{c}_n \bar{u}_{n+1} &= \bar{f}_n, \quad 0 < n < N, \\ \bar{u}_0 &= \varphi, \quad \bar{u}_N = \psi \end{aligned} \right\}$$

со следующими возмущенными правыми частями и коэффициентами:

$$\left. \begin{aligned} \bar{\varphi} &= \varphi + \kappa_{1/2}, \\ \bar{f}_l &= f_l + a_l v_{l-1} + (a_l L_{l-1/2} + b_l) (\kappa_{l+1/2} + v_l), \\ \bar{\psi} &= \psi + v_N, \\ \bar{a}_l &= a_l, \quad \bar{b}_l = b_l, \quad \bar{c}_l = c_l - (a_l L_{l-1/2} + b_l) \lambda_{l+1/2}. \end{aligned} \right\} \quad (11)$$

Докажем, что

$$|\bar{c}_l - c_l| \leq M(2M + 1)\delta < \frac{1}{6M}. \quad (12)$$

Доказательство будем вести индукцией по  $l$ . При  $l = 1$

$$|\bar{c}_1 - c_1| = |(a_1 L_{1/2} + b_1) \lambda_{1/2}| = |(a_1 \cdot 0 + b_1) \lambda_{1/2}| \leq M\delta \leq M(2M + 1)\delta < \frac{1}{6M}.$$

Пусть для  $k = 1, 2, \dots, l-1$  неравенство (12) уже доказано. Для вычисления коэффициентов  $L_{1/2}, L_{1/2}, \dots, L_{l-1/2}$  используются только  $\bar{a}_i = a_i, \bar{b}_i = b_i$  и  $\bar{c}_i$  при  $i = 1, 2, \dots, l-1$ . Поэтому можно утверждать в силу (6), что  $|L_{l-1/2}| \leq 2M$  и что, следовательно,

$$|\bar{c}_l - c_l| = |-(a_l L_{l-1/2} + b_l) \lambda_{l+1/2}| \leq (M \cdot 2M + M)\delta < \frac{1}{6M}.$$

Этим индукция завершается.

Таким образом, показано, что если  $\delta \leq \frac{1}{6M^2(2M+1)}$ , то выполнены неравенства

$$|\bar{a}_n - a_n| = 0 < \frac{1}{6M}, \quad |\bar{b}_n - b_n| = 0 < \frac{1}{6M}, \quad |\bar{c}_n - c_n| < \frac{1}{6M},$$

а значит, справедливы оценки (6) и (8):

$$\left. \begin{aligned} |L_{l-1/2}| < 2M, \\ |\bar{a}_l L_{l-1/2} + \bar{b}_l| = |a_l L_{l-1/2} + b_l| \geq \frac{1}{2M}. \end{aligned} \right\} \quad (13)$$

Мы видим, что в процессе вычислений по формулам (10) не придется делить на нуль.

Теперь из формул (11) для  $\bar{\varphi}, \bar{f}_l, \bar{\psi}$  и оценок (13) следуют неравенства

$$\begin{aligned} |\bar{\varphi} - \varphi| < \delta, \quad |\bar{\psi} - \psi| < \delta, \\ |\bar{f}_l - f_l| \leq M\delta + (M \cdot 2M + M)2\delta = M(4M + 3)\delta. \end{aligned}$$

Таким образом, допуская на каждом шаге вычислительного процесса ошибки, не превосходящие  $\delta, \delta < \frac{1}{6M^2(2M+1)}$ , мы тем самым решаем систему с возмущенными коэффициентами и правыми частями.

Эти возмущения не превосходят  $M^*\delta$ , где

$$M^* = \max \{2, (4M + 3)M\}$$

зависит только от  $M$ , причем возмущения коэффициентов не превосходят также  $1/(6M)$ .

Такие возмущения коэффициентов и правых частей приводят, как показывает оценка (10) из § 6, к погрешностям в  $u_n$ , не превосходящим  $M^{**}\delta$ . Здесь  $M^{**}$  опять-таки зависит только от  $M$ . (Если  $M \approx N^r$ , то  $M^* \approx N^{2r}$ ,  $M^{**} \approx N^{3r}$ , так что погрешность решения будет  $N^{3r}\delta$ .)

Если  $M$ , а тогда и  $M^{**}$ , не зависит от  $N$ , то, совершая при вычислениях по методу прогонки ошибки порядка  $\delta$  на каждом шаге процесса (число таких шагов пропорционально  $N$ ), мы получим в ответе ошибки не больше чем  $\text{const} \cdot \delta$ .

Таким образом, влияние на результат ошибки, допущенной на каком-либо шаге вычислений, не возрастает с ростом  $N$ . Более того, даже суммарное влияние ошибок, допущенных на всех шагах вычислений, тоже не возрастает.

Это замечательное свойство прогонки и послужило причиной ее широкого применения.

## РАЗНОСТНЫЕ СХЕМЫ ДЛЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Часть вторая книги посвящена построению и исследованию разностных схем для обыкновенных дифференциальных уравнений. При этом мы введем основные в теории разностных схем понятия сходимости, аппроксимации и устойчивости, которые носят общий характер. Знакомство с этими понятиями, полученное в связи с обыкновенными дифференциальными уравнениями, позволит в дальнейшем, при изучении разностных схем для уравнений с частными производными, сосредоточиться на многочисленных особенностях и трудностях, характерных для этого очень многообразного класса задач.

### ГЛАВА 4

#### ЭЛЕМЕНТАРНЫЕ ПРИМЕРЫ РАЗНОСТНЫХ СХЕМ

В этой главе мы рассмотрим вводные примеры разностных схем, предназначенные только для предварительного знакомства с основными понятиями теории.

#### § 8. Понятие о порядке точности и об аппроксимации

**1. Порядок точности разностной схемы.** Этот параграф посвящен вопросу сходимости решений разностных уравнений при измельчении сетки к решениям дифференциальных уравнений, которые они приближают. Мы ограничимся здесь исследованием двух разностных схем численного решения задачи

$$\left. \begin{aligned} \frac{du}{dx} + Au = 0, \quad 0 \leq x \leq 1, \\ u(0) = b. \end{aligned} \right\} \quad (1)$$

Начнем с простейшей разностной схемы, основанной на использовании разностного уравнения

$$\frac{u(x+h) - u(x)}{h} + Au(x) = 0. \quad (2)$$

Разобьем отрезок  $[0, 1]$  на шаги длины  $h$ . Удобно выбрать  $h = 1/N$ , где  $N$  — целое число. Точки деления занумеруем слева направо, так что  $x_n = nh$ ,  $n = 0, 1, \dots, N$ . Значение  $u$ , полученное по разностной схеме в точке  $x_n$ , будем обозначать  $u_n$ . Зададим начальное значение. Положим  $u_0 = b$ . Из разностного уравнения (2) вытекает соотношение

$$u_n = (1 - Ah)u_{n-1},$$

откуда находим решение уравнения (2) при начальном условии  $u_0 = b$ :

$$u_n = (1 - Ah)^n b = (1 - Ah)^{x_n/h} b. \quad (3)$$

Точное же решение задачи (1) имеет вид  $u(x) = be^{-Ax}$ . Оно принимает в точке  $x_n$  значение

$$u(x_n) = be^{-Ax_n}. \quad (4)$$

Найдем теперь оценку величины погрешности приближенного решения (3). Эта погрешность в точке  $x_n$  будет

$$\delta(x_n) = [(1 - Ah)^{x_n/h} - e^{-Ax_n}] b. \quad (5)$$

Нас интересует, как убывает  $\delta(x_n)$  при увеличении числа точек разбиения, или, что то же самое, при уменьшении шага  $h = 1/N$  разностной сетки. Для того чтобы выяснить это, представим  $(1 - Ah)^{x_n/h}$  в виде

$$\begin{aligned} (1 - Ah)^{\frac{x_n}{h}} &= e^{\frac{x_n}{h} \ln(1 - Ah)} = e^{\frac{x_n}{h} \left[ -Ah + \frac{A^2 h^2}{2} + O(h^3) \right]} = \\ &= e^{-Ax_n} e^{A^2 \frac{h}{2} x_n} e^{O(h^2)} = e^{-Ax_n} \left[ 1 + \frac{A^2 h x_n}{2} + O(h^2) \right] [1 + O(h^2)] = \\ &= e^{-Ax_n} + h \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2). \end{aligned}$$

Таким образом, равенство (3) примет вид

$$u_n = be^{-Ax_n} + hb \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2), \quad (3')$$

так что

$$\delta(x_n) = hb \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2) = O(h), \quad (6)$$

т. е. погрешность (5) стремится к нулю при  $h \rightarrow 0$  и величина погрешности имеет порядок первой степени шага.

На этом основании говорят, что *разностная схема имеет первый порядок точности* (не путать с порядком разностного уравнения, определенным в § 1).



Решим теперь задачу (1) с помощью разностного уравнения

$$\frac{u(x+h) - u(x-h)}{2h} + Au(x) = 0. \quad (7)$$

Это не так просто, как может показаться на первый взгляд. Дело в том, что рассматриваемая схема является разностным уравнением второго порядка, т. е. требует задания двух начальных условий  $u(x_0) = u(0)$  и  $u(x_1) = u(h)$ , тогда как интегрируемое уравнение (1) есть уравнение первого порядка и для него мы задаем только  $u(0) = b$ . Естественно и в разностной схеме положить  $u_0 = b$ .

Не ясно, как задавать  $u_1$ . Чтобы разобраться в этом, воспользуемся явной формой решения уравнения (7) (см. § 3 формулы (6)):

$$\begin{aligned} u_n &= u_0 \left[ \frac{q_2}{q_2 - q_1} q_1^n - \frac{q_1}{q_2 - q_1} q_2^n \right] + u_1 \left[ -\frac{1}{q_2 - q_1} q_1^n + \frac{1}{q_2 - q_1} q_2^n \right] = \\ &= \frac{q_2 u_0 - u_1}{q_2 - q_1} q_1^n - \frac{q_1 u_0 - u_1}{q_2 - q_1} q_2^n, \end{aligned} \quad (8)$$

где

$$\left. \begin{aligned} q_1 &= \sqrt{1 + A^2 h^2} - Ah = 1 - Ah + \frac{A^2 h^2}{2} + O(h^4), \\ q_2 &= (-1) \left( 1 + Ah + \frac{A^2 h^2}{2} \right) + O(h^4). \end{aligned} \right\} \quad (9)$$

Разложения (9) по формуле Тейлора корней характеристического уравнения позволяют дать приближенные представления для  $q_1^n$  и  $q_2^n$ . Проведем подробно вывод такого представления для  $q_1^n$ :

$$q_1^n = q_1^{x_n/h} = \left[ 1 - Ah + \frac{A^2 h^2}{2} + O(h^4) \right]^{x_n/h} = e^{\frac{x_n}{h} \ln \left[ 1 - Ah + \frac{A^2 h^2}{2} + O(h^4) \right]}$$

Так как  $\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} + O(z^4)$ , то

$$\ln \left[ 1 - Ah + \frac{A^2 h^2}{2} + O(h^4) \right] = -Ah + \frac{A^3 h^3}{6} + O(h^4).$$

Поэтому

$$q_1^n = e^{\frac{x_n}{h} \left[ -Ah + \frac{A^3 h^3}{6} + O(h^4) \right]} = e^{-Ax_n} \left[ 1 + h^2 \frac{A^3 x_n}{6} \right] + O(h^3). \quad (10)$$

Не будем проводить совершенно аналогичной выкладки для  $q_2^n$ , а выпишем сразу результат:

$$q_2^n = (-1)^n e^{Ax_n} + O(h^2). \quad (11)$$

Подставив приближенные выражения для  $q_1^n$  и  $q_2^n$  в формулу (8), получим

$$\begin{aligned} u_n &= \frac{q_2 u_0 - u_1}{q_2 - q_1} q_1^n - \frac{q_1 u_0 - u_1}{q_2 - q_1} q_2^n = \\ &= \frac{q_2 u_0 - u_1}{q_2 - q_1} \left[ e^{-Ax_n} + h^2 \frac{A^3 x_n}{6} e^{-Ax_n} + O(h^3) \right] - \\ &\quad - \frac{q_1 u_0 - u_1}{q_2 - q_1} (-1)^n [e^{Ax_n} + O(h^2)]. \end{aligned} \quad (12)$$

Все дальнейшие выводы мы будем получать путем исследования этой формулы.

Заметим, что если коэффициент  $\frac{q_2 u_0 - u_1}{q_2 - q_1}$  стремится при  $h \rightarrow 0$  к конечному пределу  $b$ , то первое слагаемое правой части равенства (12) стремится к искомому решению задачи (1).

Так как

$$(-1)^n [e^{Ax_n} + O(h^2)] \xrightarrow{h \rightarrow 0} \begin{cases} e^{Ax_n} & \text{при } n \text{ четном,} \\ -e^{Ax_n} & \text{при } n \text{ нечетном,} \end{cases}$$

т. е. не сходится к определенному пределу, то для сходимости к пределу при  $h \rightarrow 0$  второго слагаемого правой части равенства (12):

$$\frac{q_1 u_0 - u_1}{q_2 - q_1} (-1)^n [e^{Ax_n} + O(h^2)], \quad (13)$$

необходимо потребовать, чтобы выражение  $\frac{q_1 u_0 - u_1}{q_2 - q_1}$  стремилось к нулю при  $h \rightarrow 0$ .

Подведем итог всему сказанному.

Для того чтобы решение разностного уравнения

$$\frac{u(x+h) - u(x-h)}{2h} + Au(x) = 0$$

сходилось к решению  $u = be^{-Ax}$  краевой задачи (1), необходимо выполнение условий

$$\frac{q_1 u_0 - u_1}{q_2 - q_1} \rightarrow 0, \quad \frac{q_2 u_0 - u_1}{q_2 - q_1} \rightarrow b. \quad (14)$$

Напомним еще, что  $u_0$  мы условились задавать равным  $b$ . Условия (14) подсказывают нам, как можно задавать  $u_1$ . Оказывается, достаточно, чтобы  $u_1 \rightarrow u_0 = b$  при  $h \rightarrow 0$ . В самом деле,  $q_1 \rightarrow +1$ ,  $q_2 \rightarrow -1$  при  $h \rightarrow 0$ , и поэтому при  $h \rightarrow 0$

$$\frac{q_1 u_0 - u_1}{q_2 - q_1} \rightarrow 0, \quad \frac{q_2 u_0 - u_1}{q_2 - q_1} \rightarrow b.$$

**2. Скорость сходимости решения разностного уравнения.** Теперь перейдем к изучению скорости сходимости при различных конкретных способах выбора  $u_1 \approx u(h)$ .

Для определения  $u(h)$  естественно воспользоваться разложением решения дифференциального уравнения  $u' + Au = 0$  по формуле Тейлора. Пользуясь тем, что в силу этого уравнения  $u' = -Au$ , перепишем формулу Тейлора так:

$$u(x_1) = u(0) - hAu(0) + O(h^2) = u(0)(1 - Ah) + O(h^2).$$

Такое равенство имеет место для точного решения дифференциального уравнения. При приближенном решении, ограничиваясь двумя членами этого разложения, можно положить

$$u_1 = u_0(1 - Ah).$$

Если мы решили ограничиться только одним членом, то полагаем

$$u_1 = u_0.$$

В первом из этих двух случаев мы допускаем в начальном значении  $u_1$  ошибку порядка  $h^2$ , во втором — ошибку порядка  $h$ .

Выясним скорость сходимости в каждом из этих двух случаев задания начальных данных.

Положим

$$u_0 = b, \quad u_1 = (1 - Ah)b. \quad (15)$$

Тогда (см. формулы (9))

$$\frac{q_2 u_0 - u_1}{q_2 - q_1} = \frac{[1 - Ah + O(h^2)]b - (1 - Ah)b}{-2 + O(h^2)} = O(h^2),$$

$$\frac{q_2 u_0 - u_1}{q_2 - q_1} = \frac{\left[-1 - Ah - \frac{A^2 h^2}{2} + O(h^4)\right]b - (1 - Ah)b}{-2 + O(h^4)} = b + O(h^2). \quad (16)$$

Возвращаясь к равенству (12), легко приходим к выводу, который и является нашей целью:

$$u_n = be^{-Ax_n} + O(h^2). \quad (17)$$

Он формулируется так. Если начальное значение  $u_1$  задается с точностью до величины порядка  $h^2$ , то и погрешность в решении будет порядка  $h^2$ , т. е. разностная схема имеет *второй порядок точности*.

Можно показать, что даже если задать в качестве  $u_1$  точное значение  $be^{-Ax_1}$ , большей точности, чем порядка  $h^2$ , в решении добиться нельзя. Советуем читателю в качестве упражнения доказать высказанное утверждение.

Легко также проверить, что если в качестве  $u_0$  задавать не точно  $b$ , а любую величину вида  $b + O(h^2)$ , то скорость сходимости все равно будет второго порядка.

Перейдем к рассмотрению второго изучаемого нами случая задания начальных данных. Полагаем

$$u_1 = u_0 = b.$$

При этом

$$\frac{q_1 u_0 - u_1}{q_2 - q_1} = \frac{[1 - Ah + O(h^2)] b - b}{-2 + O(h^2)} = \frac{1}{2} Ahb + O(h^2),$$

$$\frac{q_2 u_0 - u_1}{q_2 - q_1} = \frac{-[1 + Ah + O(h^2)] b - b}{-2 + O(h^2)} = b + \frac{1}{2} Ahb + O(h^2)$$

и, следовательно,

$$\begin{aligned} u_n &= \frac{q_2 u_0 - u_1}{q_2 - q_1} [e^{-Ax_n} + O(h^2)] - \frac{q_1 u_0 - u_1}{q_2 - q_1} (-1)^n [e^{Ax_n} + O(h^2)] = \\ &= \left[ b + \frac{1}{2} Ahb + O(h^2) \right] [e^{-Ax_n} + O(h^2)] - \\ &- (-1)^n \left[ \frac{1}{2} Ahb + O(h^2) \right] [e^{Ax_n} + O(h^2)] = \\ &= be^{-Ax_n} + Ab \frac{e^{-Ax_n} - (-1)^n e^{Ax_n}}{2} h + O(h^2). \end{aligned}$$

Таким образом, если допустить в начальных данных ошибку порядка  $h$ , то и ошибка в решении будет порядка  $h$ .

Подведем итог. Мы видели, что рассматриваемая разностная схема

$$\frac{u(x+h) - u(x-h)}{2h} + Au(x) = 0,$$

в отличие от схемы

$$\frac{u(x+h) - u(x)}{h} + Au(x) = 0,$$

может дать более высокую скорость сходимости, а именно сходимость с остаточным членом порядка  $h^2$ , а не порядка  $h$ , как у второй из этих схем. Для того чтобы добиться второго порядка точности, надо, задавая точное  $u_0$ , выбирать  $u_1$  отличающимся от значения точного решения дифференциального уравнения в точке  $x = x_0 + h$  на величину порядка  $h^2$ . Можно было бы показать, что и  $u_0$  можно задавать не точно, а с ошибкой порядка  $h^2$ . От этого порядок скорости сходимости не уменьшится. Уточнение начальных данных до порядка  $h^3$  и выше не дает увеличения точности решения.

Если задавать начальные данные с ошибкой порядка  $h$ , то и решение получим с ошибкой того же порядка.

**3. Порядок аппроксимации.** Интересно понять, с чем связано то обстоятельство, что схема

$$\frac{u(x+h) - u(x)}{h} + Au(x) = 0$$

оказывается менее точной, чем схема

$$\frac{u(x+h) - u(x-h)}{2h} + Au(x) = 0.$$

Эти схемы различаются приближенными выражениями

$$\frac{u(x+h) - u(x)}{h} \quad \text{и} \quad \frac{u(x+h) - u(x-h)}{2h}$$

для производной  $du/dx$  в точке  $x$ . Естественно поэтому предполагать, что в первой схеме производная заменена менее точным выражением, чем во второй. Так оно и есть на самом деле. Заменим  $u(x+h)$  и  $u(x-h)$  их тейлоровскими разложениями:

$$u(x+h) = u(x) + u'(x)h + u''(x)\frac{h^2}{2} + u'''(x)\frac{h^3}{6} + O(h^4),$$

$$u(x-h) = u(x) - u'(x)h + u''(x)\frac{h^2}{2} - u'''(x)\frac{h^3}{6} + O(h^4).$$

Пользуясь этими разложениями, получим

$$\frac{u(x+h) - u(x)}{h} = u'(x) + u''(x)\frac{h}{2} + O(h^2),$$

$$\frac{u(x+h) - u(x-h)}{2h} = u'(x) + u'''(x)\frac{h^2}{6} + O(h^4),$$

т. е. в первом случае мы имеем аппроксимацию производной лишь с первым порядком точности, а во втором — со вторым порядком.

Рассмотренные примеры наводят на мысль, что порядок скорости сходимости решений разностных уравнений может быть сделан равным порядку аппроксимации производных дифференциального уравнения.

Однако оказывается, что в такой общей формулировке эта гипотеза неверна. На разностные схемы, для которых будет доказана ее справедливость, нам придется наложить одно весьма существенное ограничение — требование *устойчивости*. Необходимость этого ограничения станет ясна из примера, который мы рассмотрим в следующем параграфе.

## § 9. Неустойчивая разностная схема

**1. Способы аппроксимации производной.** Займемся снова разностными схемами для приближенного интегрирования простейшего дифференциального уравнения  $u' + Au = 0$ . Как мы уже видели, для составления разностной схемы, приближающей это уравнение, достаточно заменить производную  $u'$  каким-либо аппроксимирующим ее разностным отношением. Так, например, мы рассматривали схемы, для которых производная  $u'$  заменялась через

$$\frac{u(x+h) - u(x)}{h} \quad \text{или} \quad \frac{u(x+h) - u(x-h)}{2h}.$$

Очевидно также, что любое выражение вида

$$\mu \frac{u(x+h) - u(x-h)}{2h} + (1-\mu) \frac{u(x+h) - u(x)}{h}$$

будет приближать  $u'(x)$ . В самом деле, подставим в это выражение тейлоровские разложения для  $u(x+h)$  и  $u(x-h)$ :

$$u(x+h) = u(x) + u'(x)h + O(h^2),$$

$$u(x-h) = u(x) - u'(x)h + O(h^2).$$

Тогда получим

$$\begin{aligned} \mu \frac{u(x+h) - u(x-h)}{2h} + (1-\mu) \frac{u(x+h) - u(x)}{h} &= \\ &= \mu \frac{[u(x) + u'(x)h + O(h^2)] - [u(x) - u'(x)h + O(h^2)]}{2h} + \\ &+ (1-\mu) \frac{[u(x) + u'(x)h + O(h^2)] - u(x)}{h} = u'(x) + O(h). \end{aligned}$$

Пользуясь такого рода аппроксимацией производной, можно получить целое семейство разностных схем, зависящих от числового параметра.

Эти схемы будут иметь вид

$$\mu \frac{u(x+h) - u(x-h)}{2h} + (1-\mu) \frac{u(x+h) - u(x)}{h} + Au(x) = 0. \quad (1)$$

Каждому значению параметра  $\mu$  отвечает своя схема. Изучению схем, получающихся при  $\mu = 0$  и  $\mu = 1$ , был посвящен § 8.

**2. Пример неустойчивой разностной схемы.** Рассмотрим теперь еще одну схему такого вида, которая получается из (1) при  $\mu = 4$ :

$$4 \frac{u(x+h) - u(x-h)}{2h} - 3 \frac{u(x+h) - u(x)}{h} + Au(x) = 0. \quad (2)$$

Ее можно переписать еще так:

$$-2u(x-h) + (3 + Ah)u(x) - u(x+h) = 0. \quad (2')$$

Как и в ранее рассмотренных примерах, мы будем получать решение на отрезке  $[0, 1]$ , разбитом точками разностной сетки на  $N$  равных шагов, каждый длины  $h = 1/N$ . Координата  $x_n$  точки сетки определяется как  $x_n = nh = n/N$ .

Решение разностного уравнения выписывается явно формулой

$$u_n = u_0 \left[ \frac{q_2}{q_2 - q_1} q_1^n - \frac{q_1}{q_2 - q_1} q_2^n \right] + u_1 \left[ -\frac{1}{q_2 - q_1} q_1^n + \frac{1}{q_2 - q_1} q_2^n \right], \quad (3)$$

где  $q_1$  и  $q_2$  — корни характеристического уравнения

$$-2 + (3 + Ah)q - q^2 = 0.$$

Вычислим  $q_1$  и  $q_2$ :

$$\left. \begin{aligned} q_1 &= \frac{3 + Ah - \sqrt{1 + 6Ah + A^2h^2}}{2} = 1 - Ah + 2A^2h^2 + O(h^3), \\ q_2 &= \frac{3 + Ah + \sqrt{1 + 6Ah + A^2h^2}}{2} = 2(1 + Ah) + O(h^2). \end{aligned} \right\} \quad (4)$$

Мы будем пользоваться еще приближенными выражениями для  $q_1^n$  и  $q_2^n$ :

$$\left. \begin{aligned} q_1^n &= [1 - Ah + O(h^2)]^n = [1 - Ah + O(h^2)]^{x_n/h} = \\ &= e^{-Ax_n} + O(h), \\ q_2^n &= [2(1 + Ah) + O(h^2)]^n = [2(1 + Ah) + O(h^2)]^{x_n/h} = \\ &= 2^{x_n/h} [e^{Ax_n} + O(h)]. \end{aligned} \right\} \quad (5)$$

Подставив выражения (5) в формулу (3), получим

$$u_n = \frac{q_2 u_0 - u_1}{q_2 - q_1} [e^{-Ax_n} + O(h)] + \frac{q_1 u_0 - u_1}{q_1 - q_2} [e^{Ax_n} + O(h)] 2^{x_n/h}. \quad (6)$$

Прежде чем исследовать, к чему стремится  $u_n$  при  $h \rightarrow 0$ , мы должны указать, как задаются начальные значения  $u_0$  и  $u_1$  разностного решения.

Так же, как и в § 8, будем разыскивать решение, удовлетворяющее условию  $u(0) = b$ , и возьмем в качестве разностных начальных данных  $u_0 = b$  и  $u_1 = b(1 - Ah)$ . Подставим эти начальные данные в формулу (6) и упростим по отдельности каждое слагаемое.

Первое и второе слагаемые примут соответственно вид

$$\begin{aligned} \frac{q_2 u_0 - u_1}{q_2 - q_1} [e^{-Ax_n} + O(h)] &= \\ &= \frac{[2 + O(h)]b - (1 - Ah)b}{[2 + O(h)] - [1 - O(h)]} [e^{-Ax_n} + O(h)] = be^{-Ax_n} + O(h), \\ \frac{q_1 u_0 - u_1}{q_2 - q_1} [e^{Ax_n} + O(h)] 2^{x_n/h} &= \\ &= \frac{[1 - Ah + 2A^2 h^2 + O(h^3)]b - b(1 - Ah)}{[1 + O(h)] - [2 + O(h)]} [e^{Ax_n} + O(h)] 2^{x_n/h} = \\ &= -2A^2 h^2 b [e^{Ax_n} + O(h)] 2^{x_n/h}. \end{aligned}$$

Таким образом, мы получили

$$u_n = [be^{-Ax_n} + O(h)] + [-2A^2 be^{Ax_n} + O(h)] h^2 2^{x_n/h}.$$

Первое слагаемое этой формулы при  $h \rightarrow 0$ ,  $x_n = x = \text{const}$  стремится к  $be^{-Ax}$ , т. е. к искомого решению. Значит, для того чтобы к этому решению сходилось все выражение для  $u_n$ , необходимо, чтобы второе слагаемое сходилось к нулю. Однако оно при  $h \rightarrow 0$  стремится не к нулю, а к бесконечности. В самом деле,  $-2A^2 be^{Ax_n} + O(h)$  стремится к конечному и не равному нулю пределу  $-2A^2 be^{Ax}$ , а  $h^2 2^{x_n/h}$  стремится к бесконечности быстрее любой положительной степени  $1/h$ .

Мы показали, что разностная схема, аппроксимирующая дифференциальное уравнение, может иметь решение, не сходящееся при  $h \rightarrow 0$  к решению дифференциального уравнения. Можно подумать, что причина этого в недостаточно точном выборе  $u_1$ . Однако мы сейчас покажем, что сходимости не будет, даже если выбрать  $u_1$  точно равным решению дифференциального уравнения при  $x_1 = x_0 + h$ , т. е. если положить  $u_1 = u_0 e^{-Ah} = be^{-Ah}$ . Начнем с того, что упростим выражения, входящие в формулу (6):

$$\begin{aligned} \frac{q_2 u_0 - u_1}{q_2 - q_1} &= \frac{[2 + O(h)]b - be^{-Ah}}{[2 + O(h)] - [1 + O(h)]} = b + O(h), \\ \frac{q_1 u_0 - u_1}{q_1 - q_2} &= \frac{[1 - Ah + 2A^2 h^2 + O(h^3)]b - be^{-Ah}}{[1 + O(h)] - [2 + O(h)]} = -\frac{3}{2} A^2 h^2 [b + O(h)]. \end{aligned}$$

Подставив эти выражения в формулу (6), получим

$$u_n = [be^{-Ax_n} + O(h)] - \left[ \frac{3}{2} A^2 be^{Ax_n} + O(h) \right] h^2 2^{x_n/h}. \quad (7)$$



Второй член правой части этого равенства снова стремится к бесконечности, тогда как первый остается ограниченным. Поэтому стремится к бесконечности и все решение разностного уравнения.

Причина того, что разностная схема (2) не дает сходимости при  $h \rightarrow 0$ , как мы видели, состоит в том, что она может иметь быстро возрастающие при уменьшении шага  $h$  решения, даже если начальные данные заданы вполне разумно.

Такого рода разностные схемы называются *неустойчивыми*. Естественно, что они непригодны для численного решения дифференциальных уравнений.

## СХОДИМОСТЬ РЕШЕНИЯ РАЗНОСТНЫХ УРАВНЕНИЙ КАК СЛЕДСТВИЕ АППРОКСИМАЦИИ И УСТОЙЧИВОСТИ

В гл. 4 мы на примерах выяснили, что такое аппроксимация дифференциальной задачи разностной задачей и в чем состоит сходимость, благодаря которой решение дифференциальной задачи можно приближенно вычислять по разностной схеме. Мы познакомились с явлением неустойчивости, которое может сделать разностную схему расходящейся и непригодной для вычислений. Анализ поведения решений в этих элементарных вводных примерах, предназначенных только для предварительного знакомства с основными понятиями, был основан на записи решений в виде формул. Такая запись оказалась возможной лишь благодаря специальному подбору примеров.

В этой главе мы дадим строгие определения понятий сходимости, аппроксимации и устойчивости. Мы покажем, что доказательство сходимости не обязательно основывать на анализе формул для решений. Это доказательство можно разбить на проверку аппроксимации дифференциальной задачи разностной и проверку устойчивости разностной задачи.

### § 10. Сходимость разностной схемы

**1. Понятие о сетке и сеточной функции.** Пусть на некотором отрезке  $D$  поставлена некоторая дифференциальная краевая задача. Это значит, что задано дифференциальное уравнение (или система), которому должно удовлетворять решение  $u$  на отрезке  $D$  и дополнительные условия для  $u$  на одном или на обоих концах отрезка. Дифференциальную краевую задачу будем записывать в виде символического равенства

$$Lu = f, \quad (1)$$

где  $L$  — заданный дифференциальный оператор, а  $f$  — заданная правая часть. Так, например, чтобы записать в виде (1) задачу

$$\left. \begin{aligned} \frac{du}{dx} + \frac{x}{1+u^2} &= \cos x, & 0 \leq x \leq 1, \\ u(0) &= 3, \end{aligned} \right\} \quad (2)$$

ДОСТАТОЧНО ПОЛОЖИТЬ

$$Lu \equiv \begin{cases} \frac{du}{dx} + \frac{x}{1+u^2}, & 0 \leq x \leq 1, \\ u(0), & \end{cases}$$

$$f \equiv \begin{cases} \cos x, & 0 \leq x \leq 1, \\ 3. & \end{cases}$$

Задача

$$\left. \begin{aligned} \frac{d^2u}{dx^2} - (1+x^2)u &= \sqrt{x}, & 0 \leq x \leq 1, \\ u(0) &= 2, \\ \frac{du(0)}{dx} &= 1 \end{aligned} \right\} \quad (3)$$

запишется в виде (1), если положить

$$Lu \equiv \begin{cases} \frac{d^2u}{dx^2} - (1+x^2)u, & 0 \leq x \leq 1, \\ u(0), \\ \frac{du(0)}{dx}, \end{cases}$$

$$f \equiv \begin{cases} \sqrt{x}, & 0 \leq x \leq 1, \\ 2, \\ 1. \end{cases}$$

Для записи в виде (1) задачи

$$\left. \begin{aligned} \frac{d^2u}{dx^2} - (1+x^2)u &= \sqrt{x+1}, & 0 \leq x \leq 1, \\ u(0) &= 2, \\ u(1) &= 1 \end{aligned} \right\} \quad (4)$$

с краевыми условиями на обоих концах отрезка  $0 \leq x \leq 1$  надо положить

$$Lu \equiv \begin{cases} \frac{d^2u}{dx^2} - (1+x^2)u, & 0 \leq x \leq 1, \\ u(0), \\ u(1), \end{cases}$$

$$f \equiv \begin{cases} \sqrt{x+1}, & 0 \leq x \leq 1, \\ 2, \\ 1. \end{cases}$$

Краевая задача для системы дифференциальных уравнений

$$\left. \begin{aligned} \frac{dv}{dx} + xvw &= x^2 - 3x + 1, & 0 \leq x \leq 1, \\ \frac{dw}{dx} + \frac{1}{1+x^2}(v+w) &= \cos^2 x, & 0 \leq x \leq 1, \\ v(0) &= 1, \\ w(0) &= -3 \end{aligned} \right\} \quad (5)$$

будет записана в форме (1), если считать  $u$  вектор-функцией  $u = \begin{pmatrix} v \\ w \end{pmatrix}$  и положить

$$Lu \equiv \begin{cases} \frac{dv}{dx} + xvw, & 0 \leq x \leq 1, \\ \frac{dw}{dx} + \frac{1}{1+x^2}(v+w), & 0 \leq x \leq 1, \\ v(0), \\ w(0), \end{cases}$$

$$f \equiv \begin{cases} x^2 - 3x + 1, & 0 \leq x \leq 1, \\ \cos^2 x, & 0 \leq x \leq 1, \\ 1, \\ -3. \end{cases}$$

Во всех примерах мы рассматриваем задачу на отрезке  $0 \leq x \leq 1$ , а не на каком-либо другом, только для определенности.

Будем предполагать, что решение  $u(x)$  задачи (1) на отрезке  $0 \leq x \leq 1$  существует. Для вычисления этого решения с помощью метода конечных разностей, или метода сеток, надо прежде всего выбрать на отрезке  $D$  конечное число точек, совокупность которых будем называть *сеткой* и обозначать через  $D_h$ , а затем считать искомым не решение  $u(x)$  задачи (1), а таблицу  $[u]_h$  значений этого решения в точках сетки  $D_h$ . Предполагается, что сетка  $D_h$  зависит от параметра  $h > 0$ , который может принимать сколь угодно малые положительные значения. При стремлении «шага сетки»  $h$  к нулю сетка должна становиться все «гуще». Например, можно положить  $h = 1/N$ , где  $N$  — какое-нибудь натуральное число, и принять за сетку  $D_h$  совокупность точек  $x_0 = 0$ ,  $x_1 = h$ ,  $x_2 = 2h$ , ...,  $x_N = 1$ . Искомая сеточная функция  $[u]_h$  в этом случае в точке  $x_n = nh$  сетки  $D_h$  принимает значение  $u(nh)$ , которое для краткости будем обозначать  $u_n$ .

Для приближенного вычисления таблицы значений решения  $[u]_h$  в случае задачи (2) можно воспользоваться, например, системой уравнений

$$\left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + \frac{x_n}{1 + u_n^2} &= \cos x_n, & n = 0, 1, \dots, N-1, \\ u_0 &= 3, \end{aligned} \right\} \quad (6)$$

полученной в результате замены производной  $du/dx$  в точках сетки разностным отношением по приближенной формуле

$$\frac{du}{dx} \approx \frac{u(x+h) - u(x)}{h}.$$

Решение  $u^{(h)} = (u_0^{(h)}, u_1^{(h)}, \dots, u_N^{(h)})$  системы (6) определено на той же сетке  $D_h$ , что и искомая сеточная функция  $[u]_h$ . Его значения  $u_1^{(h)}, u_2^{(h)}, \dots, u_N^{(h)}$  в точках  $x_1, x_2, \dots, x_N$  последовательно вычисляются из (6) при  $n = 0, 1, \dots, N-1$ . Для краткости в уравнении (6) мы не пишем значок  $h$  при  $u_n^{(h)}$ . Как правило, мы будем так же поступать в аналогичных случаях и в дальнейшем.

В случае задачи (4) для отыскания сеточной функции  $u^{(h)}$ , приближенно совпадающей с искомой таблицей решения  $[u]_h$ , можно воспользоваться разностной схемой

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} - (1 + x_n^2)u_n &= \sqrt{x_n + 1}, \\ n = 1, 2, \dots, N-1, \\ u_0 = 2, \quad u_N = 1. \end{aligned} \right\} \quad (7)$$

Эта схема возникает в результате замены в точках сетки производной  $d^2u/dx^2$ , входящей в дифференциальное уравнение, по приближенной формуле

$$\frac{d^2u}{dx^2} \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}. \quad (8)$$

Для вычисления решения  $u^{(h)}$  задачи (7) можно воспользоваться алгоритмом исключения — прогонки, описанным в § 5.

Выпишем еще разностную схему, пригодную для вычисления решения задачи (5):

$$\left. \begin{aligned} \frac{v_{n+1} - v_n}{h} + x_n v_n \omega_n &= x_n^2 - 3x_n + 1, \\ \frac{\omega_{n+1} - \omega_n}{h} + \frac{1}{1 + x_n^2} (v_n + \omega_n) &= \cos^2 x_n, & n = 0, 1, \dots, N-1, \\ v_0 &= 1, \\ \omega_0 &= -3. \end{aligned} \right\} \quad (9)$$

Здесь  $u_0^{(h)} = \begin{pmatrix} v_0^{(h)} \\ w_j^{(h)} \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$  задано. При  $n = 0$  из уравнений (9) можно найти  $u_1^{(h)} = \begin{pmatrix} v_1^{(h)} \\ w_1^{(h)} \end{pmatrix}$ . Вообще, зная  $u_k^{(h)} = \begin{pmatrix} v_k^{(h)} \\ w_k^{(h)} \end{pmatrix}$ ,  $k = 0, 1, \dots, n$ , можно при  $k = n$  вычислить  $u_{n+1}^{(h)} = \begin{pmatrix} v_{n+1}^{(h)} \\ w_{n+1}^{(h)} \end{pmatrix}$ .

В рассмотренных примерах сетка  $D_h$  состоит из удаленных друг от друга на расстояние  $h$  точек. Ясно, что можно было бы расположить  $N + 1$  точек сетки  $D_h$ ,  $h \equiv 1/N$ , на отрезке  $[0, 1]$  не равномерно, а так, чтобы  $x_0 = 0$ ,  $x_1 = x_0 + h_0$ ,  $x_2 = x_1 + h_1, \dots, x_{n+1} = x_n + h_n$ ,  $x_N = 1$ , где  $h_n$ ,  $n = 0, 1, \dots, N - 1$ , — не равные между собой числа, однако такие, что  $\max_n h_n \rightarrow 0$  при  $h = 1/N \rightarrow 0$ . Выбором расположения узлов сетки  $D_h$  можно добиться того, чтобы искомая таблица  $[u]_h$  решения  $u(x)$  была подробнее при данном фиксированном  $N$  (или  $h = 1/N$ ) на тех участках, где  $u(x)$  более быстро изменяется. Такие участки иногда бывают заранее известны из физики или из предварительных грубых расчетов. Информация о скорости изменения  $u(x)$  выявляется также в ходе последовательного вычисления  $u_1^{(h)}, u_2^{(h)}, \dots, u_n^{(h)}$  и может быть учтена при выборе следующего узла сетки  $x_{n+1}$ .

Мы ограничимся приведенными примерами для иллюстрации понятия сетки и искомой сеточной функции (или вектор-функции) — таблицы значений решения  $[u]_h$ . Заметим только, что в качестве искомой таблицы  $[u]_h$  значений решения не обязательно рассматривать сеточную функцию, совпадающую с решением  $u$  в точках сетки. Возможны и другие способы установления соответствия между функцией и ее таблицей. Например, таблицей  $u(x)$ ,  $0 < x < 1$ , можно считать сеточную функцию  $[u]_h$ , определенную в точках  $x = \frac{h}{2}, \frac{3}{2}h, \dots, 1 - \frac{h}{2}$  равенством

$$[u]_h = \frac{1}{h} \int_{x-h/2}^{x+h/2} u(\xi) d\xi.$$

Такой способ установления соответствия удобен в случае, когда  $u(x)$  не является непрерывной функцией, но известно, что интеграл от нее по любому отрезку существует. Это может быть, например, при рассмотрении обобщенных — разрывных — решений, если существует

$$\int_a^1 u^2(x) dx.$$

Всюду в дальнейшем, если не оговорено противное, мы будем считать, что  $u$  — непрерывная функция, а под  $[u]_h$  понимать сеточную функцию, совпадающую с  $u$  в точках сетки.

Мы ставим вопрос о вычислении сеточной функции  $[u]_h$  потому, что при измельчении сетки, т. е. при  $h \rightarrow 0$ , она является все более подробной таблицей искомого решения  $u$  и дает о нем все более полное представление. Пользуясь интерполяцией, можно было бы с возрастающей при  $h \rightarrow 0$  точностью восстановить решение  $u$  всюду в области  $D$ . Ясно, что точность, с которой это можно сделать при заданном фиксированном числе и расположении узлов сетки  $D_h$ , зависит от дополнительно известных сведений о решении типа оценок для его производных, а также от расположения узлов сетки  $D_h$ .

Ограничимся этими беглыми замечаниями о восстановлении функции  $u$  по ее таблице  $[u]_h$ . Подробное рассмотрение вопросов восстановления функции по ее таблице составляет предмет теории интерполяции. Мы будем заниматься только задачей вычисления таблицы  $[u]_h$ . Поэтому *условимся считать*, что задача (1) решена точно, если найдена сеточная функция  $[u]_h$ . Однако нам не удастся вычислять ее точно. Вместо сеточной функции  $[u]_h$  будем искать другую сеточную функцию  $u^{(h)}$ , которая «сходится» к  $[u]_h$  при измельчении сетки. Для этой цели можно использовать разностные уравнения.

**2. Сходящиеся разностные схемы.** Способами построения и исследования сходящихся разностных схем мы будем заниматься на протяжении всей главы. Однако прежде надо придать точный смысл самому требованию сходимости  $u^{(h)} \rightarrow [u]_h$ , которое мы будем предъявлять к разностным схемам. Для этого рассмотрим линейное нормированное пространство функций, определенных на сетке  $D_h$ . Норма  $\|u_h\|_{U_h}$  сеточной функции  $u_h \in U_h$  есть неотрицательное число, которое принимается за меру отклонения функции  $u_h$  от тождественного нуля. Напомним, что линейное пространство  $R$  называется *нормированным*, если каждому элементу  $x$  этого пространства поставлено в соответствие неотрицательное число  $\|x\|$ , причем выполнены следующие три аксиомы нормы:

- 1°  $\|x\| \geq 0, \quad x \in R;$
- 2°  $\|\lambda x\| = |\lambda| \cdot \|x\|$ , где  $x \in R$ , а  $\lambda$  — произвольное число;
- 3°  $\|x + y\| \leq \|x\| + \|y\|$ , где  $x, y \in R$ .

Норма может быть определена различными способами. Можно, например, принять за норму функции точную верхнюю грань модуля ее значений в точках сетки, положив

$$\|u_h\|_{U_h} = \sup_n |u_h(x_n)|. \quad (10)$$

Если  $u^{(h)}$  — пара функций, как в (9), то за норму, аналогичную (10), можно принять верхнюю грань модулей обеих функций на соответствующих им сетках.

Если  $u^{(h)}$  состоит из функций, определенных на сетке  $x = 0, h, 2h, \dots, 1$ , то часто используют норму, определенную равенством

$$\|u^{(h)}\|_{U_h} = \left( h \sum_{n=0}^N |u_n|^2 \right)^{1/2}.$$

Эта норма аналогична норме

$$\|u(x)\| = \left( \int_0^1 |u(x)|^2 dx \right)^{1/2}$$

для функций  $u(x)$  с интегрируемым на отрезке  $0 \leq x \leq 1$  квадратом.

Всюду, где не оговорено противное, мы будем пользоваться нормой (10).

После того, как введено нормированное пространство  $U_h$ , приобретает смысл понятие отклонения одной функции от другой. Если  $a^{(h)}$  и  $b^{(h)}$  — две произвольные сеточные функции из  $U_h$ , то мерой их отклонения друг от друга считается норма их разности, т. е. число

$$\|a^{(h)} - b^{(h)}\|_{U_h}.$$

Теперь можно перейти к строгому определению сходящейся разностной схемы.

Пусть для приближенного вычисления решения дифференциальной краевой задачи (1), т. е. для приближенного вычисления сеточной функции  $[u]_h$  на основе использования равенства (1), составлена некоторая система уравнений, которую будем символически записывать, аналогично уравнению (1), в форме равенства

$$L_h u^{(h)} = f^{(h)}. \quad (11)$$

Примерами могут служить разностные схемы (6), (7), (9) для дифференциальных краевых задач (2), (4), (5) соответственно.

Для записи схемы (6) в форме (11) можно положить

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - u_n}{h} + \frac{nh}{1 + u_n^2}, & n = 0, 1, \dots, N-1, \\ u_0, & \end{cases}$$

$$f^{(h)} \equiv \begin{cases} \cos nh, & n = 0, 1, \dots, N-1, \\ 3, & \end{cases}$$



Схема (7) запишется в форме (11), если принять

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + [1 - (nh)^2] u_n, & n = 1, 2, \dots, N-1, \\ u_0, \\ u_N, \end{cases}$$

$$f^{(h)} \equiv \begin{cases} \sqrt{1 + nh}, & n = 1, 2, \dots, N-1, \\ 2, \\ 1. \end{cases}$$

Запишем еще в виде (11) схему (9), приняв

$$L_h u^{(h)} \equiv L_h \begin{pmatrix} v^{(h)} \\ w^{(h)} \end{pmatrix} =$$

$$= \begin{cases} \frac{v_{n+1} - v_n}{h} + (nh) v_n w_n, & n = 0, 1, \dots, N-1, \\ \frac{w_{n+1} - w_n}{h} + \frac{1}{1 + (nh)^2} (v_n + w_n), & n = 0, 1, \dots, N-1, \\ v_0, \\ w_0, \end{cases}$$

$$f^{(h)} = \begin{cases} (nh)^2 - 3nh + 1, & n = 0, 1, \dots, N-1, \\ \cos^2 nh, & n = 0, 1, \dots, N-1, \\ 1, \\ -3. \end{cases}$$

Система (11), как видим, зависит от  $h$  и должна быть выписана для всех тех  $h$ , для которых рассматривается сетка  $D_h$  и сеточная функция  $[u]_h$ . Таким образом, разностная краевая задача (11) — это не одна система, а семейство систем, зависящее от параметра  $h$ .

Будем предполагать, что при каждом рассматриваемом достаточно малом  $h$  существует решение  $u^{(h)}$  задачи (11), принадлежащее пространству  $U_h$ .

Будем говорить, что решение  $u^{(h)}$  разностной краевой задачи (11) при измельчении сетки сходится к решению и дифференциальной краевой задачи (1), если

$$\|[u]_h - u^{(h)}\|_{U_h} \rightarrow 0 \quad \text{при} \quad h \rightarrow 0. \quad (12)$$

Если, сверх того, выполнено неравенство

$$\|[u]_h - u^{(h)}\|_{U_h} \leq ch^k, \quad (13)$$

где  $c > 0$  и  $k > 0$  — некоторые постоянные, не зависящие от  $h$ , то будем говорить, что имеет место *сходимость порядка  $h^k$*  или что *разностная схема имеет  $k$ -й порядок точности*.

В § 8 были рассмотрены две разностные схемы для задачи

$$\left. \begin{aligned} \frac{du}{dx} + Au = 0, \quad 0 \leq x \leq 1, \\ u(0) = b. \end{aligned} \right\}$$

Полученные там оценки разности  $\delta(x) = u(x_k) - u_k^{(h)}$  между точным и приближенным решениями означают, что для первой из этих схем имеет место сходимость порядка  $h$ , а для второй — сходимость порядка  $h^2$ .

Обладание свойством сходимости является фундаментальным требованием, которое предъявляется к разностной схеме (11) для численного решения дифференциальной краевой задачи (1). Если оно имеет место, то с помощью разностной схемы (11) можно вычислить решение  $u$  с любой наперед заданной точностью, выбирая для этого  $h$  достаточно малым. Мы точно сформулировали понятие сходимости и подошли к центральному вопросу о том, как построить сходящуюся разностную схему (11) для вычисления решения дифференциальной краевой задачи (1). Приведенные выше примеры дополняют рассмотренные в гл. 1 и дают представление о простейшем способе построения таких схем: следует выбрать сетку и заменить производные разностными отношениями. Однако для одной и той же дифференциальной краевой задачи, как мы видели, можно получить различные разностные схемы (11), по-разному выбирая сетку  $D_h$  и по-разному заменяя производные приближающими их разностными отношениями. Мы уже видели на примере простейшего обыкновенного дифференциального уравнения из § 6, что разностная схема может оказаться непригодной для счета.

**3. Проверка сходимости разностной схемы.** Не будем пока записываться построением разностных схем и поставим задачу несколько иначе. Пусть разностная схема  $L_h u^{(h)} = f^{(h)}$ , позволяющая надеяться, что сходимость

$$\| [u]_h - u^{(h)} \|_{U_h} \rightarrow 0 \quad \text{при} \quad h \rightarrow 0$$

имеет место, на основании тех или иных соображений уже построена. Как проверить, является ли она в самом деле сходящейся?

Предположим, что разностная задача (11) имеет единственное решение  $u^{(h)} \in U_h$ . Если бы при подстановке в левую часть (11) вместо  $u^{(h)}$  сеточной функции  $[u]_h \in U_h$  равенство (11) оказалось бы в точности выполненным, то ввиду единственности

решения имело бы место равенство  $[u]_h = u^{(h)}$ , идеальное с точки зрения сходимости. Это означало бы, что решение  $u^{(h)}$  разностной задачи  $L_h u^{(h)} = f^{(h)}$  совпадает с искомой сеточной функцией  $[u]_h$ , которую мы условились считать точным решением.

Однако, как правило, систему (11) не удастся выбрать так, чтобы  $[u]_h$  в точности ей удовлетворяла. При подстановке  $[u]_h$  в уравнении (11) возникает некоторая невязка:

$$L_h [u]_h = f^{(h)} + \delta f^{(h)}. \quad (14)$$

Если эта невязка  $\delta f^{(h)}$  «стремится к нулю» при  $h \rightarrow 0$ , так что  $[u]_h$  удовлетворяет уравнению (11) все точнее, то будем говорить, что разностная схема  $L_h u^{(h)} = f^{(h)}$  аппроксимирует дифференциальную краевую задачу  $Lu = f$  на решении  $u$  последней.

В случае аппроксимации можно считать, что уравнение (14), которому удовлетворяет  $[u]_h$ , получается из уравнения (11) путем прибавления некоторой малой (при малом  $h$ ) добавки  $\delta f^{(h)}$  к правой части  $f^{(h)}$ . Следовательно, если решение  $u^{(h)}$  задачи (11) устойчиво относительно возмущения правой части  $f^{(h)}$ , т. е. мало изменяется при малом изменении правой части, то решение  $u^{(h)}$  задачи (11) и решение  $[u]_h$  задачи (14) отличаются мало, так что из аппроксимации

$$\delta f^{(h)} \rightarrow 0 \quad \text{при} \quad h \rightarrow 0$$

следует сходимость

$$u^{(h)} \rightarrow [u]_h \quad \text{при} \quad h \rightarrow 0.$$

Намеченный нами путь проверки сходимости (12) состоит в том, чтобы разбить этот трудный вопрос на два более простых: сначала проверить, имеет ли место аппроксимация задачи (1), задачей (11), а затем выяснить, устойчива ли задача (11). В этом содержится и указание на способы построения сходящихся разностных схем для численного решения задачи (1): надо строить аппроксимирующую ее разностную схему; из многих возможных способов аппроксимации надо выбирать такие, при которых разностные схемы оказываются устойчивыми.

Изложенный общий план исследования сходимости, естественно, предполагает, что введены математически строгие понятия аппроксимации и устойчивости, позволяющие доказать теорему о том, что из аппроксимации и устойчивости следует сходимость. Намеченные выше определения аппроксимации и устойчивости не являются строгими. Для определения аппроксимации надо еще уточнить, что такое невязка  $\delta f^{(h)}$  в общем случае и что такое ее величина, а для определения устойчивости — придать точный смысл словам «малому возмущению правой

части соответствует малое возмущение решения разностной задачи  $L_h u^{(h)} = f^{(h)}$ .

Строгим определениям понятий аппроксимации и устойчивости мы посвятим отдельные параграфы.

### ЗАДАЧИ

1. Разделить отрезок  $[0, 1]$  на  $N$  частей точками  $x_0 = 0, x_1, x_2, \dots, x_{N-1}, x_N = 1$  так, чтобы

$$\frac{x_{n+1} - x_n}{x_n - x_{n-1}} = q,$$

и выяснить, можно ли последовательность таких сеток при  $N \rightarrow \infty$  ( $q$  — не зависящая от  $N$  постоянная) использовать для приближенного решения задачи

$$\left. \begin{aligned} u' - u &= 0, \\ u(0) &= 1 \end{aligned} \right\}$$

с помощью разностной схемы

$$\left. \begin{aligned} \frac{u^{(h)}(x_{n+1}) - u^{(h)}(x_n)}{x_{n+1} - x_n} - u^{(h)}(x_n) &= 0 \quad \left( h = \frac{1}{N} \right) \\ u^{(h)}(x_0) &= 1. \end{aligned} \right\}$$

Стремится ли к нулю при  $N \rightarrow \infty$  максимальный из шагов  $x_{n+1} - x_n$ ?  
У к а з а н и е. Проще всего разобрать случай  $q > 1$  и убедиться, что

$$\lim_{N \rightarrow \infty} u^{(1/N)}(x_N) = \infty.$$

## § 11. Аппроксимация дифференциальной краевой задачи разностной схемой

1. Невязка  $\delta f^{(h)}$ . Придадим точный смысл понятию аппроксимации дифференциальной краевой задачи (1) из § 10

$$Lu = f \tag{1}$$

на решении  $u$  разностной схемой (11) из § 10

$$L_h u^{(h)} = f^{(h)}. \tag{2}$$

Для этого надо уточнить, что такое невязка  $\delta f^{(h)}$

$$L_h [u]_h = f^{(h)} + \delta f^{(h)}, \tag{3}$$

возникающая при подстановке сеточной функции  $[u]_h$  — таблицы искомого решения  $u$  — в уравнение (2), а также что такое ее величина.

Стремление величины невязки  $\delta f^{(h)}$  к нулю при  $h \rightarrow 0$  мы и примем затем за определение аппроксимации.

Начнем с рассмотрения примера разностной схемы для численного решения дифференциальной краевой задачи

$$\left. \begin{aligned} \frac{d^2u}{dx^2} + a(x) \frac{du}{dx} + b(x)u &= \cos x, \quad 0 \leq x \leq 1, \\ u(0) &= 1, \\ u'(0) &= 2. \end{aligned} \right\} \quad (4)$$

За сетку  $D_h$  по-прежнему примем совокупность точек  $x_n = nh$ ,  $n = 0, 1, \dots, N$ ;  $h = 1/N$ . В качестве разностной схемы для приближенного вычисления  $[u]_h$  воспользуемся совокупностью равенств

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + a(x_n) \frac{u_{n+1} - u_{n-1}}{2h} + b(x_n)u_n &= \\ &= \cos x_n, \quad n = 1, 2, \dots, N-1, \\ u_0 &= 1, \\ \frac{u_1 - u_0}{h} &= 2, \end{aligned} \right\} \quad (5)$$

возникшей при замене производных в (4) по приближенным формулам

$$\left. \begin{aligned} \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} &\approx \frac{d^2u(x)}{dx^2}, \\ \frac{u(x+h) - u(x-h)}{2h} &\approx \frac{du(x)}{dx}, \\ \frac{u(h) - u(0)}{h} &\approx \frac{du(0)}{dx}, \end{aligned} \right\} \quad (6)$$

Разностная схема (5) записывается в форме (2), если обозначить

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + a(nh) \frac{u_{n+1} - u_{n-1}}{2h} + b(nh)u_n, \\ u_0, \\ \frac{u_1 - u_0}{h}, \end{cases} \quad (7)$$

$$f^{(h)} = \begin{cases} \cos nh, \\ 1, \\ 2. \end{cases}$$

Для вычисления и оценки величины невязки  $\delta f^{(h)}$ , возникающей при подстановке  $[u]_h$  в уравнение (2), уточним формулы (6).

По формуле Тейлора имеем

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(\xi_1),$$

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{6}u'''(\xi_2),$$

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(\xi_3),$$

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(\xi_4),$$

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(\xi_5).$$

Здесь  $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$  — некоторые промежуточные точки отрезка  $[x-h, x+h]$ .

Отсюда

$$\left. \begin{aligned} \frac{u(x+h) - u(x-h)}{2h} &= u'(x) + \frac{h^2}{12} [u'''(\xi_1) + u'''(\xi_2)], \\ \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} &= u''(x) + \frac{h^2}{24} [u^{(4)}(\xi_3) + u^{(4)}(\xi_4)], \\ \frac{u(x+h) - u(x)}{h} &= u'(x) + \frac{h}{2} u''(\xi_5). \end{aligned} \right\} \quad (8)$$

**2. Вычисление невязки.** Будем считать, что решение  $u(x)$  задачи (4) имеет ограниченные производные до четвертого порядка. В силу формул (8) можно написать

$$\begin{aligned} \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + a(x) \frac{u(x+h) - u(x-h)}{2h} + \\ + b(x)u(x) = \frac{d^2u(x)}{dx^2} + a(x) \frac{du(x)}{dx} + b(x)u(x) + \\ + h^2 \left[ \frac{u^{(4)}(\xi_3) + u^{(4)}(\xi_4)}{24} + a(x) \frac{u'''(\xi_1) + u'''(\xi_2)}{12} \right]. \end{aligned}$$

Поэтому выражение

$$L_h[u]_h = \begin{cases} \frac{u(x_n+h) - 2u(x_n) + u(x_n-h)}{h^2} + \\ + a(x_n) \frac{u(x_n+h) - u(x_n-h)}{2h} + b(x_n)u(x_n), & n = 1, 2, \dots, N-1, \\ u(0), \\ \frac{u(h) - u(0)}{h} \end{cases}$$

можно переписать так:

$$L_h[u]_h = \begin{cases} \cos x_n + h^2 \left[ \frac{u^{(4)}(\xi_3) + u^{(4)}(\xi_4)}{24} + \right. \\ \quad \left. + a(x_n) \frac{u'''(\xi_1) + u'''(\xi_2)}{12} \right], & n = 1, 2, \dots, N-1, \\ 1 + 0, \\ 2 + h \frac{u''(\xi_5)}{2} \end{cases}$$

или

$$L_h[u]_h = f^{(h)} + \delta f^{(h)},$$

где

$$\delta f^{(h)} = \begin{cases} h^2 \left[ \frac{u^{(4)}(\xi_3) + u^{(4)}(\xi_4)}{24} + \frac{u'''(\xi_1) + u'''(\xi_2)}{12} \right], \\ 0, \\ h \frac{u''(\xi_5)}{2}. \end{cases} \quad (9)$$

Удобно считать, что  $f^{(h)}$  и  $\delta f^{(h)}$ , заданные формулами (7) и (9), принадлежат линейному нормированному пространству  $F_h$ , которое состоит из элементов вида

$$g^{(h)} = \begin{cases} \varphi_n & (n = 1, 2, \dots, N-1, \\ \psi_0, \\ \psi_1, \end{cases} \quad (10)$$

где  $\varphi_1, \varphi_2, \dots, \varphi_{n-1}$ , а также  $\psi_0$  и  $\psi_1$  — произвольная упорядоченная система чисел; можно считать, что  $g^{(h)}$  — это совокупность сеточной функции  $\varphi_n$ ,  $n = 1, 2, \dots, N-1$ , и упорядоченной пары чисел  $\psi_0$  и  $\psi_1$ . Сложение двух элементов пространства  $F_h$  и умножение элементов  $g^{(h)}$  на числа производятся покомпонентно. Ясно, что в рассматриваемом примере  $F_h$  есть  $(N+1)$ -мерное линейное пространство. Норма в  $F_h$  может быть введена многими способами. Если ввести в  $F_h$  норму равенством

$$\|g^{(h)}\|_{F_h} = \max(|\psi_0|, |\psi_1|, \max_n |\varphi_n|),$$

т. е. принять за норму максимум абсолютных величин всех компонент вектора  $g^{(h)}$ , то в силу (9) получим

$$\|\delta f^{(h)}\|_{F_h} \leq Ch, \quad (11)$$

где  $C$  — некоторая постоянная, зависящая от  $u(x)$ , но не зависящая от  $h$ .

Из этого неравенства следует стремление невязки  $\delta f^{(h)}$  к нулю при  $h \rightarrow 0$ .

В уравнении  $L_h u^{(h)} = f^{(h)}$ , подробно записанном равенствами (5), которое мы рассмотрели в качестве примера, на  $L_h$  можно смотреть как на оператор. Этот оператор каждой сеточной функции  $v^{(h)} = \{v_n\}$ ,  $n = 0, 1, \dots, N$ , из линейного пространства функций, определенных на сетке  $D_h$ , ставит в соответствие некоторый элемент  $g^{(h)}$  вида (10) из линейного пространства  $F_h$  по формуле

$$L_h v^{(h)} \equiv \begin{cases} \frac{v_{n+1} - 2v_n + v_{n-1}}{h^2} + a(x_n) \frac{v_{n+1} - v_{n-1}}{2h} + b(x_n) v_n, \\ v_0, \\ \frac{v_1 - v_0}{h}. \end{cases}$$

Условимся и в общем случае разностной краевой задачи (2) считать, что правые части тех скалярных уравнений, которые в совокупности записаны символическим равенством

$$L_h u^{(h)} = f^{(h)},$$

являются компонентами вектора  $f^{(h)}$  из некоторого линейного нормированного пространства  $F_h$ . Тогда на  $L_h$  можно смотреть как на оператор, ставящий в соответствие каждой сеточной функции  $u^{(h)}$  из  $U_h$  некоторый элемент  $f^{(h)}$  из  $F_h$ .

В таком случае имеет смысл выражение  $L_h[u]_h$ , возникающее в результате применения оператора  $L_h$  к сеточной функции  $[u]_h$  из  $U_h$  и являющееся элементом пространства  $F_h$ .

Невязка  $\delta f^{(h)} = L_h[u]_h - f^{(h)}$  принадлежит пространству  $F_h$ , как разность двух элементов этого пространства. Под величиной невязки следует понимать  $\|\delta f^{(h)}\|_{F_h}$ .

### 3. Аппроксимация порядка $h^k$ .

Определение. Будем говорить, что разностная схема  $L_h u^{(h)} = f^{(h)}$  аппроксимирует задачу  $Lu = f$  на решении  $u$ , если  $\|\delta f^{(h)}\|_{F_h} \rightarrow 0$  при  $h \rightarrow 0$ . Если, сверх того, имеет место неравенство

$$\|\delta f^{(h)}\|_{F_h} \leq ch^k,$$

где  $c > 0$  и  $k > 0$  — некоторые постоянные, то будем говорить, что имеет место *аппроксимация порядка  $h^k$*  или порядка  $k$  относительно величины  $h$ .

То обстоятельство, что  $u$  является решением задачи (1), дает информацию о функции  $u$ , которую можно использовать для построения системы (2), а также для проверки факта аппроксимации. Поэтому в определении аппроксимации мы и упоминаем



задачу (1). Однако подчеркнем, что приведенное определение аппроксимации задачи  $Lu = f$  на решении  $u$  разностной схемой  $L_h u^{(h)} = f^{(h)}$  само по себе не опирается на равенство  $Lu = f$  для функции  $u$ . Можно было бы говорить просто о том, что схема  $L_h u^{(h)} = f^{(h)}$  соответствует с порядком  $h^k$  функции  $u$ , не вникая в происхождение этой функции. В частности, если функция  $u$  является одновременно решением двух совсем различных задач  $L^{(1)}u = f^{(1)}$  и  $L^{(2)}u = f^{(2)}$  вида (1), то одна и та же разностная схема  $L_h u^{(h)} = f^{(h)}$  одновременно аппроксимирует или не аппроксимирует каждую из этих задач на их общем решении  $u$ .

#### 4. Примеры.

Пример 1. Разностная схема (5) ввиду оценки (11) аппроксимирует задачу (4) с первым порядком относительно  $h$ . Разностную схему (5) легко усовершенствовать так, чтобы аппроксимация стала порядка  $h^2$ . Для этого заметим, что все компоненты вектора  $\delta f^{(h)}$ , кроме последней, стремятся к нулю, как  $h^2$  (предпоследняя даже в точности равна нулю).

Только последняя компонента вектора  $\delta f^{(h)}$ , т. е. невязка от подстановки  $[u]_h$  в последнее уравнение  $\frac{u_1 - u_0}{h} = 2$  системы (5) стремится к нулю медленнее, а именно как первая степень  $h$ . Это досадное обстоятельство легко устранить. По формуле Тейлора

$$\begin{aligned} \frac{u(h) - u(0)}{h} &= u'(0) + \frac{h}{2} u''(0) + \frac{h^2}{6} u'''(\xi) = \\ &= 2 + \frac{h}{2} u''(0) + \frac{h^2}{6} u'''(\xi), \quad 0 < \xi < h. \end{aligned}$$

Но из дифференциального уравнения (4) находим

$$u''(0) = -a(0)u'(0) - b(0)u(0) + \cos 0 = -2a(0) - b(0) + 1.$$

Поэтому, заменив последнее равенство (5) равенством

$$\frac{u_1 - u_0}{h} = 2 - \frac{h}{2} [2a(0) + b(0) - 1], \quad (12)$$

получим для  $f^{(h)}$  вместо (7) выражение

$$f^{(h)} \equiv \begin{cases} \cos x_n, \\ 1, \\ 2 - \frac{h}{2} [2a(0) + b(0) - 1]. \end{cases}$$

Тогда окажется, что

$$\delta f^{(h)} \equiv \begin{cases} \frac{h^2}{12} \left[ \frac{u^{(4)}(\xi_3) + u^{(4)}(\xi_4)}{2} + (u'''(\xi_1) + u'''(\xi_2)) \right], \\ 0, \\ \frac{h^2}{6} u'''(\xi) \end{cases}$$

и  $\|\delta f^{(h)}\|_{F_h} < C_1 h^2$ , где  $C_1$  — некоторая постоянная, не зависящая от  $h$ . Порядок аппроксимации станет вторым относительно  $h$ .

Подчеркнем, что для построения разностного граничного условия (12) мы использовали не только граничные условия задачи (4), но и самое дифференциальное уравнение. Можно считать, что мы использовали граничное условие

$$u''(x) + a(x)u'(x) + b(x)u(x)|_{x=0} = \cos x|_{x=0},$$

которое является следствием дифференциального уравнения.

Пример 2. Выясним, каков порядок аппроксимации, которым обладает разностная схема

$$L_n u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - u_{n-1}}{2h} + Au_n = 1 + x_n^2, & n = 1, 2, \dots, N-1, \\ u_0 = b, \\ u_1 = b \end{cases} \quad (13)$$

на решении  $u$  задачи

$$\left. \begin{aligned} \frac{du}{dx} + Au &= 1 + x^2, \\ u(0) &= b. \end{aligned} \right\} \quad (14)$$

Подобную схему мы рассматривали в § 8 еще до того, как было введено строгое понятие аппроксимации.

Роль  $f^{(h)}$  здесь играет

$$f^{(h)} = \begin{cases} 1 + x_n^2, & n = 1, 2, \dots, N-1, \\ b, \\ b. \end{cases}$$

Далее,

$$L_h [u]_h \equiv \begin{cases} \frac{u(x_n + h) - u(x_n - h)}{2h} + Au(x_n), & n = 1, \dots, N-1, \\ u(0), \\ u(h) \end{cases}$$

ИЛИ

$$L_h[u]_h \equiv \begin{cases} \left[ \frac{du(x_n)}{dx} + Au(x_n) \right] + \frac{h^2}{6} u'''(\xi_n), & n = 1, \dots, N-1, \\ u(0), \\ u(0) + h \frac{du(\xi_0)}{dx}. \end{cases}$$

Так как для решения  $u(x)$  выполнено равенство

$$\frac{du(x_n)}{dx} + Au(x_n) = 1 + x_n^2,$$

то невязка  $\delta f^{(h)}$  имеет вид

$$\delta f^{(h)} = \begin{cases} \frac{h^2}{6} u'''(\xi_n), & n = 1, \dots, N-1, \\ 0, \\ hu'(\xi_0). \end{cases}$$

Аппроксимация задачи (14) схемой (13) имеет первый относительно  $h$  порядок. Бросается в глаза, что компоненты невязки, как и в примере 1, имеют различный порядок относительно  $h$ . Разностное уравнение

$$\frac{u_{n+1} - u_{n-1}}{2h} + Au_n = 1 + x_n^2, \quad n = 1, 2, \dots, N-1, \quad (15)$$

при подстановке  $[u]_h$  удовлетворяется с невязкой  $\frac{h^2}{6} u'''(\xi_n)$  порядка  $h^2$ . Первое граничное условие

$$u_0 = b \quad (16)$$

при подстановке  $[u]_h$  выполнено точно, а второе

$$u_1 = b \quad (17)$$

— с невязкой  $h u'(\xi_0)$  порядка первой степени  $h$ .

Погрешность аппроксимации мы оценили через

$$\max_{0 \leq x \leq 1} |u'''(x)|, \quad \max_{0 \leq x \leq 1} |u'(x)|.$$

В рассматриваемом примере точное решение

$$u(x) = u(0) e^{-Ax} + \frac{1 + x^2 - e^{-Ax}}{A} - \frac{2x}{A^2}$$

позволяет оценить эти максимумы через данные задачи  $u(0)$ ,  $A$ :

$$\begin{aligned} \max_{0 \leq x \leq 1} |u'(x)| &= \max_{0 \leq x \leq 1} \left| Au(0)e^{-Ax} + \frac{2x + Ae^{-Ax}}{A} - \frac{2}{A^2} \right| \leq \\ &\leq |u(0)| |A| (1 + e^{-A}) + \frac{2}{|A|} + \frac{2}{A^2} + (1 + e^{-A}), \\ \max_{0 \leq x \leq 1} |u'''(x)| &= \left| \left[ u(0) + \frac{1}{A} \right] A^3 e^{-Ax} \right| \leq \\ &\leq [|u(0)| |A|^3 + |A|^2] (1 + e^{-A}). \end{aligned}$$

В более сложных примерах приходится ограничиваться грубой оценкой этих производных, основанной на теореме о дифференцируемости решений обыкновенных дифференциальных уравнений в случае гладких правых частей.

**5. Разбиение разностной схемы на подсистемы.** Для подробного описания характера аппроксимации нам оказалось удобным говорить не сразу обо всей разностной схеме (13) вида (2)

$$L_h u^{(h)} = f^{(h)},$$

но отдельно о подсистемах (15), (16), (17). Эти подсистемы (две последние состоят каждая из одного уравнения) можно записать соответственно следующими символическими равенствами:

$$l_h^0 u^{(h)} = f_0^{(h)}, \quad (18)$$

$$l_h^1 u^{(h)} = f_1^{(h)}, \quad (19)$$

$$l_h^2 u^{(h)} = f_2^{(h)}. \quad (20)$$

Для этого надо положить

$$l_h^0 u^{(h)} = \frac{u_{n+1} - u_{n-1}}{2h} + Au_n, \quad n = 1, 2, \dots, N-1,$$

$$l_h^1 u^{(h)} \equiv u_0,$$

$$l_h^2 u^{(h)} \equiv u_1,$$

$$f_0^{(h)} = 1 + x_n^2,$$

$$f_1^{(h)} = b,$$

$$f_2^{(h)} = b.$$

Для удобства речи и в общем случае разностную схему (2) часто разбивают на две или несколько подсистем:

$$\left. \begin{aligned} l_h^0 u^{(h)} &= f_0^{(h)}, \\ \cdot &\cdot \cdot \cdot \\ l_h^r u^{(h)} &= f_r^{(h)}, \\ \cdot &\cdot \cdot \cdot \\ l_h^R u^{(h)} &= f_R^{(h)}, \end{aligned} \right\} \quad (21)$$

так что

$$L_h u^{(h)} \equiv \begin{cases} l_h^0 u^{(h)}, \\ \cdot \cdot \cdot \\ l_h^r u^{(h)}, \\ \cdot \cdot \cdot \\ l_h^R u^{(h)}, \end{cases} \quad f^{(h)} = \begin{cases} f_0^{(h)}, \\ \cdot \cdot \cdot \\ f_r^{(h)}, \\ \cdot \cdot \cdot \\ f_R^{(h)}. \end{cases}$$

Правую часть  $f_r^{(h)}$  каждой подсистемы  $l_h^r u^{(h)} = f_r^{(h)}$  удобно считать элементом линейного нормированного пространства  $F_h^{(r)}$ . Нормы в пространстве  $F_h$  и пространствах  $F_h^{(1)}, F_h^{(2)}, \dots, F_h^{(R)}$  удобно выбирать согласованно, чтобы имело место равенство

$$\|f^{(h)}\|_{F_h} = \max_r \|f_r^{(h)}\|_{F_h^{(r)}}. \quad (22)$$

Разбивая (2) на подсистемы (21), мы всегда будем считать, что (22) выполняется.

Удобство разбиения разностной схемы  $L_h u^{(h)} = f^{(h)}$  на подсистемы (21) состоит в том, что можно говорить о порядке соответствия каждой подсистемы в отдельности решению и задачи (1),  $Lu = f$ . За этот порядок принимается порядок убывания нормы  $\|\delta f_r^{(h)}\|_{F_h^{(r)}}$  невязки  $\delta f_r^{(h)}$

$$l_h^{(r)} [u]_h = f_r^{(h)} + \delta f_r^{(h)},$$

при  $h \rightarrow 0$ . Порядок аппроксимации всей разностной схемы  $L_h u^{(h)} = f^{(h)}$  на решении  $u$  задачи  $Lu = f$ , благодаря согласованному выбору норм (22), равен порядку убывания нормы  $\|\delta f_r^{(h)}\|_{F_h^{(r)}}$  невязки  $\delta f_r^{(h)}$  при том  $r$ , при котором она убывает медленнее всего.

В примере 2 при разбиении системы (13) на подсистемы (15)–(17), или (18)–(20), пространство  $F_h^{(0)}$  состоит из сеточных функций  $f_0^{(h)} = \{f_n\}$  с нормой  $\|f_0^{(h)}\| = \max_n |f_n|$ , определенных в точках  $x_n = nh$ ,  $n = 1, 2, \dots, N-1$ , а пространства

$F_h^{(1)}$  и  $F_h^{(2)}$  одномерны и состоят из чисел с нормой  $\|a\| = |a|$ .  
Уравнение (18):

$$I_h^{(0)}u^{(h)} = f_0^{(h)},$$

соответствует задаче (14) на решении  $u$  со вторым порядком, уравнение  $I_h^{(1)}u^{(h)} = f_1^{(h)}$  соответствует точно, а уравнение  $I_h^{(2)}u^{(h)} = f_2^{(h)}$  — с первым порядком. Чтобы повысить порядок аппроксимации, которым обладает разностная схема (13), с первого до второго относительно  $h$ , достаточно «подправить» только граничное условие  $I_h^{(2)}u^{(h)} = b$ . Заметим, что

$$I_h^{(2)}[u]_h = u(h) = u(0) + hu'(0) + \frac{h^2}{2}u''(\xi).$$

Учтем, что  $u(0) = b$  и что в силу (14)

$$u'(0) = -Au(0) + 1 = -Ab + 1.$$

Положив

$$I_h^{(2)}u^{(h)} = u_1 = b - hAb + h, \quad \text{т. е.} \quad f_2^{(h)} = b - hAb + h,$$

мы добьемся того, чтобы выполнялось условие

$$I_h^{(2)}[u]_h = u(h) = f_2^{(h)} + O(h^2),$$

т. е. чтобы имел место второй относительно  $h$  порядок соответствия граничного условия

$$I_h^{(2)}u^{(h)} = f_2^{(h)} \quad (f_2^{(h)} = b - hAb + h) \quad (23)$$

задаче (14) на решении  $u$ . Таким образом, разностная схема (15), (16), (23) аппроксимирует задачу (14) со вторым порядком относительно  $h$ .

Разбиение разностной схемы (2) на подсистемы (21) условно и делается только для удобства речи. Так, например, систему (13) можно было бы разбить на две подсистемы, отнеся к первой по-прежнему разностное уравнение (15), а ко второй — оба граничных условия (16) и (17). Мы получили бы символическую запись

$$\left. \begin{aligned} I_h^{(0)}u^{(h)} &= f_0^{(h)}, \\ I_h^{(1)}u^{(h)} &= f_1^{(h)}, \end{aligned} \right\}$$

где

$$I_h^{(1)} = \begin{Bmatrix} u_0, \\ u_1, \end{Bmatrix} \quad f_1^{(h)} = \begin{pmatrix} b \\ b \end{pmatrix}.$$

Однако при таком разбиении на подсистемы, в отличие от разбиения (15)—(17) или (18)—(20), мы лишились бы возможности коротко выразить то обстоятельство, что первое граничное условие при подстановке  $[u]_h$  выполняется точно, а второе — лишь с первым относительно  $h$  порядком.

**6. Замена производных разностными отношениями.** В рассмотренных примерах для получения разностных схем мы заменяли производные в дифференциальном уравнении разностными отношениями. Этот прием весьма универсален и позволяет построить для любой дифференциальной краевой задачи, имеющей достаточно гладкое решение  $u(x)$ , разностную схему с любым наперед заданным порядком аппроксимации.

Действительно, покажем, что производную  $d^k u/dx^k$  произвольного порядка  $k$  можно заменить разностным отношением так, чтобы погрешность от такой замены для достаточно гладкой функции  $u(x)$  была любого наперед заданного порядка  $p$  относительно шага  $h$  разностной сетки. Воспользуемся для этого методом неопределенных коэффициентов.

Напишем равенство вида

$$\frac{d^k u(x)}{dx^k} = h^{-k} \sum_{s=-s_1}^{s_2} a_s u(x+sh) + o(h^p) \quad (24)$$

и постараемся подобрать не зависящие от  $h$  неопределенные коэффициенты  $a_s$ ,  $s = -s_1, -s_1+1, \dots, s_2$ , так, чтобы оно оказалось справедливым. Пределы суммирования  $s_1 \geq 0$  и  $s_2 \geq 0$  можно взять произвольными, но так, чтобы порядок  $s_1 + s_2$  разностного отношения  $h^{-k} \sum a_s u(x+sh)$  удовлетворял неравенству  $s_1 + s_2 \geq k + p - 1$ . По формуле Тейлора

$$u(x+sh) = u(x) + sh \frac{du(x)}{dx} + \frac{(sh)^2}{2!} \frac{d^2 u(x)}{dx^2} + \dots \\ \dots + \frac{(sh)^{k+p-1}}{(k+p-1)!} \frac{d^{k+p-1} u(x)}{dx^{k+p-1}} + \frac{(sh)^{k+p}}{(k+p)!} \frac{d^{k+p} u(\xi)}{dx^{k+p}}.$$

Подставим это выражение вместо  $u(x+sh)$  в (24) и приведем подобные члены. Получим

$$\frac{d^k u(x)}{dx^k} = h^{-k} \left[ u(x) \sum a_s + \frac{du(x)}{dx} \frac{h}{1!} \sum s a_s + \dots \right. \\ \left. \dots + \frac{d^{k+p-1} u(x)}{dx^{k+p-1}} \frac{h^{k+p-1}}{(k+p-1)!} \sum s^{k+p-1} a_s \right] + \\ + \frac{h^p}{(k+p)!} \sum s^{k+p} a_s \frac{d^{k+p} u(\xi)}{dx^{k+p}}.$$

Приравнявая коэффициенты при одинаковых степенях  $h^s$ ,  $s = -k, -k+1, \dots, p-1$ , в левой и правой частях этого равенства, получим

следующую систему уравнений для определения  $a_s$ :

$$\left. \begin{aligned} \sum a_s &= 0, \\ \sum s a_s &= 0, \\ \dots &\dots \\ \sum s^{k-1} a_s &= 0, \\ \sum s^k a_s &= k!, \\ \sum s^{k+1} a_s &= 0, \\ \dots &\dots \\ \sum s^{k+p-1} a_s &= 0. \end{aligned} \right\} \quad (25)$$

Если  $s_1 + s_2 = k + p - 1$ , то выписанные  $k + p$  равенств образуют линейную систему относительно того же числа неизвестных  $a_s$ . Определитель этой системы

$$\begin{vmatrix} 1 & & & 1 & \dots & 1 \\ -s_1 & & & -s_1 + 1 & \dots & s_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ (-s_1)^{k+p-1} & & & (-s_1 + 1)^{k+p-1} & \dots & s_2^{k+p-1} \end{vmatrix}$$

есть известный определитель Вандермонда и отличен от нуля. Таким образом, существует единственный набор коэффициентов  $a_s$ , удовлетворяющий системе (25). Если  $s_1 + s_2 \geq k + p$ , то, очевидно, таких систем коэффициентов  $a_s$  много.

Так, например, существует единственное разностное отношение первого порядка вида

$$h^{-1} [a_0 u(x) + a_1 u(x+h)],$$

приближающее  $du/dx$  с первым относительно  $h$  порядком. Оно получается при

$$\frac{du}{dx} = \frac{u(x+h) - u(x)}{h} + O(h).$$

Точно так же существует единственное разностное отношение первого порядка вида

$$h^{-1} [a_{-1} u(x-h) + a_0 u(x)],$$

приближающее  $du/dx$  с первым относительно  $h$  порядком:

$$\frac{du}{dx} = \frac{u(x) - u(x-h)}{h} + O(h).$$

Среди разностных отношений второго порядка вида

$$h^{-1} [a_{-1} u(x-h) + a_0 u(x) + a_1 u(x+h)]$$

существует бесконечно много приближающих  $du/dx$  с первым порядком относительно  $h$ , но только одно со вторым порядком. Решая систему (25) для этого случая увидим, что при  $a_1 = 1/2$ ,  $a_0 = 0$ ,  $a_{-1} = -1/2$

$$\frac{du}{dx} = \frac{u(x+h) - u(x-h)}{2h} + O(h^2).$$



Если мы хотим приблизить  $d^2u/dx^2$  с порядком  $h^2$ , то  $k = 2$ ,  $p = 2$  и надо, чтобы  $s_1 + s_2 \geq 3$ . Поэтому среди разностных отношений вида

$$h^{-2}(a_{-1}u(x-h) + a_0u(x) + a_1u(x+h) + a_2u(x+2h)) \quad (26)$$

только одно является искомым. Решая систему (26) для определения коэффициентов  $a_{-1}$ ,  $a_0$ ,  $a_1$ ,  $a_2$ , получим

$$a_{-1} = a_1 = 1, \quad a_0 = -2, \quad a_2 = 0,$$

т. е. уже неоднократно использованное нами равенство

$$\frac{d^2u(x)}{dx^2} = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + O(h^2).$$

**7. Другие способы построения разностных схем.** Замена производных разностными отношениями не единственный, а часто и не лучший способ построения разностных схем. Некоторым другим способам, приводящим к наиболее употребительным разностным схемам, будет посвящен § 19. Здесь ограничимся примером.

Простейшая разностная схема

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - u_n}{h} - G(x_n, u_n) = 0, & n = 0, 1, \dots, N-1, \\ u_0 = a, \end{cases}$$

называемая *схемой Эйлера*, аппроксимирует задачу

$$\left. \begin{aligned} \frac{du}{dx} - G(x, u) &= 0, & 0 \leq x \leq 1, \\ u(0) &= a \end{aligned} \right\} \quad (27)$$

с первым порядком относительно  $h$ . При известном  $u_n$  значение  $u_{n+1}$  вычисляется по формуле  $u_{n+1} = u_n + hG(x_n, u_n)$ . Схема

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - u_n}{h} - \frac{1}{2} [G(x_n, u_n) + G(x_{n+1}, \bar{u})] = 0, \\ u_0 = a, \end{cases}$$

где  $\bar{u} = u_n + hG(x_n, u_n)$ , называется *схемой Эйлера с пересечением*. Она же является одной из схем Рунге—Кутты второго порядка аппроксимации, о которых будет подробно рассказано в § 19. Если  $u_n$  уже вычислено, то по схеме Эйлера вычисляем значение

$$\bar{u} = u_n + hG(x_n, u_n),$$

а потом осуществляем уточнение найденного  $\bar{u}$ , полагая

$$u_{n+1} = u_n + \frac{h}{2} [G(x_n, u_n) + G(x_{n+1}, \bar{u})],$$

## ЗАДАЧИ

1. Проверить, что схема Эйлера с пересчетом аппроксимирует задачу (27) на гладком решении  $u(x)$  со вторым относительно  $h$  порядком.

## § 12. Определение устойчивости разностной схемы. Сходимость как следствие аппроксимации и устойчивости

1. **Определение устойчивости.** Пусть для приближенного вычисления решения  $u$  дифференциальной краевой задачи

$$Lu = f \quad (1)$$

составлена разностная схема

$$L_h u^{(h)} = f^{(h)}, \quad (2)$$

которая аппроксимирует задачу (1) на решении  $u$  с некоторым порядком  $h^k$ . Это значит, что невязка  $\delta f^{(h)}$

$$L_h [u]_h = f^{(h)} + \delta f^{(h)},$$

возникающая при подстановке таблицы  $[u]_h$  решения  $u$  в уравнение (2), удовлетворяет оценке вида

$$\|\delta f^{(h)}\|_{F_h} \leq C_1 h^k, \quad (3)$$

где  $C_1$  — некоторая постоянная, не зависящая от  $h$ . Легко проверить, что разностная схема

$$L_h u^{(h)} \equiv \begin{cases} 4 \frac{u_{n+1} - u_{n-1}}{2h} - 3 \frac{u_{n+1} - u_n}{h} + Au_n = 0, \\ n = 1, 2, \dots, N-1, \\ u_0 = b \end{cases}$$

аппроксимирует

$$\begin{aligned} \frac{du}{dx} + Au &= 0, \\ u(0) &= b \end{aligned}$$

на решении  $u$  с первым порядком относительно  $h$ . Однако, как показано в § 9, решение  $u^{(h)}$ , доставляемое этой разностной схемой, не стремится к  $[u]_h$  при  $h \rightarrow 0$ .

Таким образом, аппроксимации, вообще говоря, недостаточно для сходимости. Нужна еще устойчивость.

**Определение 1.** Будем называть разностную схему (2) *устойчивой*, если существуют числа  $h_0 > 0$  и  $\delta > 0$  такие, что при любом  $h < h_0$  и любом  $e^{(h)} \in F_h$ ,  $\|e^{(h)}\|_{F_h} < \delta$  разностная задача

$$L_h z^{(h)} = f^{(h)} + e^{(h)}, \quad (4)$$

полученная из задачи (2) добавлением к правой части возмущения  $\varepsilon^{(h)}$ , имеет одно и только одно решение  $z^{(h)}$ , причем это решение отклоняется от решения  $u^{(h)}$  невозмущенной задачи (2) на сеточную функцию  $z^{(h)} - u^{(h)}$ , удовлетворяющую оценке

$$\|z^{(h)} - u^{(h)}\|_{U_h} \leq C \|\varepsilon^{(h)}\|_{F_h}, \quad (5)$$

где  $C$  — некоторая постоянная, не зависящая от  $h$ .

В частности, неравенство (5) означает, что малое возмущение  $\varepsilon^{(h)}$  правой части разностной схемы (2) вызывает равномерно относительно  $h$  малое возмущение  $z^{(h)} - u^{(h)}$  решения.

Пусть оператор  $L_h$ , отображающий  $U_h$  в  $F_h$ , линейный. Тогда приведенное выше определение устойчивости равносильно следующему:

Определение 2. Будем называть разностную схему (2) с линейным оператором  $L_h$  *устойчивой*, если при любом  $f^{(h)} \in F_h$  уравнение  $L_h u^{(h)} = f^{(h)}$  имеет единственное решение  $u^{(h)} \in U_h$ , причем

$$\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}, \quad (6)$$

где  $C$  — некоторая постоянная, не зависящая от  $h$ .

Докажем равносильность обоих определений устойчивости в случае линейного оператора  $L_h$ .

Сначала установим, что из устойчивости разностной схемы (2) в смысле определения 2 следует устойчивость в смысле определения 1. Пусть линейная задача (2) при всех рассматриваемых  $h < h_0$  и произвольном  $f^{(h)} \in F_h$  имеет единственное решение, причем выполнена оценка (6). Вычитая из равенства (4) равенство (2), получим

$$L_h(z^{(h)} - u^{(h)}) = \varepsilon^{(h)},$$

откуда в силу (6) следует оценка (5) при произвольном  $\varepsilon^{(h)} \in F_h$ , а значит, и устойчивость в смысле определения 1.

Покажем теперь, что устойчивость в смысле определения 1 влечет за собой устойчивость в смысле определения 2. В силу определения 1 при некоторых  $h_0 > 0$  и  $\delta > 0$  и при произвольных  $h < h_0$  и  $\varepsilon^{(h)} \in F_h$ ,  $\|\varepsilon^{(h)}\|_{F_h} < \delta$  существуют и единственные решения уравнений

$$\begin{aligned} L_h z^{(h)} &= f^{(h)} + \varepsilon^{(h)}, \\ L_h u^{(h)} &= f^{(h)}. \end{aligned}$$

Положим  $w^{(h)} \equiv z^{(h)} - u^{(h)}$  и вычтем эти равенства почленно. Получим

$$L_h w^{(h)} = \varepsilon^{(h)},$$

причем в силу (5)

$$\|\omega^{(h)}\|_{U_h} \leq C \|\varepsilon^{(h)}\|_{F_h}.$$

Очевидно, что, изменив обозначения решения и правой части уравнения  $L_h \omega^{(h)} = \varepsilon^{(h)}$ , последний результат можно сформулировать так: при произвольных  $h < h_0$  и  $f^{(h)} \in F_h$ ,  $\|f^{(h)}\|_{F_h} < \delta$  задача (2) имеет единственное решение  $u^{(h)}$ . Это решение удовлетворяет оценке (6). Однако в таком случае уравнение (2) имеет единственное решение  $u^{(h)}$  и выполнена оценка (6) не только для всех  $f^{(h)}$ , удовлетворяющих оценке  $\|f^{(h)}\|_{F_h} < \delta$ , но и вообще для всех  $f^{(h)} \in F_h$ , т. е. имеет место устойчивость в смысле определения 2.

В самом деле, пусть  $\|f^{(h)}\|_{F_h} \geq \delta$ . Докажем однозначную разрешимость и оценку (6) в этом случае. Положим

$$u^{(h)} = \frac{2\|j^{(h)}\|_{F_h}}{\delta} \tilde{u}^{(h)}, \quad f^{(h)} = \frac{2\|j^{(h)}\|_{F_h}}{\delta} \tilde{f}^{(h)}.$$

Для  $\tilde{u}^{(h)}$  получим уравнение

$$L_h \tilde{u}^{(h)} = \tilde{f}^{(h)},$$

причем

$$\|\tilde{f}^{(h)}\|_{F_h} = \frac{\delta}{2\|j^{(h)}\|_{F_h}} \|f^{(h)}\|_{F_h} = \frac{\delta}{2} < \delta.$$

Поэтому уравнение  $L_h \tilde{u}^{(h)} = \tilde{f}^{(h)}$  однозначно разрешимо, причем

$$\|\tilde{u}^{(h)}\|_{U_h} \leq C \|\tilde{f}^{(h)}\|_{F_h}.$$

В силу формул, устанавливающих связь между  $u^{(h)}$  и  $\tilde{u}^{(h)}$ , а также между  $f^{(h)}$  и  $\tilde{f}^{(h)}$ , отсюда следует однозначная разрешимость задачи (2) и справедливость оценки (6) при произвольном рассматриваемом  $f^{(h)} \in F_h$ .

**2. Зависимость между аппроксимацией, устойчивостью и сходимостью.** Докажем теперь, что из аппроксимации и устойчивости следует сходимость.

**Теорема.** Пусть разностная схема  $L_h u^{(h)} = f^{(h)}$  аппроксимирует задачу  $Lu = f$  на решении и с порядком  $h^k$  и устойчива. Тогда решение  $u^{(h)}$  разностной задачи  $L_h u^{(h)} = f^{(h)}$  сходится к  $[u]_h$ , причем имеет место оценка

$$\|[u]_h - u^{(h)}\|_{U_h} \leq (CC_1) h^k, \quad (7)$$

где  $C$  и  $C_1$  — числа, входящие в оценки (3) и (5).

Доказательство. Положим  $\varepsilon^{(h)} \equiv \delta f^{(h)}$ ,  $[u]_h \equiv z^{(h)}$ . Тогда оценка (5) примет вид

$$\| [u]_h - u^{(h)} \|_{U_h} \leq C \| \delta f^{(h)} \|_{F_h}.$$

Учитывая (3), сразу получаем доказываемое неравенство (7).

В качестве иллюстрирующего примера докажем устойчивость разностной схемы Эйлера

$$\left. \begin{aligned} \frac{u_{n+1} - u_n}{h} - G(x_n, u_n) &= \varphi_n, \quad n = 0, 1, \dots, N-1, \\ u_0 &= \psi, \end{aligned} \right\} \quad (8)$$

$x_n = nh$ ,  $h = 1/N$ , для численного решения дифференциальной краевой задачи

$$\left. \begin{aligned} \frac{du}{dx} - G(x, u) &= \varphi(x), \quad 0 \leq x \leq 1, \\ u_0 &= \psi. \end{aligned} \right\} \quad (9)$$

Будем предполагать функцию  $G(x, u)$  двух аргументов и функцию  $\varphi(x)$  такими, что существует решение  $u(x)$ , имеющее ограниченную вторую производную. Кроме того, будем считать, что  $G(x, u)$  имеет ограниченную производную по  $u$

$$\left| \frac{\partial G}{\partial u} \right| < M. \quad (10)$$

Читателю рекомендуется проверить, что разностная схема (8) аппроксимирует задачу (9) на решении  $u(x)$  с первым относительно  $h$  порядком. (Разностное уравнение соответствует задаче с первым порядком, а граничное условие  $u_0 = \psi$  — точно.) Определим нормы

$$\| u^{(h)} \|_{U_h} = \max_n |u_n|, \quad \| f^{(h)} \|_{F_h} = \max \{ |\psi|, \max_m |\varphi(x_m)| \}$$

и займемся проверкой устойчивости разностной схемы (8). Запишем ее в форме (2), положив

$$L_h u^{(h)} = \begin{cases} \frac{u_{n+1} - u_n}{h} - G(x_n, u_n), & n = 0, 1, \dots, N-1, \\ u_0, \end{cases}$$

$$f^{(h)} = \begin{cases} \varphi(x_n), & n = 0, 1, \dots, N-1, \\ \psi. \end{cases}$$

Задача

$$L_h z^{(h)} = f^{(h)} + \varepsilon^{(h)}$$



означающая устойчивость с константой  $C = 2e^M$ . В силу теоремы разностная схема (8) является сходящейся с первым относительно  $h$  порядком.

Исследуем теперь сходимость разностной схемы (7) § 10

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} - (1 + x_n^2)u_n &= \sqrt{1 + x_n}, \\ n &= 1, 2, \dots, N-1, \\ u_0 &= 2, \\ u_N &= 1 \end{aligned} \right\} \quad (13)$$

для дифференциальной краевой задачи (4) § 10. Аппроксимация со вторым относительно  $h$  порядком задачи (4) § 10 задачей (13) благодаря формуле

$$\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = u''(x) + \frac{h^2}{12} u^{(4)}(\xi)$$

здесь очевидна \*). Займемся проверкой устойчивости. Рассматриваемая задача линейна. Поэтому проверка устойчивости состоит в том, чтобы установить существование единственного решения задачи

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} - (1 + x_n^2)u_n &= g_n, \quad n = 1, 2, \dots, N-1, \\ u_0 &= \alpha, \\ u_N &= \beta \end{aligned} \right\} \quad (14)$$

при любых  $\{g_n\}$ ,  $\alpha$  и  $\beta$ , и в том, чтобы получить оценку

$$\max_n |u_n| \leq C (\max_n |g_n|, |\alpha|, |\beta|). \quad (15)$$

Задачу вида (14) мы рассматривали в § 4 (см. стр. 40). Там для задачи вида

$$\left. \begin{aligned} a_n u_{n-1} + b_n u_n - c_n u_{n+1} &= g_n, \\ u_0 &= \alpha, \\ u_N &= \beta \end{aligned} \right\}$$

в предположении

$$|b_n| \geq |a_n| + |c_n| + \delta, \quad \delta > 0$$

\*) Заметим, что если бы речь шла о решении уравнения

$$u'' - (1 + x^2)u = \sqrt{x},$$

только незначительно отличающегося от изучаемого здесь, то мы не могли бы сделать вывода об аппроксимации, так как  $|u'''(x)|$  был бы в данном случае неограничен (докажите аккуратно неограниченность  $u'''(x)$ ).

была доказана ее однозначная разрешимость и оценка

$$|u_n| < \max \left[ |\alpha|, |\beta|, \frac{1}{\delta} \max_m |g_m| \right]. \quad (16)$$

В случае задачи (14)

$$a_n = \frac{1}{h^2}, \quad c_n = \frac{1}{h^2}, \quad |b_n| = \frac{2}{h^2} + 1 + x_n^2 > |a_n| + |c_n| + 1.$$

Поэтому оценка (16) влечет за собой оценку (15) с постоянной  $C = 1$ . Устойчивость доказана.

Отметим одно обстоятельство, которое может оказаться полезным при доказательстве сходимости путем проверки аппроксимации и устойчивости.

Пусть разностная схема (2) разбита на две подсистемы

$$l_h^{(0)} u^{(h)} = f_0^{(h)}, \quad (17)$$

$$l_h^{(1)} u^{(h)} = f_1^{(h)} \quad (17')$$

так, что

$$L_h u^{(h)} \equiv \begin{cases} l_h^{(0)} u^{(h)}, \\ l_h^{(1)} u^{(h)}, \end{cases} \quad f^{(h)} = \begin{cases} f_0^{(h)}, \\ f_1^{(h)}, \end{cases} \quad \delta f^{(h)} = \begin{cases} \delta f_0^{(h)}, \\ \delta f_1^{(h)}. \end{cases}$$

Пусть, далее, разностная схема (2) аппроксимирует задачу (1) с порядком  $h^k$ , т. е. выполнено условие (3). Пусть, сверх того, подсистема (17') соответствует задаче (1) на решении  $u$  точно, т. е.  $\delta f_1^{(h)} = 0 \in F_h^{(1)}$ :

$$\delta f^{(h)} = \begin{cases} \delta f_0^{(h)}, \\ 0. \end{cases} \quad (18)$$

В таком случае для сходимости решения  $u^{(h)}$  задачи (2) к искомой сеточной функции  $[u]_h$ , т. е. для справедливости оценки (7), достаточно, чтобы оценка (5) выполнялась не при произвольных  $\varepsilon^{(h)} \in F_h$ , но лишь для всех  $\varepsilon^{(h)}$  вида

$$\varepsilon^{(h)} = \begin{cases} \varepsilon_0^{(h)}, \\ 0, \end{cases} \quad (19)$$

где  $0 \in F_h^{(1)}$ .

Доказательство дословно совпадает с проведенным выше доказательством теоремы о сходимости.

Читатель легко проверит, что в случае линейного оператора  $L_h$  требование, чтобы оценка (5) имела место лишь для всех



$\varepsilon^{(h)}$  вида (19), выполнено одновременно с требованием, чтобы оценка (6) выполнялась для всех  $f^{(h)}$  того же специального вида

$$f^{(h)} = \begin{cases} f_0^{(h)}, \\ 0, \end{cases}$$

где  $0 \in F_h^{(1)}$ .

Например, при доказательстве сходимости разностной схемы (13) можно было воспользоваться тем, что оба граничных условия

$$l_h^{(1)} u^{(h)} \equiv \begin{cases} u_0 = 2, \\ u_N = 1 \end{cases} \equiv f_h^{(1)}$$

при подстановке в них таблицы решения  $[u]_h$  задачи (4) из § 10 выполняются точно:

$$l_h^{(1)} [u]_h = \begin{cases} u(0) = 2, \\ u(1) = 1. \end{cases}$$

Поэтому проверку неравенства (15), означающего устойчивость разностной схемы (13), можно было провести не для произвольной правой части

$$f^{(h)} = \begin{cases} g_n, & n = 1, 2, \dots, N-1, \\ \alpha, \\ \beta, \end{cases}$$

а только для правых частей вида

$$f^{(h)} = \begin{cases} g_n, & n = 1, 2, \dots, N-1, \\ 0, \\ 0, \end{cases}$$

когда  $\alpha = 0$  и  $\beta = 0$ .

В задаче (13) мы справились с проверкой неравенства, означающего устойчивость, и без учета этого упрощающего обстоятельства. В более сложных задачах (для уравнений с частными производными) указанное соображение будет иногда полезно.

В заключение параграфа подчеркнем, что схема доказательства сходимости решения задачи  $L_h u^{(h)} = f^{(h)}$  к решению задачи  $Lu = f$  путем проверки аппроксимации и устойчивости носит общий характер. Под  $Lu = f$  можно понимать любое функциональное уравнение, а не только краевую задачу для обыкновенного дифференциального уравнения. Само по себе неважно, решением какой задачи является функция  $u$ . Уравнение  $Lu = f$  используется только для конструирования разностного уравнения  $L_h u^{(h)} = f^{(h)}$ . Поясним эту мысль в п. 3.

**3. Сходящаяся разностная схема для интегрального уравнения.** Построим и исследуем разностную схему для вычисления решения интегрального уравнения

$$Lu \equiv u(x) - \int_0^1 K(x, y) u(y) dy = f(x).$$

Будем предполагать, что  $|K(x, y)| < \rho < 1$ .

Зададим  $N$ , положим  $h = 1/N$  и будем искать таблицу  $[u]_h$  значений решения на сетке  $x_n = nh$ ,  $n = 0, 1, \dots, N$ . Для получения разностной схемы мы в равенстве

$$u(x_n) - \int_0^1 K(x_n, y) u(y) dy = f(x_n), \quad n = 0, 1, \dots, N.$$

приблизительно заменим интеграл суммой, пользуясь квадратурной формулой трапеций. Напомним эту формулу: для произвольной дважды дифференцируемой на отрезке  $0 \leq y \leq 1$  функции  $\varphi(y)$  справедливо приближенное равенство

$$\int_0^1 \varphi(y) dy \approx h \left( \frac{\varphi_0}{2} + \varphi_1 + \varphi_2 + \dots + \varphi_{N-1} + \frac{\varphi_N}{2} \right), \quad h = \frac{1}{N},$$

причем погрешность есть величина  $O(h^2)$ . После указанной замены интеграла получим

$$u_n^{(h)} - h \left( \frac{K(x_n, 0)}{2} u_0^{(h)} + K(x_n, h) u_1^{(h)} + \dots \right. \\ \left. \dots + K(x_n, y_{N-1}) u_{N-1}^{(h)} + \frac{K(x_n, 1)}{2} u_N^{(h)} \right) = f(x_n), \quad n = 0, 1, \dots, N. \quad (20)$$

Выписанная система равенств записывается в форме  $L_h u^{(h)} = f^{(h)}$ , если положить

$$L_h u^{(h)} = \begin{cases} g_0, \\ g_1, \\ \dots \\ g_N, \end{cases} \quad f^{(h)} = \begin{cases} f(0), \\ f(h), \\ \dots \\ f(1), \end{cases}$$

где

$$g_n = u_n^{(h)} - h \left[ \frac{K(x_n, 0)}{2} u_0^{(h)} + K(x_n, h) u_1^{(h)} + \dots + \frac{K(x_n, 1)}{2} u_N^{(h)} \right], \\ n = 0, 1, \dots, N.$$

Построенная разностная схема  $L_h u^{(h)} = f^{(h)}$  аппроксимирует задачу  $Lu = f$  на решении  $u$  со вторым порядком относительно шага  $h$ , поскольку квадратурная формула трапеций имеет второй порядок точности. Проверим устойчивость. Пусть  $u^{(h)} = (u_0, u_1, \dots, u_N)$  — какое-нибудь решение системы (20), и пусть  $u_s$  — одна из тех компонент решения, которые по модулю не меньше каждой из остальных:

$$|u_s| = \max_m |u_m|.$$

Из уравнения с номером  $n = s$  системы (20) следует

$$\begin{aligned} |f(x_s)| &= \left| u_s^{(h)} - h \left( \frac{K(x_s, 0)}{2} u_0^{(h)} + K(x_s, h) u_1^{(h)} + \dots + \frac{K(x_s, 1)}{2} u_N^{(h)} \right) \right| \geq \\ &\geq |u_1^{(h)}| + h \left( \frac{\rho}{2} + \rho + \dots + \rho + \frac{\rho}{2} \right) |u_s^{(h)}| = (1 - Nh\rho) |u_s^{(h)}| = (1 - \rho) |u_s^{(h)}|. \end{aligned}$$

Поэтому

$$\|u^{(h)}\|_{U_h} = \max_m |u_m^{(h)}| = |u_s^{(h)}| \leq \frac{1}{1-\rho} |f(x_s)| \leq \frac{1}{1-\rho} \|j^{(h)}\|_{F_h}. \quad (21)$$

В частности, при  $f(x_n) \equiv 0$  отсюда следует, что система (20) не имеет нетривиальных решений, а следовательно, однозначно разрешима при любой правой части  $f^{(h)}$ . Неравенство (21) означает устойчивость (6) с постоянной  $C = 1/(1-\rho)$ . Решение  $u^{(h)}$  задачи  $L_h u^{(h)} = f^{(h)}$  в силу теоремы о сходимости удовлетворяет неравенству

$$\|[u]_h - u^{(h)}\|_{U_h} = \max_m |u(mh) - u_m^{(h)}| \leq Ah^2,$$

где  $A$  — некоторая постоянная.

### § 13. О выборе норм

Понятия сходимости, аппроксимации и устойчивости, введенные в §§ 10—12, имеют смысл, если тем или иным способом введены нормы в пространствах  $U_h$  и  $F_h$ , которым принадлежат соответственно решение  $u^{(h)}$  и правая часть  $f^{(h)}$  разностной схемы

$$L_h u^{(h)} = f^{(h)}$$

для приближенного вычисления решения  $u$  дифференциальной краевой задачи  $Lu = f$ .

Обсудим вопрос о степени произвола, с какой можно выбирать нормы в пространствах  $U_h$  и  $F_h$ . Начнем с нормы  $\|\cdot\|_{U_h}$ , по величине которой оценивается уклонение приближенного решения  $u^{(h)}$  от сеточной функции  $[u]_h$ , т. е. от таблицы значений решения  $u$ . Во всех рассмотренных примерах мы пользовались нормой, определенной равенством

$$\|z^{(h)}\|_{U_h} = \max_k |z_k^{(h)}|. \quad (1)$$

Максимум берется по всем точкам сетки  $D_h$ , на которой определена сеточная функция  $z^{(h)} \in U_h$ . Можно было бы, конечно, положить

$$\|z^{(h)}\|_{U_h} = h \max_k |z_k^{(h)}|, \quad (2)$$

или

$$\|z^{(h)}\|_{U_h} = \frac{1}{h} \max_k |z_k^{(h)}|, \quad (3)$$

или даже

$$\|z^{(h)}\|_{U_h} = 2^{-1/h} \max_k |z_k^{(h)}|.$$

Последняя норма может показаться удобной, так как в ней оказывается сходящейся разностная схема

$$\left. \begin{aligned} 4 \frac{u_{n+1} - u_{n-1}}{2h} - 3 \frac{u_{n+1} - u_n}{h} + Au_n &= 0, \\ n &= 0, 1, \dots, N-1, \\ u_0 &= a, \\ u_1 &= ae^{-Ah} \end{aligned} \right\}$$

для решения задачи

$$\left. \begin{aligned} u' + Au &= 0, \quad 0 \leq x \leq 1, \\ u(0) &= a, \end{aligned} \right\}$$

построенная в § 9 как пример непригодной схемы. Действительно, в силу равенства

$$u(nh) - u_n = ae^{-Ax_n} - u_n = -\left[\frac{3}{2} A^2 ae^{Ax_n} + O(h)\right] h^2 2^{x_n/h} + O(h),$$

вытекающего из соотношения (7) § 9, величина

$$\|[u]_h - u^{(h)}\|_{U_h} = 2^{-1/h} \max_m |u(mh) - u_m^{(h)}|$$

стремится к нулю при измельчении сетки. Но ясно, что стремление этой величины к нулю ни в каком разумном смысле не означает стремления к нулю погрешности  $z^{(h)} = [u]_h - u^{(h)}$ , поскольку при этом разности  $u(nh) - u_n$  могут стремительно (почти как  $2^{1/h}$ ) возрастать, что и имеет место в рассматриваемом примере.

Нормы (2) и (3) также не стоит рекомендовать, так как они недостаточно характеризуют погрешности  $[u]_h - u^{(h)}$ .

Обычно принято выбирать норму в пространстве  $U_h$  так, чтобы при стремлении шага  $h$  к нулю она переходила в ту или иную норму функций, заданных на всем отрезке, т. е. чтобы выполнялось равенство

$$\lim_{h \rightarrow 0} \|[u]_h\|_{U_h} = \|u\|_U, \quad (4)$$

где  $\|\cdot\|_U$  — норма в пространстве функций на отрезке, которому принадлежит решение  $u(x)$ . Норма

$$\|z^{(h)}\|_{U_h} = \max_m |z_m^{(h)}|$$

этому условию удовлетворяет, если в качестве  $U$  рассматривать пространство непрерывных функций, в котором

$$\|u\|_U = \max_{0 \leq x \leq 1} |u(x)|,$$

а сеточную функцию  $[u]_h$  определять как совпадающую с  $u(x)$  в точках сетки.

Норма

$$\|z^{(h)}\|_{U_h} = \sqrt{h \sum_m |z_m^{(h)}|^2} \quad (5)$$

также является разумной. Она удовлетворяет условию (4), если за  $U$  принять пространство непрерывных функций с нормой

$$\|u\|_U = \sqrt{\int_0^1 u^2(x) dx},$$

а сеточную функцию  $[u]_h$  по-прежнему определить как совпадающую с  $u(x)$  в точках сетки.

В случае разрывного решения  $u(x)$ , обладающего, однако, интегрируемым квадратом, за  $U$  можно принять пространство функций с интегрируемым квадратом и с нормой

$$\|u\|_U = \sqrt{\int_0^1 u^2(x) dx},$$

но значение  $u_n$  сеточной функции  $[u]_h$  определять не по формуле  $u_n = u(nh)$ , которая может не иметь смысла, а по формуле

$$u_n = \frac{1}{h} \int_{x_n - h/2}^{x_n + h/2} u(x) dx.$$

Тогда и для разрывной функции  $u(x)$  будет

$$\lim_{h \rightarrow 0} \|[u]_h\|_{U_h} = \sqrt{\int_0^1 u^2 dx}.$$

Ясно, что сходимость

$$\lim_{h \rightarrow 0} \|[u]_h - u^{(h)}\|_{U_h} = 0$$

в смысле нормы (1), т. е. равномерная сходимость, влечет за собою сходимость в смысле нормы (5), т. е. сходимость в среднем, но из сходимости в среднем не следует равномерная сходимость. Поэтому из числа разумных норм, удовлетворяющих условию (4), выбирают ту, в которой удается доказать

сходимость изучаемой конкретной разностной схемы. Для этого выбора нет общего рецепта.

В случае обыкновенных дифференциальных и соответствующих разностных уравнений, которыми мы занимаемся в этой главе, обычно достаточно удобны нормы (1), (5) или норма типа

$$\|z^{(h)}\|_{U_h} = \max \left[ \max_m |z_m^{(h)}|, \max_m \frac{|z_{m+1}^{(h)} - z_m^{(h)}|}{h} \right], \quad (6)$$

в которой учтена скорость изменения сеточной функции при переходе от точки к точке. Равенство (4) при этой норме выполняется, если за  $U$  принять пространство непрерывно дифференцируемых функций с нормой

$$\|u\|_U = \max \left[ \max_{0 \leq x \leq 1} |u(x)|, \max_{0 \leq x \leq 1} |u'(x)| \right].$$

В случае уравнений с частными производными и соответствующих разностных схем иногда удобно пользоваться довольно замысловатыми нормами, приспособленными для конкретных задач.

Перейдем к вопросу о выборе нормы в пространстве  $F_h$ , которому принадлежит правая часть разностной схемы  $L_h u^{(h)} = f^{(h)}$ . Подчеркнем, что сходимость разностной схемы  $\| [u]_h - u^{(h)} \|_{U_h} \rightarrow 0$  при избранной норме  $\| \cdot \|_{U_h}$  не зависит от того, как выбрана норма  $\| \cdot \|_{F_h}$  и выбрана ли эта норма вообще. Считать  $F_h$  линейным нормированным пространством нам приходится только для того, чтобы свести доказательство сходимости и проверку порядка точности разностной схемы к проверке аппроксимации с некоторым порядком и проверке устойчивости.

Обсуждение вопроса о выборе нормы в  $F_h$  проведем в предположении линейности разностной схемы  $L_h u^{(h)} = f^{(h)}$ . Это делается лишь для того чтобы избежать громоздкости, не вызванной существом дела.

Пусть при каком-нибудь фиксированном выборе нормы  $\| \cdot \|^{(1)}$  разностная схема  $L_h u^{(h)} = f^{(h)}$  аппроксимирует задачу  $Lu = f$  на решении  $u$  с некоторым порядком  $h^k$  и устойчива. Тогда в силу теоремы о сходимости разностная схема  $L_h u^{(h)} = f^{(h)}$  является сходящейся и имеет порядок точности  $h^k$ :

$$\| [u]_h - u^{(h)} \|_{U_h} < Ch^k. \quad (7)$$

Аппроксимация, напомним, означает выполнение неравенства вида

$$\| L_h [u]_h - f^{(h)} \|_{F_h}^{(1)} \leq C_1 h^k. \quad (8)$$

Устойчивость означает, что задача  $L_h u^{(h)} = f^{(h)}$ , однозначно разрешима при любом  $f^{(h)} \in F_h$ , причем

$$\|u^{(h)}\|_{U_h} \leq C_2 \|f^{(h)}\|_{F_h}^{(1)}. \quad (9)$$

Если выбрать другую норму  $\|\cdot\|_{F_h}$  положив

$$\|f^{(h)}\|_{F_h}^{(2)} = h \|f^{(h)}\|_{F_h}^{(1)}, \quad (10)$$

то, очевидно, неравенства (8) и (9) заменятся соответственно неравенствами

$$\left. \begin{aligned} \|L_h [u]_h - f^{(h)}\|_{F_h}^{(2)} &\leq Ch^{k+1}, \\ \|u^{(h)}\|_{U_h} &\leq \frac{C}{h} \|f^{(h)}\|_{F_h}^{(2)}. \end{aligned} \right\} \quad (11)$$

Таким образом, аппроксимация будет уже не порядка  $k$  относительно шага  $h$ , а на единицу более высокого порядка  $k+1$ . Судя по этому, можно было бы ошибочно заключить, что порядок точности разностной схемы не  $h^k$ , а  $h^{k+1}$ . Дело в том, что неравенство (9) уже не означает устойчивости, которая при новом выборе нормы, вообще говоря, теряется.

Если бы мы вместо (10) ввели норму  $\|\cdot\|_{F_h}^{(2)}$  равенством

$$\|f^{(h)}\|_{F_h}^{(2)} = \frac{1}{h} \|f^{(h)}\|_{F_h}^{(1)},$$

то вместо (8) и (9) получили бы соответственно

$$\|L_h [u]_h - f^{(h)}\|_{F_h}^{(2)} \leq C_1 h^{k-1}, \quad (12)$$

$$\|u^{(h)}\|_{U_h} \leq C_2 h \|f^{(h)}\|_{F_h}^{(2)}. \quad (13)$$

Неравенство (13) гарантирует устойчивость, так как  $C_2 h$  можно заменить не зависящей от  $h$  постоянной  $C_2$ , лишь усилив неравенство. Неравенство (12) означает аппроксимацию порядка  $k-1$  относительно шага  $h$ .

Таким образом, при сделанном выборе нормы  $\|\cdot\|_{F_h}^{(2)}$  мы могли бы на основании теоремы о сходимости гарантировать лишь  $(k-1)$ -й порядок точности разностной схемы  $L_h u^{(h)} = f^{(h)}$ , на единицу более низкий, чем гарантированный неравенством (7). Утеря информации о порядке точности произошла из-за неудачного выбора нормы в пространстве  $F_h$ .

Чтобы правильно выявить порядок точности разностной схемы, надо так выбрать норму  $\|\cdot\|_{F_h}$ , чтобы порядок аппроксимации оказался как можно более высоким, но устойчивость при этом еще не утерялась. Для такого выбора нормы  $\|\cdot\|_{F_h}$  нет общего правила\*). Более того, не всегда можно выбрать норму так, чтобы имела место и аппроксимация и устойчивость, иначе, вопреки примеру из § 9, всякая разностная схема была бы сходящейся.

Приведем, однако, одно соображение общего характера, способствующее правильному выбору нормы в линейном пространстве  $F_h$ . При выборе нормы  $\|\cdot\|_{F_h}$  надо учитывать характер непрерывной зависимости решения дифференциальной краевой задачи  $Lu = f$ , на основе которой построена разностная схема  $L_h u^{(h)} = f^{(h)}$ , от правой части  $f$ .

Например, в случае задачи

$$\frac{du}{dx} + Au = \varphi(x), \quad u(0) = a, \quad 0 \leq x \leq 1$$

при внесении изменений  $\delta\varphi(x)$  и  $\delta a$  в правые части уравнения и граничного условия соответственно решение  $u(x)$  изменяется на величину  $\delta u(x)$  того же порядка.

Рассмотрим теперь разностную схему

$$L_h u^{(h)} = \begin{cases} \frac{u_{n+1} - u_n}{h} + Au_n = \varphi(x_n), & n = 0, 1, \dots, N-1, \\ u_0 = a, \end{cases}$$

так что

$$f^{(h)} = \begin{cases} \varphi(x_n), & n = 0, 1, \dots, N-1, \\ a. \end{cases}$$

Норму в  $U_h$ , как обычно, зададим равенством

$$\|u^{(h)}\|_{U_h} = \max_m |u_m^{(h)}|.$$

Устойчивости можно ожидать только в том случае, если норма

$$\|f^{(h)}\|_{F_h} = \left\| \begin{matrix} \varphi(x_n) \\ a \end{matrix} \right\|_{F_h}$$

существенно зависит и от  $\varphi(x_n)$  и от  $a$ . Например, она может иметь вид

$$\|f^{(h)}\|_{F_h} = \max [ |a|, \max_m |\varphi_m| ]. \quad (14)$$

\*) Мы имеем в виду и случай разностных схем для уравнений с частными производными.



Устойчивость в этой норме доказана в § 12, где рассмотрена более общая нелинейная задача.

Нельзя ожидать устойчивости, если норма выбрана, скажем, по формуле

$$\|f^{(h)}\|_{F_h} = \max [h|a|, \max_m |\varphi_m|],$$

куда  $a$  входит по мере уменьшения  $h$  со все более малым весом.

Устойчивость в смысле этой нормы означала бы более слабую зависимость решения  $u^{(h)}$  от  $a$ , чем зависимость от  $a$  решения  $u$  дифференциального уравнения. Между тем, при малом  $h$  в силу сходимости (сходимость имела бы место в случае устойчивости, поскольку аппроксимация тоже есть) решение разностного уравнения мало отличается от решения дифференциального уравнения и при изменении начального значения  $a$  должно меняться примерно так, как меняется решение  $u(x)$ .

Более четко: при сделанном выборе нормы задача

$$\left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + Au_n &= \varphi_n, \quad n = 0, 1, \dots, N-1, \\ u_0 &= 0 \end{aligned} \right\}$$

аппроксимирует задачу

$$\frac{du}{dx} + Au = \varphi(x), \quad u(0) = a$$

на решении  $u(x)$  при любом  $a$ . Значит, в случае устойчивости функция  $u^{(h)}$ , не зависящая от  $a$ , должна была бы сходиться к решению  $u(x)$ , каково бы ни было заданное  $a$ . Но  $u^{(h)}$  не может сходиться одновременно к разным функциям  $u(x)$ .

В случае разностной схемы

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + A \frac{u_{n+1} - u_{n-1}}{2h} + Bu_n &= \varphi_n, \\ n &= 1, \dots, N, \\ u_0 &= a, \\ \frac{u_1 - u_0}{h} &= b. \end{aligned} \right\} \quad (15)$$

для задачи

$$\left. \begin{aligned} \frac{d^2u}{dx^2} + A \frac{du}{dx} + Bu &= \varphi(x), \\ u(0) &= a, \\ \frac{du(0)}{dx} &= b \end{aligned} \right\}$$

из тех же соображений норма

$$\|f^{(h)}\|_{F_h} = \left\| \begin{array}{c} \varphi(x_m) \\ a \\ b \end{array} \right\|_{F_h}$$

должна существенно зависеть от  $\varphi$ ,  $a$  и  $b$ . Она может иметь вид

$$\|f^{(h)}\|_{F_h} = \max[|a|, |b|, \max_m |\varphi_m|], \quad (16)$$

но нельзя ожидать устойчивости при выборе в качестве нормы  $\|f^{(h)}\|_{F_h}$ , скажем, величины

$$\|f^{(h)}\|_{F_h} = \max[|a|, h|b|, \max_m |\varphi_m|].$$

Преобразуем схему (15) к несколько иному виду:

$$L_h u^{(h)} = \begin{cases} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + A \frac{u_{n+1} - u_{n-1}}{2h} + Bu_n = \varphi(x_n), \\ u_0 = a, \\ u_1 = a + bh, \end{cases} \quad (17)$$

так что

$$f^{(h)} = \begin{cases} \varphi(x_n), & n = 1, 2, \dots, N-1, \\ a, \\ a + bh. \end{cases}$$

Норму в  $F_h$  теперь следует ввести, определив ее для произволь-

ного элемента  $g^{(h)} = \begin{cases} \psi_m, \\ \alpha, \\ \beta \end{cases}$  по формуле типа

$$\|g^{(h)}\|_{F_h} = \max\left[|\alpha|, \frac{|\beta - \alpha|}{h}, \max_m |\psi_m|\right], \quad (18)$$

куда  $|\beta - \alpha|$  входит с возрастающим при  $h \rightarrow 0$  весом  $1/h$ . Действительно, изменение  $\alpha$  или  $\beta$  на величину  $h$  равносильно изменению  $u_0$  или  $u_1$  на величину  $h$ , но при этом  $\frac{|u_1 - u_0|}{h}$  изменится на величину порядка 1. Последнее, если схема устойчива, повлечет за собой изменение решения уравнения

$$\frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + A \frac{u_{n+1} - u_{n-1}}{2h} + Bu_n = \varphi_n$$

на величину порядка 1, так как изменение  $\frac{|u_1 - u_0|}{h}$  на величину  $O(1)$  аналогично изменению правой части условия  $\frac{du(0)}{dx} = b$  дифференциальной задачи на величину порядка 1.

Нельзя ожидать устойчивости определив норму по формуле

$$\|g^{(h)}\|_{F_h} = \max [|\alpha|, |\beta|, \max_m |\psi_m|],$$

т. е. так, как она была определена выше, когда мы пользовались пространством  $F_h$  для оснастки разностной схемы (15). Порядок аппроксимации, которым обладают схемы (15) и (17) при нормах (16) и (18) соответственно, для обеих схем одинаков — первый относительно  $h$ . Устойчивость схем (15) и (17) при нормах (16) и (18) будет доказана в § 14.

#### § 14. Достаточный признак устойчивости разностных схем решения задачи Коши

В этом параграфе мы покажем, как провести исследование устойчивости разностных схем  $L_h u^{(h)} = f^{(h)}$  решения дифференциальной задачи с начальными условиями (задачи Коши). Мы сделаем это с помощью рассмотрения характерных примеров разностных схем, приближающих задачи

$$Lu = \begin{cases} \frac{du}{dx} + Au = \varphi(x), & 0 \leq x \leq 1, \\ u(0) = a, \end{cases} \quad (1)$$

$$Lu \equiv L \begin{pmatrix} v \\ w \end{pmatrix} = \begin{cases} \frac{dv}{dx} + Av + Bw = p(x), & 0 \leq x \leq 1, \\ \frac{dw}{dx} + Cv + Dw = q(x), & 0 \leq x \leq 1, \\ v(0) = a, \\ w(0) = b, \end{cases} \quad (2)$$

$$Lu \equiv \begin{cases} \frac{d^2u}{dx^2} + A \frac{du}{dx} + Bu = \varphi(x), & 0 \leq x \leq 1, \\ u(0) = a, \\ \frac{du(0)}{dx} = b. \end{cases} \quad (3)$$

Чтобы понятие устойчивости разностной схемы  $L_h u^{(h)} = f^{(h)}$  имело смысл, должны быть определены линейные нормированные пространства  $U_h$  и  $F_h$ . Этим пространствам принадлежат подлежащая приближенному вычислению таблица  $[u]_h$  искомого



Воспользуемся теперь тем, что выражение  $(1 - Ah)^n$  ограничено для  $n \leq N = 1/h$ ,

$$|(1 - Ah)^n| < C_1. \quad (9)$$

С помощью неравенства (9) из формулы (6) для  $u_n$  заключаем:

$$\begin{aligned} |u_n| &\leq C_1 |u_0| + hNC_1 \max_m |\varphi_m| = \\ &= C_1 |a| + C_1 \max_m |\varphi_m| \leq 2C_1 \|f^{(n)}\|_{F_h}. \end{aligned} \quad (10)$$

Ввиду произвольности  $n = 0, 1, \dots, N$  отсюда следует

$$\|u^{(h)}\|_{U_h} \leq 2C_1 \|f^{(h)}\|_{F_h}, \quad (11)$$

и устойчивость доказана.

**2. Каноническая запись разностной схемы.** Введем новые обозначения, положив

$$u_n = y_n, \quad R_h = (1 - Ah), \quad \rho_n = \varphi_n. \quad (12)$$

Тогда равенства (5) запишутся в виде

$$\left. \begin{aligned} y_{n+1} &= R_h y_n + h\rho_n, \\ y_0 &\text{ задано.} \end{aligned} \right\} \quad (13)$$

Пользуясь обозначениями (12), повторим все сделанные выше выкладки. Равенства (6) примут вид

$$\left. \begin{aligned} y_1 &= R_h y_0 + h\rho_0, \\ y_2 &= R_h^2 y_0 + h[R_h \rho_0 + \rho_1], \\ y_3 &= R_h^3 y_0 + h[R_h^2 \rho_0 + R_h \rho_1 + \rho_2], \\ &\dots \\ y_n &= R_h^n y_0 + h[R_h^{n-1} \rho_0 + R_h^{n-2} \rho_1 + \dots + \rho_{n-1}]. \end{aligned} \right\} \quad (6')$$

Отсюда

$$\max_n |y_n| \leq \max_n |R_h^n| \cdot [|y_0| + hN \max_n |\rho_n|].$$

Нормы  $\|\cdot\|_{U_h}$  и  $\|\cdot\|_{F_h}$  теперь запишутся равенствами

$$\|u^{(h)}\|_{U_h} = \max_n |y_n|, \quad (7')$$

$$\|f^{(h)}\|_{F_h} = \max [|y_0|, \max_n |\rho_n|]. \quad (8')$$

Тогда, учитывая  $Nh = 1$ , можно написать

$$\|u^{(h)}\|_{U_h} \leq \max_n |R_h^n| \cdot 2 \|f^{(h)}\|_{F_h}.$$

Доказательство устойчивости будет завершено, если будет установлена равномерная относительно  $h$  ограниченность совокупности чисел  $|R_h^n|$ , т. е. оценка

$$|R_h^n| \leq C, \quad n = 1, 2, \dots, N. \quad (9')$$

Но

$$|R_h^n| \leq (1 - Ah)^n \leq (1 + |A|h)^N \leq e^{|A|} = C,$$

что завершает доказательство.

Запись разностной схемы в форме (13) позволила свести доказательство устойчивости к получению оценки для  $|R_h^n|$ . Это удобно. Мы и все другие разностные схемы решения задач с начальными условиями будем приводить к каноническому виду (13), понимая под  $y_n$ ,  $\rho_n$  и  $R_n$  различные выражения, в каждом примере свои.

Запишем, например, в форме (13) разностную схему

$$L_h u^{(h)} \equiv \begin{cases} \frac{v_{n+1} - v_n}{h} + Av_n + Bw_n = p_n, & n = 0, 1, \dots, N-1, \\ \frac{w_{n+1} - w_n}{h} + Cv_n + Dw_n = q_n, & n = 0, 1, \dots, N-1, \\ v_0 = a, \\ w_0 = b, \end{cases} \quad (14)$$

приближающую задачу Коши (2) для системы дифференциальных уравнений. Здесь

$$f^{(h)} = \begin{cases} p_n, & n = 0, 1, \dots, N-1, \\ q_n, & n = 0, 1, \dots, N-1, \\ a, \\ b, \end{cases}$$

Запишем разностную схему (14) в форме

$$L_h u^{(h)} \equiv \begin{cases} \frac{\begin{bmatrix} v_{n+1} \\ w_{n+1} \end{bmatrix} - \begin{bmatrix} v_n \\ w_n \end{bmatrix}}{h} + \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{bmatrix} v_n \\ w_n \end{bmatrix} = \begin{bmatrix} p_n \\ q_n \end{bmatrix}, & n = 0, 1, \dots, N-1, \\ \begin{bmatrix} v_0 \\ w_0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}, \end{cases}$$

где  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$  — матрица второго порядка. Придадим этому векторному разностному уравнению вид рекуррентного соотношения:

$$\left. \begin{aligned} \begin{bmatrix} v_{n+1} \\ w_{n+1} \end{bmatrix} &= \begin{pmatrix} 1 - Ah & -Bh \\ -Ch & 1 - Dh \end{pmatrix} \begin{bmatrix} v_n \\ w_n \end{bmatrix} + h \begin{bmatrix} p_n \\ q_n \end{bmatrix}, \\ \begin{bmatrix} v_0 \\ w_0 \end{bmatrix} &= \begin{bmatrix} a \\ b \end{bmatrix}. \end{aligned} \right\}$$

Если положить

$$y_n = u_n = \begin{bmatrix} v_n \\ w_n \end{bmatrix}, \quad R_h = \begin{pmatrix} 1 - Ah & -Bh \\ -Ch & 1 - Dh \end{pmatrix}, \quad \rho_n = \begin{bmatrix} p_n \\ q_n \end{bmatrix},$$

то последняя запись приобретет требуемый вид (13).

**3. Устойчивость как ограниченность норм степеней оператора перехода.** Сделаем замечание, которое одинаково применимо к уравнениям вида (13) независимо от размерности линейного пространства  $Y$ , которому принадлежат векторы  $y_n$  и  $\rho_n$ , и от вида линейного оператора  $R_h$ : из записи (13) следует запись (6').

Если в пространстве  $Y$ , которому принадлежат  $\rho_n$  и  $y_n$ , введена какая-либо норма  $\|\cdot\|_Y$ , то из равенств (6') вытекает оценка

$$\|y_n\|_Y \leq \|R_h^n\|_Y \cdot \|y_0\|_Y + h [\|R_h^{n-1}\|_Y \cdot \|\rho_0\|_Y + \dots + \|\rho_{n-1}\|_Y]. \quad (15)$$

Напомним, что *нормой*  $\|T\|$  линейного оператора  $T$ , отображающего какое-либо линейное нормированное пространство  $Y$  в себя, называется число

$$\|T\| = \sup_{x \in Y} \frac{\|Tx\|}{\|x\|} \quad \text{или} \quad \|T\| = \sup_{\|x\|=1, x \in Y} \|Tx\|.$$

Отсюда и из свойств нормы векторов следует:

$$\begin{aligned} \|Tx\| &\leq \|T\| \cdot \|x\|, \\ \|\lambda T\| &= |\lambda| \cdot \|T\|, \quad \text{где } \lambda - \text{любое число,} \\ \|T^m\| &\leq \|T\|^m. \end{aligned}$$

Первые два из этих свойств использованы для получения оценки (15).

Из (15), очевидно, вытекает

$$\max_{0 \leq n \leq N} \|y_n\|_Y \leq \max_{0 \leq n \leq N} \|R_h^n\|_Y [\|y_0\|_Y + Nh \max_{0 \leq n \leq N} \|\rho_n\|_Y]. \quad (16)$$

Пусть разностная схема  $L_h u^{(h)} = f^{(h)}$  приведена к каноническому виду (13), и пусть нормы, введенные в пространствах  $U_h$ ,

$F_h$  и  $Y$ , подобраны так, что выполнены неравенства

$$\left. \begin{aligned} \|u^{(h)}\|_{U_h} &\leq C_2 \max_{0 \leq n \leq N} \|y_n\|_Y, \\ \|y_0\|_Y &\leq C_2 \|f^{(h)}\|_{F_h}, \\ \|\rho_n\|_Y &\leq C_2 \|f^{(h)}\|_{F_h}. \end{aligned} \right\} \quad (17)$$

Тогда для устойчивости

$$\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h} \quad (18)$$

достаточно, чтобы нормы  $\|R_h^m\|_Y$  степеней операторов  $R_h$  были равномерно по  $h$  ограничены, т. е. чтобы выполнялось условие

$$\|R_h^n\|_Y \leq C_3, \quad n = 1, 2, \dots, N.$$

При этом в качестве числа  $C$ , входящего в определение устойчивости (18), можно взять число

$$C = 2C_2^2 C_3.$$

Доказательство этого утверждения содержится в следующей цепочке очевидных неравенств, написанных с учетом  $N/h = 1$ , условий (17) и (18):

$$\begin{aligned} \|u^{(h)}\|_{U_h} &\leq C_2 \max_n \|y_n\|_Y \leq C_2 \max_n \|R_h^n\| [C_2 + C_2] \|f^{(h)}\|_{F_h} \leq \\ &\leq C_2 C_3 [C_2 + C_2] \|f^{(h)}\|_{F_h}, \end{aligned}$$

или

$$\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}.$$

#### 4. Примеры исследования устойчивости.

Пример 1. Займемся теперь анализом устойчивости разностной схемы (14) для системы дифференциальных уравнений. Нормы в  $U_h$  и  $F_h$  введем равенствами

$$\begin{aligned} \|u^{(h)}\|_{U_h} &= \left\| \begin{Bmatrix} v_n \\ w_n \end{Bmatrix} \right\|_{U_h} = \max_n [\max |v_n|, \max |w_n|], \\ \|f^{(h)}\|_{F_h} &= \left\| \begin{Bmatrix} p_n \\ q_n \\ a \\ b \end{Bmatrix} \right\|_{F_h} = \max [|\alpha|, |b|, \max_n |p_n|, \max_n |q_n|]. \end{aligned}$$

Как мы видели, после введения обозначений

$$y_n = \begin{bmatrix} v_n \\ w_n \end{bmatrix}, \quad R_h = \begin{pmatrix} 1 - Ah & -Bh \\ -Ch & 1 - Dh \end{pmatrix}, \quad \rho_n = \begin{bmatrix} p_n \\ q_n \end{bmatrix}, \quad y_0 = \begin{bmatrix} a \\ b \end{bmatrix}$$



рассматриваемая система разностных уравнений принимает канонический вид (13).

Введем норму в двумерном пространстве  $Y$ , которому принадлежат  $y_n$  и  $\rho_n$ , положив

$$\|y_n\|_Y = \left\| \begin{bmatrix} y_n^{(1)} \\ y_n^{(2)} \end{bmatrix} \right\|_Y = \max [ |y_n^{(1)}|, |y_n^{(2)}| ].$$

Нормы в  $U_h$ ,  $F_h$  и  $Y$  оказываются удовлетворяющими условиям (17). Поэтому для проверки устойчивости достаточно показать, что

$$\|R_h^n\|_Y \leq M, \quad n = 1, 2, \dots, N, \quad M = \text{const.}$$

Заметим, что при выбранной нами в  $Y$  норме векторов норма любого линейного оператора

$$T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$$

задается формулой

$$\|T\| = \max [ |t_{11}| + |t_{12}|, |t_{21}| + |t_{22}| ], \quad (19)$$

поскольку  $\max_{\|x\|=1} \|Tx\|_Y = \|T\|_Y$  достигается хотя бы при одном из двух векторов  $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  или  $x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ .

В силу формулы (19) для  $\|T\|$ , получим

$$\begin{aligned} \|R_h\|_Y &= \left\| \begin{pmatrix} 1 - Ah & -Bh \\ -Ch & 1 - Dh \end{pmatrix} \right\|_Y \leq \\ &\leq \max [ |1 - Ah| + |Bh|, |1 - Dh| + |Ch| ] = 1 + C'h. \end{aligned}$$

Следовательно,

$$\|R_h^n\|_Y \leq \|R_h\|_Y^n \leq (1 + C'h)^N \leq e^{C'h} = M, \quad n = 1, 2, \dots, N,$$

и устойчивость доказана.

Пример 2. Рассмотрим схему

$$\left. \begin{aligned} \frac{u_{n+1} - u_{n-1}}{2h} + Au_n &= \varphi_n, \quad n = 1, 2, \dots, N-1, \\ u_0 &= \alpha, \\ u_1 &= \beta, \end{aligned} \right\} \quad (20)$$

которая при  $\alpha = a$ ,  $\beta = (1 - Ah)a + h\varphi_0$  аппроксимирует со вторым порядком относительно  $h$  задачу Коши (1).

Нормы  $\|\cdot\|_{U_h}$  и  $\|\cdot\|_{F_h}$  введем равенствами

$$\|u^{(h)}\|_{U_h} = \max_n |u_n|,$$

$$\|f^{(h)}\|_{F_h} = \left\| \begin{bmatrix} \varphi_n \\ \alpha \\ \beta \end{bmatrix} \right\|_{F_h} = \max [|\alpha|, |\beta|, \max_n |\varphi_n|].$$

Для исследования устойчивости постараемся записать разностную схему в форме (13) и свести доказательство к получению оценки  $\|R_h^n\|_Y \leq C$ . Перепишем разностное уравнение (20) в виде

$$u_{n+1} = u_{n-1} - 2Ahu_n + 2h\varphi_n.$$

Записи его в форме (13) мешает то, что оно связывает не два, а три последовательных значения:  $u_{n-1}$ ,  $u_n$ ,  $u_{n+1}$ . Чтобы преодолеть эту трудность, положим

$$y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}.$$

Тогда пара равенств

$$\left. \begin{aligned} u_{n+1} &= u_{n-1} - 2Ahu_n + 2h\varphi_n, \\ u_n &= u_n \end{aligned} \right\} \quad (21)$$

выражает компоненты вектора  $y_n$  через компоненты вектора  $y_{n-1}$ :

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = \begin{pmatrix} -2Ah & 1 \\ 1 & 0 \end{pmatrix} \begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix} + h \begin{bmatrix} 2\varphi_n \\ 0 \end{bmatrix}.$$

Мы записали задачу (20) в форме (13), где

$$R_h = \begin{pmatrix} -2Ah & 1 \\ 1 & 0 \end{pmatrix}, \quad \rho_n = \begin{bmatrix} 2\varphi_n \\ 0 \end{bmatrix},$$

$$y_0 = \begin{bmatrix} (1 - Ah)a + h\varphi_0 \\ a \end{bmatrix}.$$

Введем норму в двумерном пространстве  $Y$ , которому принадлежат  $y_n$  и  $\rho_n$ , по формуле

$$\left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_Y = \max (|\alpha|, |\beta|).$$

Тогда нормы

$$\|u^{(h)}\|_{U_h}, \quad \|f^{(h)}\|_{F_h}, \quad \|\rho_n\|_Y, \quad \|y_0\|_Y,$$

как легко видеть, удовлетворяют условиям (17). Поэтому оценка

$$\|R_h^n\|_Y \leq \|R_h\|_Y^n = \left\| \begin{pmatrix} -2Ah & 1 \\ 1 & 0 \end{pmatrix} \right\|_Y^n \leq \leq (1 + 2|Ah|)^n \leq e^{2|A|}, \quad n = 1, 2, \dots, N,$$

доказывает устойчивость.

Пример 3. Исследуем устойчивость разностной схемы

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + A \frac{u_{n+1} - u_{n-1}}{2h} + Bu_n &= \varphi_n, \\ n &= 1, 2, \dots, N-1, \\ u_0 &= a, \\ \frac{u_1 - u_0}{h} &= b, \end{aligned} \right\} \quad (22)$$

приближающей при естественном выборе норм задачу Коши (3). Нормы  $\|u^{(h)}\|_{U_h}$  и  $\|f^{(h)}\|_{F_h}$  определим равенствами

$$\left. \begin{aligned} \|u^{(h)}\|_{U_h} &= \max_n |u_n|, \\ \|f^{(h)}\|_{F_h} &= \left\| \begin{array}{c} \varphi_n \\ a \\ b \end{array} \right\|_{F_h} = \max[|a|, |b|, \max_n |\varphi_n|]. \end{aligned} \right\} \quad (23)$$

Для приведения исследуемой схемы к каноническому виду (13) положим, как в примере 2,

$$y_n = \begin{bmatrix} u_{n+1} \end{bmatrix}; \quad (24)$$

Тогда компоненты вектора  $y_n = \begin{bmatrix} y_n^{(1)} \\ y_n^{(2)} \end{bmatrix}$  однозначно выражаются через компоненты вектора  $y_{n-1}$  в силу заданного разностного уравнения по формулам

$$\left. \begin{aligned} y_{n+1}^{(1)} &= \frac{2}{2+Ah} \left[ (2 - Bh^2)y_n^{(1)} - \frac{2-Ah}{2}y_n^{(2)} + h^2\varphi_{n+1} \right], \\ y_{n+1}^{(2)} &= y_n^{(1)}. \end{aligned} \right\} \quad (25)$$

Таким образом,

$$y_{n+1} = R_h y_n + h\varphi_n, \quad n = 0, 1, \dots, N-2,$$

где

$$R_h = \begin{pmatrix} \frac{4-2Bh^2}{2+Ah} & -\frac{2-Ah}{2+Ah} \\ 1 & 0 \end{pmatrix}, \quad \rho_n = \begin{bmatrix} \frac{2h^2}{2+Ah} \varphi_{n+1} \\ 0 \end{bmatrix}. \quad (26)$$

В силу условий  $u_0 = a$ ,  $\frac{u_1 - u_0}{h} = b$  (см. (22)) вычислим вектор  $y_0$ :

$$y_0 = \begin{bmatrix} a + bh \\ a \end{bmatrix}, \quad (27)$$

чем и завершим приведение исследуемой разностной схемы к виду (13).

Легко видеть, что если норму вектора  $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$  определить как  $\max(|\alpha|, |\beta|)$ , то нам не удастся так просто доказать устойчивость с нашим оператором  $R_h$ , так как  $\|R_h\| \approx 2$  и  $\|R_h\|^n \rightarrow \infty$ . Поэтому норму в пространстве  $Y$  определим не так, как в примере 2. Именно, положим

$$\left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_{Y_h} = \max \left[ |\alpha|, \left| \frac{\beta - \alpha}{h} \right| \right].$$

Мы поставили значок  $h$  при  $Y$ , чтобы подчеркнуть, что норма теперь зависит от  $h$ . При сделанном выборе норм между  $\|u^{(h)}\|_{U_h}$ ,  $\|f^{(h)}\|_{F_h}$ ,  $\|\rho_n\|_{Y_h}$ ,  $\|y_0\|_{L_h}$  выполнены соотношения (17). Остается проверить выполнение условия  $\|R_h^n\|_{Y_h} \leq C$ ,  $n = 1, 2, \dots, N$ . Нам известна формула (19), выражающая норму оператора через элементы задающей его матрицы, если норма в пространстве  $Y$  задана формулой

$$\left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_Y = \max [|\alpha|, |\beta|].$$

Сведем задачу вычисления нормы оператора в пространстве  $Y_h$  к этому случаю:

$$\left\| \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right\|_{Y_h} = \left\| \begin{pmatrix} 1 & 0 \\ 1/h & -1/h \end{pmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_Y = \|S \begin{bmatrix} \alpha \\ \beta \end{bmatrix}\|_Y,$$

где  $S = \begin{pmatrix} 1 & 0 \\ 1/h & -1/h \end{pmatrix}$ . Покажем, что для любого линейного преобразования  $T$ , действующего в пространстве  $Y$ , справедливо равенство  $\|T\|_{Y_h} = \|STS^{-1}\|_Y$ . В самом деле,

$$\left\| T \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_{Y_h} = \|ST \begin{bmatrix} \alpha \\ \beta \end{bmatrix}\|_Y = \|STS^{-1} (S \begin{bmatrix} \alpha \\ \beta \end{bmatrix})\|_Y.$$

Далее,

$$\begin{aligned} \|T\|_{Y_h} &= \max_x \frac{\|Tx\|_{Y_h}}{\|x\|_{Y_h}} = \max_{x \in Y} \frac{\|STS^{-1}Sx\|_Y}{\|Sx\|_Y} = \\ &= \max_{v \in Y} \frac{\|STS^{-1}v\|_Y}{\|v\|_Y} = \|STS^{-1}\|_Y. \end{aligned}$$

Теперь заметим, что

$$\|R_h^n\|_{Y_h} = \|SR_h^n S^{-1}\|_Y = \|(SR_h S^{-1})^n\|_Y \leq \|SR_h S^{-1}\|_Y^n.$$

Так как  $S^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & -h \end{pmatrix}$ , то

$$SR_h S^{-1} = \begin{pmatrix} 1 - \frac{2B}{2+Ah} h & \frac{2-Ah}{2+Ah} h \\ -\frac{2B}{2+Ah} h & 1 - \frac{2A}{2+Ah} h \end{pmatrix}.$$

Поэтому

$$\|SR_h S^{-1}\|_Y \leq 1 + Ch,$$

где  $C$  — какая-нибудь не зависящая от  $h$  постоянная, выбранная из условия

$$1 + Ch \geq \max \left[ \left| 1 - \frac{2B}{2+Ah} h \right| + \left| \frac{2-Ah}{2+Ah} h \right|, \left| \frac{2B}{2+Ah} h \right| + \left| 1 - \frac{2A}{2+Ah} h \right| \right].$$

В частности, при достаточно малых  $h$  этому условию удовлетворяет, очевидно, число  $C = 1 + 2|A| + 2|B|$ .

Итак,

$$\|R_h^n\|_{Y_h} \leq \|SR_h S^{-1}\|_Y^n \leq (1 + Ch)^N \leq e^C, \quad n = 1, 2, \dots, N,$$

что гарантирует устойчивость исследуемой схемы.

**5. Неединственность канонической записи.** Приведение разностной схемы к каноническому виду (13) можно осуществить многими способами. Полагая  $y'_n = Ty_n$ , где  $T$  — произвольное линейное преобразование пространства  $Y$ , которому принадлежат  $y_n$  и  $\rho_n$ , перейдем к записи

$$\left. \begin{aligned} y_{n+1} &= R'_h y'_n + h \rho'_n, \\ y'_0 &\text{ задано.} \end{aligned} \right\} \quad (13')$$

Здесь  $R'_h = TR_h T^{-1}$ ,  $\rho'_n = T \rho_n$ ,  $y'_0 = Ty_0$ .

Если бы в примере 3 вместо  $y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}$  мы положили  $y_n = \begin{bmatrix} u_n \\ \frac{u_{n+1} - u_n}{h} \end{bmatrix}$ , то пришли бы к записи схемы в виде (13), где

$$R_h = \begin{pmatrix} 1 & h \\ -\frac{2hB}{2+hA} & \frac{2-hA-2h^2B}{2+hA} \end{pmatrix}, \quad \rho_n = \begin{bmatrix} 0 \\ \frac{2}{2+hA} \Phi_{n+1} \end{bmatrix}.$$

При выборе нормы в  $Y$  по формуле  $\left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_Y = \max(|\alpha|, |\beta|)$  были бы выполнены условия (17). Ограниченность  $\|R_h^n\|_Y$  очевидна:

$$\|R_h^n\|_Y \leq \|R_h\|_Y^n \leq (1+Ch)^N < e^C,$$

где  $C$  выбрано из условия

$$1 + Ch \geq \max \left( 1 + h, \left| \frac{2hB}{2+hA} \right| + \left| \frac{2-hA-2h^2B}{2+hA} \right| \right) = \\ = \max \left( 1 + h, 1 + \frac{2(|A|+|B|h)}{2-|A|h} h \right).$$

Имеется произвол также и в выборе размерности пространства  $Y$ . Мы могли бы, скажем, вместо  $y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}$  положить  $y_n = \begin{bmatrix} u_{n+2} \\ u_{n+1} \\ u_n \end{bmatrix}$ , что в этом примере, впрочем, не упростило бы исследования устойчивости.

Подведем итог нашим рассуждениям. Из приведенных примеров вытекает, что для исследования устойчивости разностной схемы  $L_h u^{(h)} = \hat{f}^{(h)}$  решения задачи Коши с постоянными коэффициентами удобно привести эту разностную схему к виду (13):

$$\left. \begin{aligned} y_{n+1} &= R_h y_n + h \rho_n, \quad n = 0, 1, \dots, \\ y_0 &\text{ задано.} \end{aligned} \right\}$$

Если в пространстве, которому принадлежит  $y_n$  и  $\rho_n$ , удается ввести норму так, чтобы выполнялись условия

$$\left. \begin{aligned} \|u^{(h)}\|_{U_h} &\leq C_2 \max_n \|y_n\|, \\ \|\rho_n\| &\leq C_2 \|\hat{f}^{(h)}\|_{F_h}, \\ \|y_0\| &\leq C_2 \|\hat{f}^{(h)}\|_{F_h}, \end{aligned} \right\} \quad (28)$$

то для устойчивости достаточно, чтобы нормы степеней оператора  $R_h$  были равномерно по  $h$  ограничены,

$$\|R_h^n\| \leq C_3, \quad n = 1, 2, \dots, N.$$

Для этого достаточно, очевидно, чтобы имело место неравенство

$$\|R_h\| < 1 + C'h,$$

где  $C'$  не зависит от  $h$ .

В этом случае постоянная  $C$  в определении устойчивости

$$\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}$$

может быть взята в виде

$$C = 2C_2^2 C_3. \quad (29)$$

### ЗАДАЧИ

1. Доказать устойчивость следующих разностных схем решения задачи  $u' + Au = \varphi(x)$ ,  $u(0) = a$ . Найти константу  $C$ , входящую в определение  $\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}$  устойчивости.

$$\text{а) } \left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + A(nh)u_n = \varphi_n, \quad n = 0, 1, 2, \dots, N-1, \\ u_0 = a, \end{aligned} \right\}$$

если  $|A(x)| \leq M = \text{const}$ , а нормы введены равенствами

$$\|u^{(h)}\|_{U_h} = \max_n |u_n|, \quad \|f^{(h)}\|_{F_h} = \max [ |a|, \max_n |\varphi_n| ].$$

$$\text{б) } \left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + Au_{n+1} = \varphi_n, \\ u_0 = a. \end{aligned} \right\} \text{Нормы — как в а).}$$

$$\text{в) } \left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + A \frac{u_n + u_{n+1}}{2} = \varphi \left[ \left( n + \frac{1}{2} \right) h \right], \quad n = 0, \dots, N-1, \\ u_0 = a, \end{aligned} \right\}$$

$$\|u^{(h)}\|_{U_h} = \max_n |u_n|, \quad \|f^{(h)}\|_{F_h} = \max \left( |a|, \max_n \left| \varphi \left[ \left( n + \frac{1}{2} \right) h \right] \right| \right).$$

2. Решить задачу 1 в предположении, что  $u_n = \begin{bmatrix} u_n^{(1)} \\ u_n^{(2)} \end{bmatrix}$  — вектор,

$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  — матрица,  $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ ,  $\varphi_n = \begin{bmatrix} \varphi_n^{(1)} \\ \varphi_n^{(2)} \end{bmatrix}$  — векторы. Нормы

заданы в виде

$$\|u^{(h)}\|_{U_h} = \max_n (|u_n^{(1)}| + |u_n^{(2)}|),$$

$$\|f^{(h)}\|_{F_h} = \max [ |a_1| + |a_2|, \max_n (|\varphi_n^{(1)}| + |\varphi_n^{(2)}|) ], \quad |A_{ij}(x)| \leq M.$$

3. Привести к каноническому виду:  $y_{n+1} = R_h y_n + h \rho_n$ ,  $y_0$  задано — разностное уравнение

$$\frac{u_{n+2} - 2u_{n+1} + 3u_n - 4u_{n-1}}{h} - 5u_n = \varphi_n, \quad n = 1, 2, \dots,$$

положив  $y_n = \begin{bmatrix} u_{n+2} \\ u_{n+1} \\ u_n \end{bmatrix}$ .

### § 15. Необходимый спектральный признак устойчивости

В §14 мы показали, что приведение разностной схемы решения задачи Коши с постоянными коэффициентами

$$L_h u^{(h)} = f^{(h)} \quad (1)$$

к виду

$$\left. \begin{array}{l} y_{n+1} = R_h y_n + h \rho_n, \quad n = 0, 1, \dots, \\ y_0 \text{ задано} \end{array} \right\} \quad (2)$$

может быть использовано для доказательства устойчивости: при определенных условиях (условия (17) § 14) оценка

$$\|R_h^n\|_Y \leq C, \quad n = 1, 2, \dots, N, \quad (3)$$

достаточна для устойчивости.

Здесь мы покажем, что эта оценка (3) при некоторых естественных условиях необходима для устойчивости. Покажем также, что, независимо от выбора нормы, для оценки (3) необходимо, чтобы спектр матрицы  $R_h$ , т. е. совокупность корней уравнения

$$\det(R_h - \lambda E) = 0, \quad (4)$$

лежал в круге

$$|\lambda| < 1 + Ch, \quad (5)$$

где  $C$  не зависит от  $h$ .

Перейдем к реализации намеченной программы.

**1. Ограниченность норм степеней оператора перехода необходима для устойчивости.** Описанные нами способы приведения разностных уравнений к виду (2) таковы, что в случае нулевых правых частей разностных уравнений выражение  $\rho_n$  также тождественно равно нулю.

Пусть постоянные  $M_1 = M_1(h) > 0$  и  $M_2 = M_2(h) > 0$  выбраны так, чтобы выполнялись неравенства

$$\|u^{(h)}\|_{U_h} \geq M_1 \max_n \|y_n\|, \quad (6)$$

$$\|y_0\| \geq M_2 \|f^{(h)}\|_{F_h} \quad (7)$$



(последнее при условии  $\rho_n \equiv 0$ ). Тогда при нулевых правых частях разностного уравнения (или системы разностных уравнений) уравнение (2) примет вид

$$y_{n+1} = R_h y_n,$$

откуда

$$y_n = R_h^n y_0. \quad (8)$$

Далее, в силу (6) и (8)

$$\|u^{(h)}\|_{U_h} \geq M_1 \max_n \|R_h^n y_0\|. \quad (9)$$

Из определения нормы линейного оператора следует, что вектор  $y_0$  из конечномерного пространства всегда можно выбрать так, чтобы при данном  $n$  было  $\|R_h^n y_0\| = \|R_h^n\| \cdot \|y_0\|$ . Поэтому при некотором  $y_0$  (зависящем от  $h$ )

$$\max_n \|R_h^n y_0\| = \max_n \|R_h^n\|_Y \cdot \|y_0\|. \quad (10)$$

При таком выборе  $y_0$  в силу (9) и (10) получим

$$\|u^{(h)}\|_{U_h} \geq M_1 \max_n \|R_h^n\| \cdot \|y_0\| \geq M_1 M_2 \max_n \|R_h^n\| \cdot \|f^{(h)}\|_{F_h}. \quad (6')$$

Из последней оценки следует, что в случае устойчивости разностной схемы (1) постоянная  $C'$ , входящая в определение устойчивости

$$\|u^{(h)}\|_{U_h} < C' \|f^{(h)}\|_{F_h}$$

неизбежно должна удовлетворять оценке

$$C' \geq M_1 M_2 \max_n \|R_h^n\|. \quad (6'')$$

Отсюда видно, в частности, что если нормы  $\|u^{(h)}\|_{U_h}$ ,  $\|f^{(h)}\|_{F_h}$  и  $\|y_n\|$  согласованы так, что выполнены условия (6) и (7), то условие (3) необходимо для устойчивости. Условие (3) равносильно тому, что решение  $\{y_n\}$  однородного уравнения  $y_{n+1} = R_h y_n$  при любом  $y_0$  удовлетворяет неравенству

$$\|y_n\| \leq C \|y_0\|, \quad n = 1, 2, \dots, N. \quad (11)$$

В примерах 1 и 2 § 14 числа  $M_1$  и  $M_2$  можно было выбрать не зависящими от  $h$  (равными единице), как без труда проверит читатель. Это указывает на естественность сформулированных условий.

В примере 3 из § 14 для разностной схемы (22) при использовании равенства (24) и при нормах (23) условие  $\|u^{(h)}\|_{U_h} \geq$

$\geq M_1 \max_n \|y_n\|$  выполняется, только если  $M_1 \leq h/2$ . Но если изменить выбор нормы  $\|u^{(h)}\|_{U_h}$ , положив

$$\|u^{(h)}\|_{U_h} = \max \left[ \max_n |u_n|, \max_n \left| \frac{u_{n+1} - u_n}{h} \right| \right], \quad (12)$$

то можно положить  $M_1 = 1$ ,  $M_2 = 1$ , и оценка (3) необходима для устойчивости. При таком изменении норм продолжают выполняться сформулированные в § 14 условия (17), при которых оценка (3) достаточна для устойчивости.

**2. Спектральный признак устойчивости.** Для оценки  $\|R_h^n\|$  можно пользоваться собственными значениями матрицы  $R_h$ , т. е. корнями  $\lambda$  уравнения

$$\det \|R_h - \lambda E\| = 0.$$

Если  $\lambda$  — собственное значение, то существует собственный вектор  $y$  такой, что  $R_h y = \lambda y$ . Поэтому

$$R_h^n y = \lambda^n y, \quad \|R_h^n y\| = |\lambda|^n \|y\|, \quad \|R_h^n\| \geq |\lambda|^n.$$

Таким образом, для ограниченности  $\|R_h^n\|$  необходимо, чтобы были ограничены степени собственных значений  $|\lambda^n|$ ,  $n = 1, 2, \dots, N$ . Для этого все собственные значения должны лежать в круге

$$|\lambda| \leq 1 + ch \quad (13)$$

на комплексной плоскости, где  $c$  не зависит от  $h$ . В противном случае на произвольном  $c$  и некотором достаточно малом  $h$

$$\|R_h^n\| \geq |\lambda|^N > (1 + ch)^{\frac{1}{h}} = e^{\frac{1}{h} \ln(1+ch)} \geq e^c \left(1 - \frac{ch}{2}\right) > e^{\frac{c}{2}}.$$

Сформулированный признак оценки степеней  $\|R_h^n\|$  по расположению спектра (т. е. совокупности собственных значений) оператора  $R_h$  не зависит, очевидно, от выбора нормы в пространстве, где действует оператор  $R_h$ .

Спектральный признак устойчивости (13) не зависит также от способа приведения схемы (1) к виду (2). Если приведение осуществлено иначе,  $y'_{n+1} = R'_h y'_n + h \rho'_n$ , так что  $y' = T y$ ,  $R'_h = T R_h T^{-1}$ , где  $T$  — произвольный невырожденный линейный оператор, то спектры операторов  $R_h$  и  $R'_h$  совпадают. В самом деле,

$$\begin{aligned} \det(R'_h - \lambda E) &= \det(T R_h T^{-1} - \lambda E) = \det[T(R_h - \lambda E)T^{-1}] = \\ &= \det T \det(R_h - \lambda E) \det T^{-1} = \det(R_h - \lambda E). \end{aligned}$$

Поэтому уравнения  $\det(R_h - \lambda E) = 0$  и  $\det(R'_h - \lambda E) = 0$  имеют одинаковые корни  $\lambda$ .

**3. Обсуждение спектрального признака устойчивости.** Выше было показано, что при выборе норм в соответствии с условиями (6) и (7) расположение спектра оператора  $R_h$  в круге

$$|\lambda| \leq 1 + ch, \quad (13)$$

необходимое для ограниченности  $\|R_h^n\|$ , необходимо также и для устойчивости.

Пусть условие (13) грубо нарушено, так что при достаточно малых  $h > 0$  имеется собственное число  $\lambda$ , по модулю существенно превосходящее единицу, скажем,

$$|\lambda| > 1 + h^{1-\varepsilon},$$

где  $\varepsilon > 0$  не зависит от  $x$ . Тогда разностная схема (1) неустойчива при любом разумном выборе норм  $\|u^{(h)}\|_{U_h}$  и  $\|f^{(h)}\|_{F_h}$ , даже если и не ограничивать свободу этого выбора условиями (6) и (7).

Это высказывание нельзя назвать теоремой хотя бы потому, что оно оперирует термином «разумный», не получившим точного определения. Объявляем, что мы имеем в виду.

При любом разумном выборе нормы  $\|u^{(h)}\|_{U_h}$  можно так подобрать положительное  $k_1$ , чтобы при всех достаточно малых  $h$  выполнялось неравенство

$$\|u^{(h)}\|_{U_h} \geq h^{k_1} \max_n |u_n|. \quad (14)$$

В противном случае, очевидно, не может быть выполнено равенство (4) из § 13:

$$\lim_{h \rightarrow 0} \| [u]_h \|_{U_h} = \| u \|_{U}.$$

Далее, при любом разумном выборе нормы  $\|f^{(h)}\|_{F_h}$  можно так подобрать  $k_2 > 0$ , чтобы при всех достаточно малых  $h$  выполнялось неравенство

$$\|f^{(h)}\|_{F_h} < h^{-k_2} F, \quad (15)$$

где через  $F$  обозначен максимум модулей компонент элемента  $f^{(h)}$  пространства  $F_h$ . В противном случае разностная схема (1) не может аппроксимировать задачу  $Lu = f$  на решении  $u$ : ведь мы видели, что компоненты невязки  $\delta f^{(h)}$ , возникающей при подстановке  $[u]_h$  в левую часть приближающей задачи разностной схемы (1), стремятся к нулю не быстрее, чем некоторая степень шага  $h$ .

Приведем теперь разностную схему (1) к виду (2), полагая для этого

$$y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}, \quad \left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_{\mathcal{Y}} = \max \{ |\alpha|, |\beta| \}. \quad (16)$$

Для определенности мы считаем, что рассматривается разностное уравнение, которое связывает три последовательных значения  $u_{n-1}$ ,  $u_n$ ,  $u_{n+1}$ .

Если правую часть разностного уравнения, на основе которого построена схема (1), положить равной нулю, то при некотором  $r > 0$  будет выполнено неравенство

$$\max \{ |u_0|, |u_1| \} = \|y_0\|_{\mathcal{Y}} \geq h^r F, \quad (17)$$

поскольку соотношения, связывающие  $u_1$  и  $u_0$  и входящие в разностную схему, имеют вид

$$\left. \begin{array}{l} u_0 = a, \\ u_1 = b, \end{array} \right\} \text{ либо } \frac{u_1 - u_0}{h} = b, \left. \begin{array}{l} u_0 = a, \\ \end{array} \right\}$$

либо им аналогичный.

Теперь ясно, что всегда можно добиться справедливости неравенств (6) и (7), положив  $M_1(h) = h^{k_1}$ ,  $M_2(h) = h^{r+k_2}$ . В самом деле (см. также (14) и (17)),

$$\begin{aligned} \|u^{(h)}\|_{U_h} &\geq h^{k_1} \max_n |u_n| = h^{k_1} \max_n \|y_n\|, \\ \|y_0\| &\geq h^r F = h^r h^{k_2'} (h^{-k_2} F) \geq h^{r+k_2} \|f^{(h)}\|_{F_h}. \end{aligned}$$

Таким образом, неравенство (6') примет вид

$$\|u^{(h)}\|_{U_h} \geq h^{r+k_1+k_2} \max_n \|R_h^n\| \cdot \|f^{(h)}\|_{F_h} \geq h^{r+k_1+k_2} (1 + h^{1-\varepsilon})^{1/h} \|f^{(h)}\|_{F_h}.$$

Это означает неустойчивость, потому что при любых  $r$ ,  $k_1$ ,  $k_2$  и  $\varepsilon > 0$ , как легко видеть,

$$h^{r+k_1+k_2} (1 + h^{1-\varepsilon})^{1/h} \rightarrow \infty \text{ при } h \rightarrow 0.$$

Этим мы закончим изложение соображений, показывающих, что если среди собственных значений матрицы  $R_h$  есть корень, удовлетворяющий неравенству  $|\lambda| > 1 + h^{1-\varepsilon}$ , то она неустойчива при любом разумном выборе норм.

Воспользуемся необходимым спектральным признаком устойчивости (13) и докажем, что схема, рассмотренная в § 9, действительно неустойчива. В § 9 строгого исследования неустойчивости не могло быть проведено хотя бы потому, что там в нашем распоряжении еще не было аккуратных определений.

Интересующая нас разностная схема приближает задачу

$$\left. \begin{array}{l} u' + Au = 0, \quad 0 \leq x \leq 1, \\ u(0) = a \end{array} \right\} \quad (18)$$

и имеет вид

$$\left. \begin{array}{l} 4 \frac{u_{n+1} - u_{n-1}}{2h} - 3 \frac{u_{n+1} - u_n}{h} + Au_n = 0, \quad n = 1, 2, \dots, N+1, \\ u_0 = a, \\ u_1 = (1 - Ah)a. \end{array} \right\} \quad (19)$$

Положив  $y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}$ , приведем схему (19) к виду (2), где

$$R_h = \begin{pmatrix} 3 + Ah & -2 \\ 1 & 0 \end{pmatrix}, \quad \rho_n \equiv 0.$$

Собственные значения матрицы  $R_h$  суть корни квадратного уравнения  $\det(R_h - \lambda E) = 0$ :

$$\lambda_{1,2} = \frac{3 + Ah}{2} \pm \sqrt{\left(\frac{3 + Ah}{2}\right)^2 - 2}.$$

Первый корень  $\lambda_1(h)$  при  $h \rightarrow 0$  стремится к числу 2, так что при малых  $h$

$$|\lambda_1| > \frac{3}{2} > 1.$$

Поэтому нельзя ожидать устойчивости ни при каком разумном выборе норм.

В частности, если ввести нормы равенствами

$$\|u^{(h)}\|_{U_h} = \max_n |\varphi_n|,$$

$$\|f^{(h)}\|_{F_h} = \left\| \begin{array}{c} \varphi_n \\ \alpha \\ \beta \end{array} \right\|_{F_h} = \max [|\alpha|, |\beta|, \max_n |\varphi_n|],$$

$$\|y_n\|_Y = \left\| \begin{array}{c} y_n^{(1)} \\ y_n^{(2)} \end{array} \right\|_Y = \max [ |y_n^{(1)}|, |y_n^{(2)}| ],$$

то будут выполнены оба условия (6), (7), при которых неравенство (3) необходимо для устойчивости. Однако  $\|R_h^n\| > (3/2)^n \rightarrow \infty$ , если  $n = 1/h$ ,  $h \rightarrow 0$  и устойчивости нет.

Как мы видели, грубое нарушение необходимого спектрального признака устойчивости (13):

$$|\lambda| \leq 1 + ch,$$

например наличие собственного числа  $\lambda^*$  оператора  $R_h$ , удовлетворяющего оценке

$$|\lambda^*| > 1 + h^{1-\varepsilon},$$

свидетельствует о непоправимой за счет выбора норм неустойчивости.

Подчеркнем, однако, что расположение спектра оператора  $R_h$  внутри круга  $|\lambda| < 1 + ch$  еще не гарантирует устойчивости. Устойчивость в этом случае может зависеть от удачного выбора норм, как показывает пример следующей разностной схемы, которую мы уже рассматривали в § 14 с несколько иной точки зрения.

Разностную схему решения задачи  $u'' = \varphi(x)$ ,  $u(0) = a$ ,  $u'(0) = b$  выберем так:

$$\left. \begin{aligned} \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} &= \varphi_n, & n = 1, 2, \dots, N-1, \\ u_0 &= a, \\ \frac{u - u_0}{h} &= b. \end{aligned} \right\}$$

Положив  $y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}$ , запишем ее в виде (2), где

$$R_h = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad \rho_n = \begin{bmatrix} h\varphi_n \\ 0 \end{bmatrix}, \quad y_0 = \begin{bmatrix} a + hb \\ a \end{bmatrix}.$$

Оба собственных числа матрицы  $R_h$  равны единице. В случае  $\varphi_n \equiv 0$  решение  $\{u_n\}$  задачи имеет вид

$$u_n = u_0 + (u_1 - u_0)n, \quad n = 0, 1, 2, \dots, N.$$

Используем два набора норм:

$$\begin{aligned} 1) \quad & \|y_n\|_Y = \max(|y_n^{(1)}|, |y_n^{(2)}|), \\ & \|u^{(h)}\|_{U_h} = \max_n \|y_n\|_Y, \\ & \|f^{(h)}\|_{F_h} = \left\| \begin{matrix} \varphi_n \\ a \\ b \end{matrix} \right\|_{F_h} = \max[\|y_0\|_Y, \max_m |\varphi_m|]; \\ 2) \quad & \|y_n\|_{Y_h} = \max\left(|y_n^{(1)}|, \left|\frac{y_n^{(2)} - y_n^{(1)}}{h}\right|\right), \\ & \|u^{(h)}\|_{U_h} = \max_n \|y_n\|_{Y_h}, \\ & \|f^{(h)}\|_{F_h} = \max[\|y_0\|_{Y_h}, \max_m |\varphi_m|]. \end{aligned}$$

Читатель легко убедится, что в обоих случаях выполнены условия (6), (7) и условия (28) из § 14, при которых устойчивость равносильна оценке

$$\|R_h^n\| \leq C, \quad n = 1, 2, \dots, N-1.$$

При выборе норм по формулам 1) эта оценка не выполняется. Например, полагая  $y_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ,  $\|y_0\| = 1$ , получим

$$y_n = \begin{bmatrix} n+1 \\ n \end{bmatrix}, \quad \|y_n\| = n+1, \quad \|R_h^n\| \geq n+1 \rightarrow \infty$$

при  $n = 1/h$ ,  $h \rightarrow 0$ .

При выборе норм по формулам 2) устойчивость имеется: при произвольном  $y_0 = \begin{bmatrix} u_1 \\ u_0 \end{bmatrix}$  имеем

$$\begin{aligned} \|y_n\|_{Y_h} &= \|R_h^n y_0\| = \left\| \begin{array}{c} u_0 + (u_1 - u_0)(n-1) \\ u_0 + (u_1 - u_0)n \end{array} \right\|_{Y_h} = \\ &= \max \left[ |u_0 + (u_1 - u_0)(n+1)| \cdot \left| \frac{u_1 - u_0}{h} \right| \right]. \end{aligned}$$

Но  $n+1 \leq 1/h$ , поэтому

$$\|y_n\|_{Y_h} = \|R_h^n y_0\|_{Y_h} \leq |u_0| + \left| \frac{u_1 - u_0}{h} \right| \leq 2 \|y_0\|_{Y_h}$$

и

$$\|R_h^n\| < 2, \quad n = 1, 2, \dots, N-1.$$

На практике часто ограничиваются проверкой того, выполняется ли необходимый спектральный признак устойчивости. Если он выполнен, дальнейшую проверку пригодности схемы устанавливают путем экспериментального счета по этой схеме, не заботясь о явном конструировании норм. Способам такой проверки посвящен § 18.

### ЗАДАЧИ

1. Пусть разностное уравнение второго порядка  $au_{n-1} + bu_n + cu_{n+1} = \varphi_n$  приведено к виду  $y_{n+1} = R_h y_n + h\varphi_n$  с помощью замены  $y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}$ . Показать, что корни характеристического уравнения  $a + b\lambda + c\lambda^2 = 0$  и собственные значения матрицы  $R_h$  совпадают.

2. Записать разностное уравнение второго порядка  $au_{n-1} + bu_n + cu_{n+1} = \varphi_n$  в виде  $y_{n+1} = R_h y_n + h\varphi_n$  с помощью замены  $y_n = \begin{bmatrix} u_{n+2} \\ u_{n+1} \\ u_n \end{bmatrix}$ . Един-

ственно ли такое приведение? Показать, что собственными значениями матрицы  $R_h$  являются корни характеристического уравнения  $a + b\lambda + c\lambda^2 = 0$  и еще число  $\lambda = 0$ , так что выполнение спектрального признака устойчивости

$$|\lambda| \leq 1 + c\dot{h} \text{ не зависит от выбора } y_n = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} \text{ или } y_n = \begin{bmatrix} u_{n+2} \\ u_{n+1} \\ u_n \end{bmatrix}.$$

3. Пусть собственные векторы  $v^{(1)}$  и  $v^{(2)}$  матрицы второго порядка  $R_h$ , отвечающие собственным значениям  $\lambda_1$  и  $\lambda_2$  соответственно,

$$R_h v^{(1)} = \lambda_1 v^{(1)}, \quad R_h v^{(2)} = \lambda_2 v^{(2)},$$

при  $h \rightarrow 0$  стремятся к различным неколлинеарным предельным положениям. Тогда условия  $|\lambda_1| < 1 + ch$ ,  $|\lambda_2| < 1 + ch$  не только необходимы, но и достаточны для оценки вида  $\|R_h^n\| < C$ ,  $n = 1, 2, \dots, N$ , если  $\left\| \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right\|_{Y'} = \max[|\alpha|, |\beta|]$ . Доказать.

## § 16. Ошибки округления

1. Ошибки в коэффициентах. Если разностная схема

$$L_h u^{(h)} = f^{(h)} \quad (1)$$

аппроксимирует задачу  $Lu = f$  на решении  $u$  и устойчива, то имеет место сходимость. Однако задуманная разностная схема никогда не реализуется точно из-за ошибок округления в задании ее коэффициентов и правых частей.

Пусть, например, требуется решить задачу

$$\left. \begin{aligned} u' + Au &= \cos x, & 0 \leq x \leq 1, \\ u(0) &= a \end{aligned} \right\}$$

по разностной схеме

$$\left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + Au_n &= \cos x_n, & n = 0, 1, \dots, N-1, \\ u_0 &= a. \end{aligned} \right\} \quad (2)$$

Значения  $\cos x_n$ ,  $a$ ,  $A$  и коэффициент  $1/h$  задаются с теми или иными ошибками округления. В общем случае вместо (1) мы имеем дело с разностной схемой

$$L_h v^{(h)} + (\Delta^{(h)} L_h) v^{(h)} = f^{(h)} + \Delta^{(h)} f^{(h)}, \quad (3)$$

где  $\Delta^{(h)} L_h$  и  $\Delta^{(h)} f^{(h)}$  — погрешности в задании оператора  $L_h$  и правой части  $f^{(h)}$ , вызванные округлениями. Для схемы (2) оператор  $\Delta^{(h)} L_h$  имеет вид

$$(\Delta^{(h)} L_k) v^{(h)} = \begin{cases} \Delta^{(h)} \left( \frac{1}{h} \right) (v_{n+1} - v_n) + (\Delta^{(h)} A) v_n, & n = 0, 1, \dots, N-1, \\ 0 \cdot v_0. \end{cases}$$

Погрешность  $\Delta^{(h)} f^{(h)}$  задается формулой

$$\Delta^{(h)} f^{(h)} = \begin{cases} \Delta^{(h)} \cos x_n, & n = 0, 1, \dots, N-1, \\ \Delta^{(h)} a. \end{cases}$$

Здесь  $\Delta^{(h)} M$  — погрешность, допущенная при задании величины  $M$ .

Чтобы избежать чисто технических трудностей, ограничимся случаем, когда операторы  $L_h$  и  $\Delta^{(h)} L_h$  линейны, а пространство  $U_h$  имеет конечную размерность, как в рассмотренной схеме (2). При этих предположениях исследуем, каковы допустимые ошибки округления и как должна возрастать точность задания раз-



ностной схемы по мере измельчения сетки, т. е. при стремлении  $h$  к нулю.

**Теорема.** Если устойчивая разностная схема (1) аппроксимирует задачу  $Lu = f$  на решении и с некоторым порядком  $h^k$ :

$$\|L_h[u]_h - f^{(h)}\|_{F_h} \leq ch^k,$$

то при условиях

$$\left. \begin{aligned} \|\Delta^{(h)}L_h v^{(h)}\|_{F_h} &\leq c_1 h^k \|v^{(h)}\|_{U_h}, \\ \|\Delta^{(h)}f^{(h)}\|_{F_h} &\leq c_2 h^k \end{aligned} \right\} \quad (4)$$

разностная схема (3) тоже аппроксимирует задачу  $Lu = f$  с порядком  $h^k$  и тоже устойчива.

Таким образом, при условиях (4) порядок точности разностной схемы (3), по которой фактически производится счет, есть  $h^k$  и совпадает с порядком точности задуманной разностной схемы (1).

В предположении, что норма  $\|\cdot\|_{U_h}$  выбрана в соответствии с условием (4) из § 13, т. е. так, что

$$\lim_{h \rightarrow 0} \| [u]_h \|_{U_h} = \| u \|_U,$$

величина  $\| [u]_h \|_{U_h}$  остается ограниченной при  $h \rightarrow 0$ ,  $\| [u]_h \|_{U_h} \leq P < \infty$ . Обозначим

$$\begin{aligned} \tilde{L}_h \tilde{u}^{(h)} &\equiv L_h u^{(h)} + (\Delta^h L_h) u^{(h)}, \\ \tilde{f}^{(h)} &\equiv f^{(h)} + \Delta^{(h)} f^{(h)} \end{aligned}$$

и убедимся, что схема  $\tilde{L}_h u^{(h)} = \tilde{f}^{(h)}$  имеет порядок аппроксимации  $h^k$ . В самом деле, имеем

$$\begin{aligned} \|\tilde{L}_h [u]_h - \tilde{f}^{(h)}\|_{F_h} &= \|L_h [u]_h - f^{(h)} + (\Delta^{(h)} L_h [u]_h - \Delta^{(h)} f^{(h)})\|_{F_h} \leq \\ &\leq \|L_h [u]_h - f^{(h)}\|_{F_h} + \|\Delta^{(h)} L_h [u]_h\|_{F_h} + \|\Delta^{(h)} f^{(h)}\|_{F_h} \leq \\ &\leq ch^k + c_1 P h^k + c_2 h^k \leq \tilde{c} h^k. \end{aligned}$$

Для доказательства теоремы нам будет полезна следующая известная

**Лемма.** Пусть  $A$  и  $B$  — два линейных оператора, отображающих некоторое конечномерное линейное нормированное пространство  $X$  в другое линейное нормированное пространство  $G$ .

Пусть, далее, при произвольном  $g \in G$  существует решение  $x \in X$  уравнения

$$Ax = g,$$

причем

$$\|x\|_X \leq c \|g\|_G, \quad (5)$$

а также при любом  $x \in X$  выполнено неравенство

$$\|Bx\|_G \leq \frac{q}{c} \|x\|_X, \quad (6)$$

где  $c > 0$ ,  $0 < q < 1$ ,  $c$  и  $q$  — некоторые числа. Тогда уравнение

$$(A + B)\tilde{x} = g$$

имеет единственное решение при любом  $g \in G$  и выполнено неравенство

$$\|\tilde{x}\|_X \leq \frac{c}{1-q} \|g\|_G. \quad (7)$$

Доказательство. Заметим, что  $X$  и  $G$  имеют одинаковую размерность, так как иначе не при всяком  $g \in G$  была бы разрешима задача  $Ax = g$ .

Далее, если  $x_0$  — какое-нибудь решение уравнения

$$(A + B)x_0 = g,$$

то

$$\begin{aligned} Ax_0 &= g - Bx_0, \\ x_0 &= A^{-1}g - A^{-1}Bx_0, \end{aligned}$$

где  $A^{-1}g$  и  $A^{-1}Bx_0$  — решения уравнений  $Ax = g$  и  $Ax = Bx_0$ ,

$$\begin{aligned} \|x_0\|_X &\leq \|A^{-1}g\|_X + \|A^{-1}(Bx_0)\|_X \leq \\ &\leq c \|g\|_G + c \|Bx_0\| \leq c \|g\|_G + c \frac{q}{c} \|x_0\|_X. \end{aligned}$$

Отсюда

$$\|x_0\|_X \leq \frac{c}{1-q} \|g\|_G.$$

Из последнего неравенства следует, что при  $g = 0$  существует только тривиальное решение  $x_0 = 0$  уравнения  $(A + B)x = g$ , а значит, существует единственное решение при произвольном  $g \in G$ , и справедлива оценка (7).

Доказательство теоремы. Воспользуемся леммой и примем за операторы  $A$  и  $B$  соответственно  $L_h$  и  $\Delta^{(h)}L_h$ . Существование решения задачи  $Ax = g$  и оценка (5) равносильны предположению устойчивости схемы (1). Оценка (6) имеет ме-

сто в силу (4) при любом положительном  $q$ , если только  $h$  достаточно мало.

Разрешимость уравнения  $(A + B)x = g$  при любом  $g \in G$  и оценка (7) в точности равносильны факту устойчивости разностной схемы (3).

Отметим, что ограничения (4) на ошибки округления при задании устойчивой разностной схемы являются вполне разумными: если, уменьшая  $h$ , мы хотим получить ответ с точностью до  $h^k$ , т. е. с числом десятичных знаков порядка  $\ln(1/h)$ , то и коэффициенты разностной схемы надо задавать все более точно, увеличивая число знаков, с которыми они задаются, тоже со скоростью возрастания величины  $\ln(1/h)$ . Такое возрастание обычно вполне реализуемо, так как  $\ln(1/h)$  — медленно растущая функция. Если уменьшать шаги, не увеличивая числа десятичных знаков, с которыми заданы коэффициенты и правые части, то никакого повышения точности не получится.

**2. Ошибки в вычислениях.** После того как разностная схема задана, нужно еще вычислить ее решение  $u^{(h)}$ . Предположим, что разностные уравнения мы умеем решать точно. Тогда, если применяемая разностная схема аппроксимирует дифференциальное уравнение и устойчива, то при достаточно мелком шаге решение  $u^{(h)}$  мало отличается от искомого точного решения  $[u]_h$ .

При этом совершенно безразличен тот порядок действий (алгоритм), который используется для вычисления  $u^{(h)}$ , так как ответ не зависит от порядка действий.

Но в действительности, избрав какой-нибудь алгоритм для вычисления решения  $u^{(h)}$ , мы на каждом шаге осуществления этого алгоритма допускаем ошибки округления, которые оказывают влияние на результаты, получаемые на последующих шагах вычислений. При фиксированном  $h$  и конечномерном пространстве  $U_h$  алгоритм состоит из конечной последовательности арифметических действий. Результат каждого арифметического действия (вычисление суммы, разности, произведения или частного) непрерывным образом зависит от величин, над которыми это действие осуществляется. Поэтому, ведя вычисления с «достаточно большим» числом десятичных знаков, мы можем вычислить  $u^{(h)}$  с любым наперед заданным числом десятичных знаков. Число запасных десятичных знаков, с которым ведутся вычисления для получения  $u^{(h)}$  с заданным числом знаков, зависит и от избранного алгоритма и от  $h$ . Так, например, в § 7 показано, что при решении прогонкой хорошо обусловленной краевой задачи число запасных десятичных знаков вовсе не возрастает при  $h \rightarrow 0$ . Иногда, казалось бы, разумные алгоритмы для решения устойчивых задач могут требовать быстро возрастающего числа запасных десятичных знаков, пропорционального

$1/h$ . Пример такого алгоритма приведен в п. 2 § 5. С уменьшением  $h$  это число, вообще говоря, должно возрастать. Алгоритмы, в которых это число возрастает слишком быстро, считаются неустойчивыми и практически непригодны для счета. Вопрос об исследовании устойчивости алгоритмов сложный. Примером такого исследования является обоснование прогонки (§ 7). Но в простейших случаях удается понять, каково требуемое число запасных десятичных знаков, опираясь лишь на сведения об устойчивости разностной схемы и доказанную в п. 1 теорему о возможности задавать разностную схему приближенно.

Пусть, например, мы ведем вычисления по разностной схеме

$$\frac{u^{(h)}(x+h) - u^{(h)}(x)}{h} + Au^{(h)}(x) = f^{(h)}(x).$$

Находя  $u^{(h)}(x+h)$  по рекуррентной формуле

$$u^{(h)}(x+h) = u^{(h)}(x)(1 - Ah) + hf^{(h)}(x)$$

и ведя расчет с конечным числом десятичных знаков, можем допустить в  $u^{(h)}(x+h)$  некоторую ошибку  $\delta$ . Удобно считать, что ошибка допущена не в значении  $u^{(h)}(x+h)$ , а в использованной правой части  $f^{(h)}(x)$ , т. е. считать, что мы  $u^{(h)}(x+h)$  вычислили точно, но вместо  $f^{(h)}(x)$  использовали величину  $f^{(h)}(x) + \delta/x$ . Так как такие ошибки совершаются во всех точках  $x$ , то величину  $\delta$  следует считать зависящей от  $x$ , так что  $\delta = \delta(x)$ . Таким образом, в этом примере ошибку округления при вычислениях можно считать погрешностью  $\delta(x)/h$  в задании правой части. Рассматриваемая схема имеет первый порядок аппроксимации и устойчива. Поэтому, чтобы не испортить сходимости со скоростью  $h$ , мы должны вести вычисления с возрастающей точностью, а именно так, чтобы

$$\Delta^{(h)} f^{(h)} = \frac{\delta(x)}{h}$$

было порядка  $h$ .

Для этого  $\delta(x)$  должно быть порядка  $h^2$ . Этого можно добиться, ведя вычисления  $u^{(h)}$  с возрастающим при  $h \rightarrow 0$  как  $\ln(1/h)$  числом запасных десятичных знаков.

На этом примере мы показали, что в простых случаях ошибки округления при вычислении решения  $u^{(h)}$  с точностью до множителя вида  $h^m$  можно считать ошибками задания правых частей  $f^{(h)}$ . Из доказанной выше теоремы следует, что тогда для устойчивых схем эти ошибки не мешают сходимости без потери порядка точности, если число десятичных знаков, с которыми ведется счет, медленно возрастает, как  $c \ln(1/h)$ , где  $c$  — некоторая постоянная.

### § 17. Количественная характеристика устойчивости

Начнем с рассмотрения хорошо известного примера разностной схемы

$$\left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + Au_n &= 0, \\ u_0 &= 1 \end{aligned} \right\} \quad (1)$$

для дифференциальной краевой задачи

$$u' + Au = 0, \quad u(0) = 1.$$

Ее решение имеет вид

$$u_n = e^{-Ax_n} + h \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2)$$

(см. (3') из § 8; полагаем  $b = 1$ ). Выражение (6) из § 8

$$\delta(x_n) = h \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2)$$

представляет собою остаточный член, т. е. ошибку от замены значения  $e^{-Ax_n}$  точного решения дифференциального уравнения решением  $u_n^{(h)}$  разностной задачи. Остаточный член стремится к нулю, как первая степень  $h$ ; эта схема имеет первый порядок точности. Выбор шага  $h$  зависит от точности, которую мы хотим достичь. Ясно, что модуль отношения ошибки к точному решению  $|\delta(x_n)/u(x_n)|$  должен быть во всяком случае меньше единицы, чтобы приближенное решение можно было считать сколько-нибудь точным.

Посмотрим, при каких  $h$  это условие выполняется. В выражении  $\delta(x_n)$  пренебрежем слагаемым  $O(h^2)$  и рассмотрим отношение ошибки  $\delta(x_n)$  в точке  $x_n$  к точному решению:

$$\frac{\delta(x_n)}{u(x_n)} \approx \frac{h \frac{A^2 x_n}{2} e^{-Ax_n}}{e^{-Ax_n}} = h \frac{A^2 x_n}{2}.$$

Возьмем  $A = 20$  и будем рассматривать это отношение в точке  $x_n = 1$ . Тогда из условия  $|\delta(1)/u(1)| < 1$  получим

$$h < 0,2 \cdot 10^{-3}.$$

Теперь выясним, какие шаги требуются для интегрирования той же задачи  $u' + Au = 0$  по схеме второго порядка точности

$$\left. \begin{aligned} \frac{u_{n+1} - u_{n-1}}{2h} + Au_n &= 0, \\ u_0 &= 1, \\ u_1 &= 1 - Ah, \end{aligned} \right\} \quad (2)$$

если по-прежнему  $A = 20$  и ставится та же цель удовлетворить условию

$$\left| \frac{\delta(1)}{u(1)} \right| < 1 \quad (3)$$

Решение этой задачи имеет вид (см. равенство (12) из § 8 при  $b = 1$ ):

$$u_n = e^{-Ax_n} + h^2 \left[ \frac{2Ax_n - 3}{12} A^2 e^{-Ax_n} + (-1)^n \frac{A^2}{4} e^{Ax_n} \right] + O(h^3).$$

Ошибка, таким образом, имеет вид

$$\delta(x_n) = h^2 \left[ \frac{2Ax_n - 3}{12} A^2 e^{-Ax_n} + (-1)^n \frac{A^2}{4} e^{Ax_n} \right] + O(h^3).$$

Пренебрежем слагаемым  $O(h^3)$ , выпишем отношение ошибки  $\delta(x_n)$  к точному решению  $u(x_n) = e^{-Ax_n}$  и определим шаг  $h$  из условия (3). Этот шаг окажется столь малым, что если условно принять машинное время расчета по схеме (1) за одну секунду, то по схеме (2) придется затратить около четырех суток!

Дело в том, что оценку практической пригодности той или иной схемы для решения определенной задачи следует делать не только по степени  $h$ , входящей в выражение погрешности, но еще и по коэффициенту при этой степени.

Теперь постараемся понять, как можно судить о пригодности той или иной разностной схемы  $L_h u^{(h)} = f^{(h)}$  из исследования ее устойчивости. Для краткости записей будем считать оператор  $L_h$  линейным. Напомним (см. § 12), что разностная схема называется устойчивой, если при любом  $f^{(h)} \in F_h$  она однозначно разрешима, причем решение  $u^{(h)} \in U_h$  удовлетворяет оценке

$$\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}.$$

Доказывая в § 12 теорему о том, что из аппроксимации и устойчивости следует сходимость, мы получили для погрешности  $z^{(h)} = [u]_h - u^{(h)}$  неравенство

$$\|z^{(h)}\|_{U_h} \leq CC_1 h^k, \quad (4)$$

в котором  $C_1 h^k$  представляет собой оценку величины погрешности аппроксимации:

$$\|L_h [u]_h - f^{(h)}\|_{F_h} \leq C_1 h^k.$$

Пусть ошибка аппроксимации  $C_1 h^k$  мала. Из оценки для  $\|z^{(h)}\|_{U_h}$  видно, что для малости величины  $\|[u]_h - u^{(h)}\|_{U_h}$  надо

еще, чтобы не был слишком велик коэффициент  $C$ , характеризующий устойчивость.

Поэтому, если мы хотим выяснить пригодность той или иной разностной схемы для решения интересующей нас задачи, мало знать, что схема устойчива. Нужно еще знать примерно величину коэффициента  $C$ , суждение о которой можно получить способами, указанными в §§ 14, 15, экспериментальными расчетами или каким-нибудь косвенным образом.

Подсчитаем, например, коэффициент  $C$  для разностных схем (1) и (2) решения задачи  $u' + Au = \varphi(x)$ ,  $u(0) = a$ , о которых шла речь в начале параграфа.

Сначала рассмотрим схему

$$\frac{u_{n+1} - u_n}{h} + Au_n = \varphi_n, \quad n = 0, 1, \dots, N-1,$$

$$u_0 = a$$

при нормах

$$\|u^{(h)}\|_{U_h} = \max_n |u_n|, \quad \|f^{(h)}\|_{F_h} = \max [ |a|, \max_n |\varphi_n| ].$$

Приведем ее к виду

$$\left. \begin{aligned} y_{n+1} &= R_h y_n + h \rho_n, \\ y_0 &\text{ задано,} \end{aligned} \right\}$$

положив  $y_n = u_n$ ,  $R_h = (1 - Ah)$ ,  $\rho_n = \varphi_n$ . Положим  $\|y_n\| = |y_n|$ . Тогда условия (17) из § 14 выполнены:

$$\left. \begin{aligned} \|u^{(h)}\|_{U_h} &\leq C_2 \max_n \|y_n\|, \\ \|\rho_n\| &\leq C_2 \|f^{(h)}\|_{F_h}, \\ \|y_0\| &\leq C_2 \|f^{(h)}\|_{F_h}, \end{aligned} \right\} \quad (5)$$

причем можно положить  $C_2 = 1$ .

Далее, очевидно,  $\|R_h^n\| = (1 - Ah)^n$ . Поэтому можно положить  $C = 2 \max [1, (1 - Ah)^N]$ . Отсюда

$$C = \begin{cases} 2, & \text{если } A > 0, \\ 2(1 - Ah)^N, & \text{если } A \leq 0. \end{cases}$$

Покажем, что число  $C$  нельзя взять существенно меньшим. Нормы выбраны нами так, что выполнены и условия (6) и (7) из § 15:

$$\|u^{(h)}\|_{U_h} \geq M_1 \max_n \|y_n\|, \quad (6)$$

а при  $\varphi_n = 0$  ( $\rho_n = 0$ ) также

$$\|y_0\| \geq M_2 \|f^{(h)}\|_{F_h}, \quad (7)$$

причем можно положить  $M_1 = M_2 = 1$ . Поэтому постоянная  $C$  обязана удовлетворять, как установлено в § 15, оценке  $C \geq M_1 M_2 \max \|R_h^n\|$ :

$$C \geq \begin{cases} 1, & \text{если } A > 0, \\ (1 - Ah)^N, & \text{если } A \leq 0. \end{cases}$$

Теперь оценим постоянную  $C$ , входящую в определение устойчивости  $\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}$ , для разностной схемы (2). Запишем ее в виде

$$\left. \begin{aligned} y_{n+1} &= R_h y_n + h \rho_n, & n &= 0, 1, \dots, \\ y_0 &\text{ задано,} \end{aligned} \right\}$$

положив для этого

$$\begin{aligned} y_n &= \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}, & \rho_n &= \begin{bmatrix} 2\varphi_n \\ 0 \end{bmatrix}, \\ R_h &= \begin{pmatrix} -2Ah & 1 \\ 1 & 0 \end{pmatrix}, & y_0 &= \begin{bmatrix} 1 - Ah \\ 1 \end{bmatrix}. \end{aligned}$$

Выберем нормы

$$\begin{aligned} \|u^{(h)}\|_{U_h} &= \max_n |u_n|, \\ \|f^{(h)}\| &= \left\| \begin{bmatrix} \varphi_n \\ \alpha \\ \beta \end{bmatrix} \right\| = \max [|\alpha|, |\beta|, \max_n |\varphi_n|], \\ \|y_n\| &= \left\| \begin{bmatrix} y_n^{(1)} \\ y_n^{(2)} \end{bmatrix} \right\| = \max [ |y_n^{(1)}|, |y_n^{(2)}| ]. \end{aligned}$$

Тогда выполнены условия (5)–(7), причем  $C_2 = M_1 = M_2 = 1$ . Поэтому в силу п. 3 из § 14 в качестве постоянной  $C$  можно взять число  $C = 2C_2^2 \max_n \|R_h^n\| = 2 \max_n \|R_h^n\|$ ; но в силу (6'') из § 15 нельзя более чем вдвое уменьшить его: заведомо должно быть

$$C \geq M_1 M_2 \max_n \|R_h^n\| = \max_n \|R_h^n\|.$$

Оценка сверху для величины  $\max \|R_h^n\|$  была получена в § 14:

$$\|R_h^n\| \leq \left\| \begin{pmatrix} -2Ah & 1 \\ 1 & 0 \end{pmatrix} \right\|^N \leq (1 + 2|A|h)^N.$$



Поэтому можно положить

$$C = 2e^{2|A|} \geq 2(1 + 2|A|h)^{1/h}.$$

Оценку снизу для  $\max \|R_h^n\|$  получим из условия  $\|R_h^n\| \geq |\lambda|^n$ , где  $\lambda$  — большее по модулю из двух собственных чисел матрицы  $R_h$ . Решая уравнение  $\det(R_h - \lambda E) = 0$ , найдем собственные числа:

$$\lambda_2 = 1 - Ah + \frac{A^2 h^2}{2} + o(h^2) = 1 - Ah + O(h^2),$$

$$\lambda_1 = -1 - Ah - \frac{A^2 h^2}{2} + o(h^2) = -1 - Ah + O(h^2),$$

так что

$$\max(|\lambda_1|, |\lambda_2|) = 1 + |A|h + O(h^2),$$

$$\max_n \|R_h^n\| = (1 + |A|h)^{1/h} + O(h).$$

Поэтому найденную выше постоянную  $C = 2e^{2|A|}$  заведомо нельзя заменить числом меньшим, чем  $(1 + |A|h)^{1/h} \approx e^{|A|}$ , т. е. нельзя существенно уменьшить.

При  $A=20$  для первой схемы  $C=2$ , а для второй заведомо  $C \geq e^{20} \geq 10^8$ .

При  $A \approx 1$  или  $A < 0$  в свойствах устойчивости обеих схем нет коренного различия: постоянная  $C$  для обеих схем примерно одинакова.

Легко понять механизм, в силу которого при  $A \gg 1$  постоянная  $C$  для второй схемы много больше единицы, в то время как для первой  $C=2$ .

Общее решение однородного уравнения  $u_{n+1} - (1 - Ah)u_n = 0$ , соответствующего схеме (1), есть  $\bar{u}_n = \alpha q^n$ , где  $q$  — корень характеристического уравнения  $q - (1 - Ah) = 0$ ,  $q = 1 - Ah$  (рис. 4). Общее решение однородного уравнения

$$u_{n+1} + 2Ahu_n - u_{n-1} = 0,$$

соответствующего схеме (2), есть

$$u_n = \alpha q_1^n + \beta q_2^n,$$

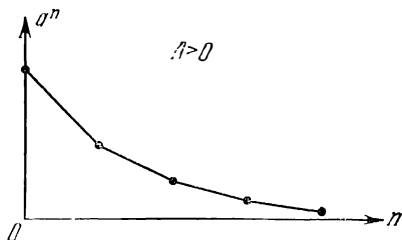


Рис. 4.

Где  $q_1$  и  $q_2$  — корни характеристического уравнения

$$q^2 + 2Ahq - 1 = \det(R_h - qE) = 0,$$

$$q_1 = 1 - Ah + \frac{A^2 h^2}{2} + o(h^2),$$

$$q_2 = -1 - Ah - \frac{A^2 h^2}{2} + o(h^2).$$

Корень  $q_1$  «похож» на корень  $q = 1 - Ah$ , и ему соответствует решение  $q_1^n$ , похожее на решение  $q^n$  первого уравнения. Но «паразитический» корень  $q_2 = -1 - Ah + O(h^2)$  дает быстро возрастающее «паразитическое» решение  $q_2^n$  (рис. 5), которое и обуславливает большое значение  $C$ .

При отрицательных  $A$  будет  $q > 1$ ,  $q_1 > 1$ ,  $|q_2| < 1$ . Решения

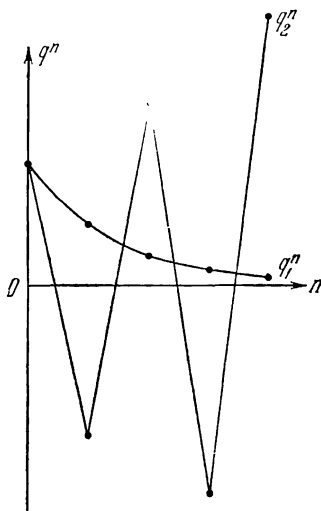


Рис. 5.

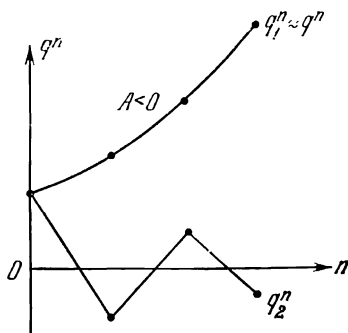


Рис. 6.

$q^n$  и  $q_1^n$ , соответствующие корням  $q$  и  $q_1$ , примерно одинаково быстро растут, а паразитическое решение  $q_2^n$  затухает, не оказывая влияния на характер устойчивости второй схемы (рис. 6).

Отметим, что большое значение  $C$  при  $A \ll 0$  неизбежно для любой разностной схемы, приближающей задачу  $u' + Au = 0$ ,  $u(0) = a$ . В самом деле, при малых  $h$  решение устойчивой разностной задачи похоже на решение дифференциальной задачи, к которому оно при  $h \rightarrow 0$  сходится. Но решение дифференциальной задачи  $u = u_0 e^{-Ax}$  таково, что  $\max |u(x)| = |u_0| e^{-Ax}$ , т. е.  $\max |u(x)|$  в большое число  $e^{-A}$  раз превосходит модуль  $|u_0|$  начального значения  $u_0$ .

Мы должны еще отметить, что большой коэффициент  $C$  ведет не только к необходимости расчетов с мелким шагом, но и

к большому числу десятичных знаков, с которым приходится вести вычисления.

В самом деле, в § 16 мы показали, что ошибки округления можно включить в ошибки при задании правых частей, которые оцениваются величиной  $C_1 h^k$ . Увеличение этих ошибок вызывает увеличение коэффициента  $C_1$ , что при большом  $C$  в силу (4) может катастрофически сказаться на точности результата.

В заключение этого параграфа мы хотели бы еще предостеречь читателя от ложного впечатления о схемах второго порядка точности, которое могло у него создаться из рассмотренного примера. Мы вовсе не хотим опорочить все такие схемы, описывая недостатки одной из них. Читателю будет очень полезно провести самостоятельное изучение схемы второго порядка точности вида

$$\left. \begin{aligned} \frac{u_{n+1} - u_n}{h} + A \frac{u_{n+1} + u_n}{2} &= 0, \\ u_0 &= 1. \end{aligned} \right\}$$

Стремясь добиться, чтобы при  $A = 1$  погрешность  $\delta(1)$  была меньше, чем  $u(1) = e^{-A}$ , он убедится, что эта схема накладывает менее жесткое ограничение на шаг  $h$ , чем схема первого порядка точности (1).

Кроме того, советуем прикинуть, с каким шагом надо интегрировать задачу  $u' + u = 0$ ,  $u(0) = 1$ , чтобы получить в  $u(1)$  погрешность не более  $10^{-5}$ . Если читатель проделает эту прикидку для рассмотренных в начале параграфа разностных схем (1) и (2), то увидит, что схема первого порядка точности (1) требует значительно более мелкого шага, чем схема второго порядка точности (2).

Таким образом, выгодность или невыгодность той или иной схемы зависит не только от нее самой, но и от задачи, к которой она применяется.

## § 18. Прием исследования устойчивости нелинейных задач

Способы исследования устойчивости, изложенные в §§ 14 и 15, были непосредственно приспособлены для разностных уравнений с постоянными коэффициентами. Поэтому может показаться, что нельзя использовать приведенный в этих параграфах материал для анализа схем интегрирования даже уравнений  $\frac{du}{dx} = G(x, u)$  с довольно общей функцией  $G$ . Однако это не так.

Пусть интересующая нас интегральная кривая уравнения

$$\frac{du}{dx} = G(x, u) \quad (1)$$

проходит через точку с координатами  $x = x_0$ ,  $u = u_0$ . Вблизи этой точки имеем

$$G(x, u) \approx \frac{\partial G(x, u)}{\partial u} (u - u_0) + \frac{\partial G(x, u)}{\partial x} (x - x_0) + G(x_0, u_0), \quad (2)$$

и поэтому уравнение (1) с определенной точностью может быть заменено более простым:

$$\frac{du}{dx} - Au = \varphi(x), \quad (3)$$

где

$$A = \left. \frac{\partial G(x, u)}{\partial u} \right|_{\substack{x=x_0 \\ u=u_0}},$$

$$\varphi(x) = G(x_0, u_0) + \left. \frac{\partial G(x, u)}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}} (x - x_0) - u_0 \left. \frac{\partial G(x, u)}{\partial u} \right|_{\substack{x=x_0 \\ u=u_0}}.$$

Естественно, что схемы, которые мы хотим применять для нахождения нужного решения, должны хорошо интегрировать уравнение (3), аппроксимирующее уравнение (1) вблизи некоторой точки, через которую проходит интегральная кривая. Конечно, для разных точек этой кривой величина коэффициента  $A$ , полученного описанным способом линеаризации исходного уравнения, будет различной. Поэтому, отбирая ту или иную разностную схему, мы должны будем ее проверить на уравнении (3) не с одним значением  $A$ , а с целым набором таких значений, достаточно полно описывающим диапазон изменения  $\partial G/\partial u$  вдоль интегральной кривой. В подавляющем большинстве практически встречающихся случаев такого исследования оказывается достаточно для выявления всех существенных недостатков и достоинств схемы, относящихся к характеру сходимости полученных по ней приближенных решений.

Такой же метод построения модельных уравнений может быть применен и в случаях систем уравнений и уравнений высших порядков.

На практике вычисление решения задачи Коши для обыкновенных дифференциальных уравнений без особенностей обычно производится по одной-двум довольно универсальным, хорошо апробированным схемам, для которых на современных вычислительных машинах имеются стандартные программы.

Если приходится с очень большой точностью решать задачи специального вида, то применяются многочисленные специальные схемы, приспособленные именно для этих задач, но уступающие универсальным схемам при решении другого круга задач.

УПОТРЕБИТЕЛЬНЫЕ РАЗНОСТНЫЕ СХЕМЫ

§ 19. Схемы Рунге — Кутты и Адамса

Изложим здесь некоторые употребительные разностные схемы решения задачи Коши для дифференциального уравнения первого порядка

$$\left. \begin{aligned} \frac{du}{dx} - G(x, u) = 0, \quad 0 \leq x \leq 1, \\ u(0) = a. \end{aligned} \right\} \quad (1)$$

В конце параграфа эти схемы будут перенесены на системы дифференциальных уравнений первого порядка, к которым сводится общий случай уравнений и систем любого порядка.

Выберем на отрезке  $0 \leq x \leq 1$  сетку точек

$$0 = x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N = 1, \quad x_n = nh, \quad h = 1/N,$$

и будем составлять разностные схемы для приближенного отыскания таблицы  $[u]_h$  значений решения  $u(x)$  на выбранной сетке.

С простейшей употребительной схемой мы уже встречались. Это — схема Эйлера

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{n+1} - u_n}{h} - G(x_n, u_n) = 0, & n = 0, 1, \dots, N-1, \\ u_0 = a, \end{cases} \quad (2)$$

обладающая первым порядком аппроксимации (и точности).

Вычисления по этой схеме имеют простой геометрический смысл. Если  $u_n$  уже вычислено, то вычисление

$$u_{n+1} = u_n + hG(x_n, u_n)$$

равносильно сдвигу из точки  $(x_n, u_n)$  в точку  $(x_{n+1}, u_{n+1})$  на плоскости  $Oxu$  по касательной к интегральной кривой  $u = u(x)$  дифференциального уравнения  $u' = G(x, u)$ , проходящей через точку  $(x_n, u_n)$ .

Среди схем более высокого порядка аппроксимации наиболее употребительны различные варианты схем Рунге — Кутта и Адамса, которые мы опишем и сопоставим.

**1. Схемы Рунге — Кутта.** Пусть значение  $u_n$  приближенного решения в точке  $x_n$  уже найдено и требуется вычислить  $u_{n+1}$  в точке  $x_{n+1} = x_n + h$ . Задаем целое  $l$  и выписываем выражения

$$\begin{aligned}
 k_1 &= G(x_n, u_n), \\
 k_2 &= G(x_n + \alpha h, u_n + \alpha h k_1), \\
 k_3 &= G(x_n + \beta h, u_n + \beta h k_2), \\
 &\dots \\
 k_l &= G(x_n + \gamma h, u_n + \gamma h k_{l-1}).
 \end{aligned}$$

Затем полагаем

$$L_h u^{(h)} = \begin{cases} \frac{u_{n+1} - u_n}{h} - (p_1 k_1 + \dots + p_l k_l) = 0, & n = 0, 1, \dots, N-1, \\ u_0 = a. \end{cases}$$

Коэффициенты  $\alpha, \beta, \dots, \gamma, p_1, p_2, \dots, p_l$  подбираем так, чтобы получить при заданном  $l$  аппроксимацию возможно более высокого порядка. Зная  $u_n$ , можно вычислить  $k_1, \dots, k_l$ , а затем

$$u_{n+1} = u_n + h(p_1 k_1 + \dots + p_l k_l).$$

Простейшей схемой Рунге — Кутта является схема Эйлера ( $l = 1$ ).

Схема Рунге — Кутта

$$L_h u^{(h)} = \begin{cases} \frac{u_{n+1} - u_n}{h} - \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) = 0, \\ n = 0, 1, \dots, N-1, \\ u_0 = a, \end{cases} \quad (3)$$

где

$$\begin{aligned}
 k_1 &= G(x_n, u_n), \\
 k_2 &= G\left(x_n + \frac{h}{2}, u_n + \frac{k_1 h}{2}\right), \\
 k_3 &= G\left(x_n + \frac{h}{2}, u_n + \frac{k_2 h}{2}\right), \\
 k_4 &= G(x_n + h, u_n + k_3 h),
 \end{aligned}$$

имеет четвертый порядок аппроксимации.

## Схема Рунге — Кутта

$$L_h u^{(h)} = \begin{cases} \frac{u_{n+1} - u_n}{h} - \left[ \frac{2\alpha - 1}{2\alpha} k_1 + \frac{1}{2\alpha} k_2 \right] = 0, \\ n = 0, 1, \dots, N - 1, \\ u_0 = a, \end{cases} \quad (4)$$

где

$$k_1 = G(x_n, u_n), \quad k_2 = G(x_n + \alpha h, u_n + \alpha h k_1),$$

при любом фиксированном  $\alpha$  имеет второй порядок аппроксимации.

Мы докажем только утверждение о схеме (4). Доказательство утверждения о схеме (3) аналогично, но более громоздко.

Решение  $u(x)$  уравнения  $u' = G(x, u)$  удовлетворяет тождествам

$$\begin{aligned} \frac{du}{dx} &\equiv G(x, u(x)), \\ \frac{d^2u}{dx^2} &\equiv \frac{d}{dx} G(x, u) = \frac{\partial G}{\partial x} + \frac{\partial G}{\partial u} G. \end{aligned}$$

Поэтому из формулы Тейлора

$$\frac{u(x_n + h) - u(x_n)}{h} = u'(x_n) + \frac{h}{2} u''(x_n) + O(h^2)$$

для решения  $u(x)$  следует равенство

$$\frac{u(x_{n+1}) - u(x_n)}{h} - \left[ G + \frac{h}{2} \left( \frac{\partial G}{\partial x} + \frac{\partial G}{\partial u} G \right) \right]_{\substack{x=x_n \\ u=u(x_n)}} = O(h^2). \quad (5)$$

Но, разлагая по  $h$  функцию двух переменных по формуле Тейлора и удерживая члены первой степени, получим

$$\begin{aligned} \left. \frac{2\alpha - 1}{2\alpha} k_1 + \frac{1}{2\alpha} k_2 \right|_{\substack{x=x_n \\ u=u(x_n)}} &= \frac{2\alpha - 1}{2\alpha} G + \frac{1}{2\alpha} G(x + \alpha h, u + \alpha h G) \Big|_{\substack{x=x_n \\ u=u(x_n)}} = \\ &= \frac{2\alpha - 1}{2\alpha} G + \frac{1}{2\alpha} \left[ G + \frac{\partial G}{\partial x} \alpha h + \frac{\partial G}{\partial u} \alpha h G + O(h^2) \right]_{\substack{x=x_n \\ u=u(x_n)}} = \\ &= G + \frac{h}{2} \left( \frac{\partial G}{\partial x} + \frac{\partial G}{\partial u} G \right) \Big|_{\substack{x=x_n \\ u=u(x_n)}} + O(h^2). \quad (6) \end{aligned}$$

Поэтому при подстановке в левую часть равенства (4) вместо  $u_n$  и  $u_{n+1}$  соответственно значений  $u(x_n)$  и  $u(x_{n+1})$  решения  $u(x)$  получится выражение, совпадающее с левой частью равенства (5) с точностью до  $O(h^2)$ . Следовательно, это выражение имеет второй порядок относительно  $h$ . Поскольку значение  $u_0 = a$  задано точно, этим завершается доказательство того, что схема (4) имеет второй порядок аппроксимации.

Для получения  $u_{n+1}$  по схеме Рунге — Кутта при уже известном  $u_n$  приходится  $l$  раз вычислять значения функции  $G(x, u)$ . Эти значения больше не используются.

**2. Схемы Адамса.** В схемах Адамса, одну из разновидностей которых мы сейчас опишем, для вычисления каждого следующего значения  $u_{n+1}$  достаточно дополнительно вычислить значение  $G(x, u)$  лишь в одной точке независимо от порядка аппроксимации. Кроме того, приходится проделать небольшое число вычитаний и сложений, которые требуют во много раз меньше времени, чем каждое вычисление сколько-нибудь сложной функции  $G(x, u)$ .

Обозначим

$$\begin{aligned}\nabla f_n &= f_n - f_{n-1}, \\ \nabla^2 f_n &= \nabla(\nabla f_n) = \nabla f_n - \nabla f_{n-1} = f_n - 2f_{n-1} + f_{n-2}, \\ \nabla^3 f_n &= \nabla \nabla^2 f_n = f_n - 3f_{n-1} + 3f_{n-2} - f_{n-3}\end{aligned}$$

и положим  $G_n = G(x_n, u_n)$ . Выпишем несколько разностных уравнений, используемых в схемах Адамса для вычисления  $u_{n+1}$ , если  $u_n, u_{n-1}, \dots$  уже вычислены:

$$\frac{u_{n+1} - u_n}{h} - G_n = 0, \quad n = 0, 1, \dots, N-1, \quad (7)$$

$$\frac{u_{n+1} - u_n}{h} - G_n - \frac{1}{2} \nabla G_n = 0, \quad n = 1, 2, \dots, N-1, \quad (8)$$

$$\frac{u_{n+1} - u_n}{h} - G_n - \frac{1}{2} \nabla G_n - \frac{5}{12} \nabla^2 G_n = 0, \quad n = 2, 3, \dots, N-1, \quad (9)$$

$$\frac{u_{n+1} - u_n}{h} - G_n - \frac{1}{2} \nabla G_n - \frac{5}{12} \nabla^2 G_n - \frac{3}{8} \nabla^3 G_n = 0 \quad (10)$$

$$n = 3, 4, \dots, N-1.$$

Первое из этих уравнений — разностное уравнение Эйлера. При подстановке в левые части уравнений (7) — (10) вместо  $u_{n+1}, u_n, u_{n-1}, \dots$  значений  $u((n+1)h), u(nh), \dots$  точного решения  $u(x)$  в равенствах (7) — (10) возникают невязки порядка  $h, h^2, h^3$  и  $h^4$  соответственно.

Формулы Адамса получаются следующим образом. Пусть  $u(x)$  — решение уравнения

$$\frac{du}{dx} = G(x, u).$$

Обозначим

$$G(x, u(x)) \equiv F(x).$$



Тогда

$$u(x_n + h) - u(x_n) = \int_{x_n}^{x_n+h} u' dx = \int_{x_n}^{x_n+h} F(x) dx.$$

Из теории интерполяции функций известно, что существует один и только один многочлен  $P_k(x, F)$  степени не выше  $k$ , принимающий в  $(k+1)$ -й точке  $x_n, x_{n-1}, \dots, x_{n-k}$  заданные значения  $F(x_n), F(x_{n-1}), \dots, F(x_{n-k})$  соответственно. Этот многочлен  $P_k(x, F)$  в случае достаточно гладкой функции  $F(x)$  уклоняется от  $F(x)$  на отрезке  $x_n \leq x \leq x_n + h$  на величину порядка  $h^{k+1}$ , так что

$$\max |P_k(x, F) - F(x)| = O(h^{k+1}). \quad (11)$$

Разностная формула Адамса имеет вид

$$\frac{u_{n+1} - u_n}{h} - \frac{1}{h} \int_{x_n}^{x_n+h} P_k(x, F) dx = 0. \quad (12)$$

При подстановке в левую часть вместо

$$u_n, u_{n+1}, G(x_{n-s}, u_{n-s})$$

соответственно значений

$$u(x_n), u(x_{n+1}), G(x_{n-s}, u(x_{n-s}))$$

получим невязку порядка  $h^{k+1}$ :

$$\begin{aligned} & \left| \frac{u(x_n + h) - u(x_n)}{h} - \frac{1}{h} \int_{x_n}^{x_n+h} P_k(x, F) dx \right| = \\ & = \left| \left[ \frac{u(x_n + h) - u(x_n)}{h} - \frac{1}{h} \int_{x_n}^{x_n+h} F(x) dx \right] + \frac{1}{h} \int_{x_n}^{x_n+h} [F(x) - P_k(x, F)] dx \right| \leq \\ & \leq 0 + \max |F(x) - P_k(x, F)| = O(h^{k+1}) \end{aligned}$$

При  $k=0$  интерполяционный многочлен

$$P_0(x, F) = G(x_n, u_n) = \text{const}$$

и формула (12) превращается в (7).

При  $k=1$

$$P_1(x, F) = \frac{1}{h} [(x - x_{n-1}) G_n - (x - x_n) G_{n-1}].$$

Далее,

$$\begin{aligned} \frac{1}{h} \int_{x_n}^{x_n+h} P_1(x, F) dx &= \frac{1}{h^2} \frac{(x - x_{n-1})^2}{2} \Big|_{x_n}^{x_n+h} G_n - \frac{1}{h^2} \frac{(x - x_n)^2}{2} \Big|_{x_n}^{x_n+h} G_{n-1} = \\ &= \frac{1}{h^2} \left( \frac{4h^2}{2} - \frac{h^2}{2} \right) G_n - \frac{1}{h^2} \frac{h^2}{2} G_{n-1} = G_n + \frac{1}{2} \nabla G_n. \end{aligned}$$

Следовательно, формула (12) превращается в (8). Аналогично при  $k = 2$  и  $k = 3$  из (12) получаются формулы (9) и (10) соответственно.

Для вычисления по схеме (7) достаточно знать  $u_0 = a$ . Для того чтобы начать вычисления по схеме (8), надо заранее знать, кроме  $u_0 = a$ , также еще и  $u_1$ . Схема (9) требует использования  $u_0$ ,  $u_1$  и  $u_2$ , а схема (10) — четырех значений:  $u_0$ ,  $u_1$ ,  $u_2$  и  $u_3$ . Эти значения могут быть найдены по схеме Рунге — Кутта; с помощью схемы Эйлера с мелким шагом; с помощью разложения решения в окрестности точки  $x = 0$  в степенной ряд. Нестандартное начало счета является одним из недостатков схем Адамса по сравнению со схемами Рунге — Кутта. Отмечавшимся ранее достоинством схем Адамса является то, что для вычисления  $u_{n+1}$  нужно, в дополнение к уже вычисленным в процессе отыскания  $u_n$ ,  $u_{n-1}$ , ... значениям  $G_s$ ,  $\nabla G_s$ , ...,  $\nabla^k G_s$ , вычислить только одно значение функций  $G_n = G(x_n, u_n)$  и произвести некоторое число вычитаний для вычисления  $\nabla G_n$ , ...,  $\nabla^k G_n$ .

Итак, преимущества методов Адамса перед методами Рунге — Кутта заключаются в меньшей трудоемкости вычислений на один шаг. Основные недостатки — нестандартное начало счета, невозможность (без усложнения формул) в процессе счета изменить, начиная с какой-то точки  $x_n$ , шаг  $h$ ,  $x_{n+1} = x_n + h$ , с которым ведутся вычисления. Последнее обстоятельство существенно в тех случаях, когда решение и его производные на некоторых участках меняются быстро, а на других изменяются медленно.

Схема Рунге — Кутта, если такого рода обстоятельства выясняются в процессе счета, может, например, по заданной подпрограмме автоматически уменьшить шаг или увеличить шаг на гладких участках, чтобы не производить лишней работы.

По-видимому, наиболее рационально использование обоих методов — Рунге — Кутта и Адамса — с автоматическим переходом с одного из них на другой в процессе счета. При этом начинать счет надо по схеме Рунге — Кутта. В программе должен быть предусмотрен автоматический выбор шага, при котором расчет ведется с нужной точностью. При этом программа выбора шага должна предусматривать некоторый «консерватизм при выборе шага: диктовать изменение шага только в случае,

если требуется «довольно сильно» его изменить. Если оказывается, что при вычислении нескольких последовательных значений  $u_n$  по схеме Рунге — Кутта не происходит изменения шага, целесообразен автоматический переход на счет по более экономной схеме Адамса. Как только вновь появляется необходимость изменить шаг, программа расчета должна переходить на вычисления по схеме Рунге — Кутта и т. д.

Для контроля правильности выбора шага обычно параллельно ведут вычисления с некоторым заданным и вдвое меньшим шагом. В пределах требуемой точности решения должны совпадать. В противном случае надо вести вычисления с более мелким шагом. Нужно также предусмотреть пробу на возможность увеличить шаг.

**3. Замечания об устойчивости.** Для задачи  $u' + Au = 0$ , линейной и с постоянным коэффициентом  $A$ , схемы Рунге — Кутта после исключения  $k_1, k_2, \dots$  окажутся схемами первого порядка,

$$u_{n+1} - a(h)u_n = 0.$$

Корень характеристического уравнения  $\lambda - a(h) = 0$  равен  $\lambda = a(h)$ .

В случае  $u_n = u(x_n)$  для  $u_{n+1}$  получается задание, совпадающее с точным решением  $u(x_n + h)$  с точностью до  $h^{p+1}$ , где  $p$  — порядок аппроксимации. Поскольку

$$u(x_n + h) = u(x_n)e^{-Ah} = u(x_n) \left( 1 - Ah + \frac{A^2h^2}{2} - \dots \right),$$

а

$$u_{n+1} = a(h)u_n,$$

то

$$\lambda = a(h) = e^{-Ah} + O(h^{p+1}).$$

Таким образом,

$$|\lambda(h)| < 1 + ch.$$

Степени  $\lambda^n(h)$  ведут себя «правильно»: они растут, если  $A < 0$  и решение дифференциального уравнения растет. Они убывают, если  $A > 0$  и решение  $e^{-Ax}$  убывает.

В случае схемы Адамса (8)

$$\frac{u_{n+1} - u_n}{h} + Au_n + \frac{A}{2}(u_n - u_{n-1}) = 0 \quad (13)$$

характеристическое уравнение имеет вид

$$\lambda - \left( 1 - \frac{3Ah}{2} \right) \lambda - \frac{Ah}{2} = 0.$$

Отсюда

$$\lambda = \frac{1}{2} \left( 1 - \frac{3Ah}{2} \right) \pm \frac{1}{2} \sqrt{\left( 1 - \frac{3Ah}{2} \right)^2 + 2Ah},$$

$$\lambda_1 = 1 - Ah + O(h^2),$$

$$\lambda_2 = O(h),$$

Таким образом, решение  $u_n = \lambda_1^n$  ведет себя при измельчении  $h$ , как  $u(x_n) = e^{-A_n h}$ , а «паразитическое» решение  $\lambda_2^n$ , вызванное выбором разностного уравнения второго порядка, стремится к нулю, так как  $|\lambda_2| = O(h)$ , и на устойчивость не влияет.

Читателю полезно сравнить схему (13) со схемой второго порядка (2) из § 17:

$$\frac{u_{n+1} - u_{n-1}}{2h} + Au_n = 0.$$

Для нее

$$\lambda_1 = 1 - Ah + \frac{A^2 h^2}{2} + O(h^3), \quad \lambda_2 = -1 - Ah + O(h^2).$$

«Паразитический» корень  $\lambda_2$  при положительном  $A$  по модулю больше корня  $\lambda_1$ , что и приводило к большой постоянной в оценке устойчивости для этой схемы и к практической непригодности ее при больших  $A$ , установленной в § 17.

**4. Обобщение на системы уравнений.** Все описанные схемы численного решения задачи Коши для дифференциального уравнения первого порядка (1) автоматически переносятся на системы уравнений первого порядка. Для этого в записи (1)

$$\left. \begin{aligned} \frac{du}{dx} - G(x, u) &= 0, \\ u(0) &= a \end{aligned} \right\}$$

надо понимать под  $u(x) = \bar{u}(x)$  и  $G(x, u) = \bar{G}(x, \bar{u})$  вектор-функции и под  $a = \bar{a}$  заданный вектор. Тогда схемы Рунге — Кутты (3), (4) и схемы Адамса (7) — (10) сохранят смысл и останутся применимыми.

Например, система уравнений

$$\left. \begin{aligned} \frac{dv}{dx} - (x + v^2 + \sin w) &= 0, \\ \frac{dw}{dx} + xv w &= 0, \\ v(0) &= a_1, \\ w(0) &= a_2 \end{aligned} \right\}$$

запишется в форме

$$\left. \begin{aligned} \frac{d\bar{u}}{dx} - G(x, \bar{u}) &= 0, \\ \bar{u}(0) &= \bar{a}, \end{aligned} \right\}$$

если положить

$$\begin{aligned} \bar{u}(x) &= \begin{pmatrix} v(x) \\ w(x) \end{pmatrix}, \\ \bar{G}(x, \bar{u}) &= \begin{pmatrix} x + v^2 + \sin w \\ -xvw \end{pmatrix}, \\ \bar{a} &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \end{aligned}$$

Формула для  $\bar{u}_{n+1}$  в схеме Эйлера

$$\bar{u}_{n+1} = \bar{u}_n + h\bar{G}(x_n, \bar{u}_n)$$

подробно запишется так:

$$\left. \begin{aligned} v_{n+1} &= v_n + h(x_n + v_n^2 + \sin w_n), \\ w_{n+1} &= w_n + h(-x_n v_n w_n). \end{aligned} \right\}$$

Все рассуждения о порядке аппроксимации, изложенные мелким шрифтом (стр. 159), тоже сохраняются. При этом в формуле (6) под производной вектора  $G(G_1, \dots, G_k)$  по вектору  $u(u_1, \dots, u_k)$ ,  $\frac{\partial G}{\partial u}$ , надо понимать матрицу

$$\begin{pmatrix} \frac{\partial G_1}{\partial u_1} & \cdots & \frac{\partial G_1}{\partial u_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_k}{\partial u_1} & \cdots & \frac{\partial G_k}{\partial u_k} \end{pmatrix}.$$

Произвольная система дифференциальных уравнений, разрешенных относительно старших производных, сводится к системе уравнений первого порядка

$$\frac{d\bar{u}}{dx} = \bar{G}(x, \bar{u})$$

путем замены искоемых функций. Как это делается, ясно из следующего примера. Система

$$\left. \begin{aligned} \frac{d^2v}{dx^2} + \sin(xv' + v^2 + w) &= 0, \\ \frac{dw}{dx} + \sqrt{x^2 + v^2 + (v')^2 + w^2} &= 0, \\ v(0) &= a, \\ v'(0) &= b, \\ w(0) &= c \end{aligned} \right\}$$

приводится к требуемому виду, если положить

$$\begin{aligned} u_1(x) &= v(x), \\ u_2(x) &= \frac{dv}{dx}, \\ u_3(x) &= w(x). \end{aligned}$$

Получим

$$\left. \begin{aligned} \frac{du_1}{dx} - u_2 &= 0, \\ \frac{du_2}{dx} + \sin(xu_2 + u_1^2 + u_3) &= 0, \\ \frac{du_3}{dx} + \sqrt{x^2 + u_1^2 + u_2^2 + u_3^2} &= 0, \\ u_1(0) &= a, \\ u_2(0) &= b, \\ u_3(0) &= c. \end{aligned} \right\}$$

**З а м е ч а н и е.** Разработаны разностные схемы типа схем Рунге — Кутты, применимые непосредственно для уравнения второго порядка и не требующие предварительного сведения этих уравнений к системам первого порядка.

## § 20. Методы решения краевых задач

Примером краевой задачи является задача

$$\left. \begin{aligned} y'' &= f(x, y, y'), & 0 \leq x \leq 1, \\ y(0) &= Y_0, & y(1) = Y_1 \end{aligned} \right\} \quad (1)$$

с граничными условиями на обоих концах отрезка  $0 \leq x \leq 1$ , на котором надо найти решение  $y = y(x)$ . На этом примере мы схематически изложим некоторые способы численного решения краевых задач.

**1. Метод стрельбы.** В § 19 указаны удобные способы численного решения задачи Коши, т. е. задачи вида

$$\left. \begin{aligned} y'' &= f(x, y, y'), & 0 \leq x \leq 1, \\ y(0) &= Y_0, & \left. \frac{dy}{dx} \right|_{x=0} = \operatorname{tg} \alpha, \end{aligned} \right\} \quad (2)$$

где  $Y_0$  — ордината точки  $(0, Y_0)$ , из которой выходит интегральная кривая, а  $\alpha$  — угол наклона интегральной кривой к оси  $Ox$  при выходе из точки  $(0, Y_0)$  (рис. 7, а). При фиксированном  $Y_0$  решение задачи (2) имеет вид  $y = y(x, \alpha)$ . При  $x = 1$  решение  $y(x, \alpha)$  зависит только от  $\alpha$ :

$$y(x, \alpha)|_{x=1} = y(1, \alpha).$$

Используя указанное замечание о решении задачи Коши (2), мы можем теперь переформулировать задачу (1) следующим образом: найти такой угол  $\alpha = \alpha^*$ , при котором интегральная кривая, выходящая из точки  $(0, Y_0)$  под углом  $\alpha$  к оси абсцисс, попадет в точку  $(1, Y_1)$ :

$$y(1, \alpha) = Y_1. \quad (3)$$

Решение задачи (2) при этом  $\alpha = \alpha^*$  совпадает с искомым решением задачи (1). Дело сводится, таким образом, к решению

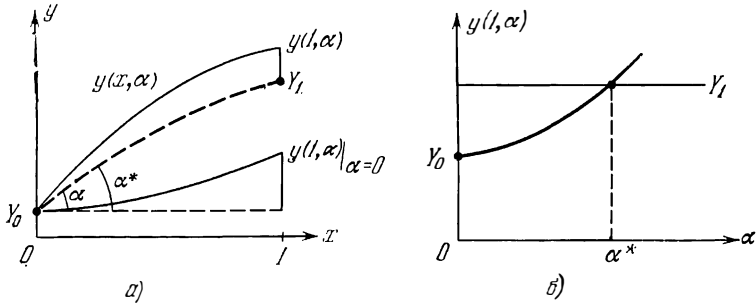


Рис. 7.

уравнения (3) (рис. 7, б). Уравнение (3) есть уравнение вида  $F(\alpha) = 0$ , где  $F(\alpha) = y(1, \alpha) - Y_1$ . Оно отличается от привычных уравнений лишь тем, что функция  $F(\alpha)$  задана не аналитическим выражением, а с помощью алгоритма решения задачи (2).

Сведение решения краевой задачи (1) к решению задачи Коши (2) и составляет сущность метода стрельбы.

Для решения уравнения (3) можно использовать метод деления отрезка пополам, метод хорд, метод касательных (метод Ньютона) и т. д. Например, при использовании метода деления отрезка пополам мы задаем  $\alpha_0$  и  $\alpha_1$  так, чтобы разности

$$y(1, \alpha_0) - Y_1 \quad \text{и} \quad y(1, \alpha_1) - Y_1$$

имели разные знаки. Затем полагаем

$$\alpha_2 = \frac{\alpha_0 + \alpha_1}{2}.$$

Вычисляем  $y(1, \alpha_2)$ . Вычисляем затем  $\alpha_3$  по одной из формул

$$\alpha_3 = \frac{\alpha_1 + \alpha_2}{2} \quad \text{или} \quad \alpha_3 = \frac{\alpha_0 + \alpha_2}{2}$$

в зависимости от того, имеют ли разности

$$y(1, \alpha_2) - Y_1 \quad \text{и} \quad y(1, \alpha_1) - Y_1$$

соответственно разные или одинаковые знаки. Затем вычисляем  $y(1, \alpha_3)$ . Процесс продолжается до тех пор, пока не будет достигнута требуемая точность,  $|y(1, \alpha_n) - Y_1| < \varepsilon$ .

В случае использования метода хорд задаем  $\alpha_0$  и  $\alpha_1$ , а затем последующие  $\alpha_i$  вычисляем по рекуррентной формуле

$$\alpha_{n+1} = \alpha_n - \frac{F(\alpha_n)}{F(\alpha_n) - F(\alpha_{n-1})} (\alpha_n - \alpha_{n-1}), \quad n = 1, 2, \dots$$

Метод стрельбы, сводящий решение краевой задачи (1) к вычислению решений задачи Коши (2), хорошо работает в том случае, если решение  $y(x, \alpha)$  «не слишком сильно» зависит от  $\alpha$ . В противном случае он становится вычислительно неустойчивым, даже если решение задачи (1) зависит от входных данных «умеренно».

Поясним взятые в кавычки слова на примере следующей линейной краевой задачи:

$$\left. \begin{aligned} y'' - a^2 y &= 0, & 0 \leq x \leq 1, \\ y(0) &= Y_0, & y(1) = Y_1 \end{aligned} \right\} \quad (1')$$

при постоянном  $a^2$ . Выпишем решение этой задачи:

$$y(x) = \frac{e^{-ax} - e^{-a(2-x)}}{1 - e^{-2a}} Y_0 + \frac{e^{-a(1-x)} - e^{-a(1+x)}}{1 - e^{-2a}} Y_1.$$

Коэффициенты при  $Y_0$  и  $Y_1$  с ростом  $a$  остаются ограниченными на отрезке  $0 \leq x \leq 1$  функциями; при всех  $a > 0$  они не превосходят единицу. Поэтому небольшие ошибки при задании  $Y_0$  и  $Y_1$  ведут к столь же небольшим погрешностям в решении. Рассмотрим теперь задачу Коши

$$\left. \begin{aligned} y'' - a^2 y &= 0, & 0 \leq x \leq 1, \\ y(0) &= Y_0, & y'(0) = \operatorname{tg} \alpha. \end{aligned} \right\} \quad (2')$$

Ее решение имеет вид

$$y(x) = \frac{aY_0 + \operatorname{tg} \alpha}{2a} e^{ax} + \frac{aY_0 - \operatorname{tg} \alpha}{2a} e^{-ax}.$$

Если при задании  $\operatorname{tg} \alpha$  допущена погрешность  $\varepsilon$ , то значение решения при  $x = 1$  получит приращение

$$\Delta y(1) = \frac{\varepsilon}{2a} e^a - \frac{\varepsilon}{2a} e^{-a}. \quad (4)$$

При больших  $a$  вычитаемое в равенстве (4) пренебрежимо мало, но коэффициент при  $\varepsilon$  в первом слагаемом  $e^a/(2a)$  стано-



вится большим. Поэтому метод стрельбы при решении задачи (1'), будучи формально приемлемой процедурой, при больших  $a$  становится практически непригодным. Это перекликается с соображениями п. 2 § 5, где был приведен пример вычислительно неустойчивого алгоритма для решения разностной краевой задачи.

**2. Метод прогонки.** Для решения краевой задачи

$$\left. \begin{aligned} y'' - p(x)y &= f(x), & 0 \leq x \leq 1, \\ y(0) &= Y_0, & y(1) = Y_1 \end{aligned} \right\}$$

при  $p(x) \gg 1$  можно воспользоваться разностной схемой

$$\left. \begin{aligned} \frac{y_{m+1} - 2y_m + y_{m-1}}{h^2} - p(x_m)y_m &= f(x_m), \\ 0 < m < M, & \quad Mh = 1, \\ y_0 = Y_0, & \quad y_M = Y_1 \end{aligned} \right\}$$

и решать разностную задачу прогонкой. Условия применимости прогонки при  $p(x) > 0$ , как легко проверит читатель, выполнены.

**3. Метод Ньютона.** Метод стрельбы при решении хорошо поставленной краевой задачи может оказаться, как мы видели, неприменимым из-за вычислительной неустойчивости. Но метод прогонки, даже формально, можно применять только для решения линейных задач.

Метод Ньютона сводит решение нелинейной задачи к серии линейных задач и состоит в следующем. Пусть известна некоторая функция  $y_0(x)$ , удовлетворяющая граничным условиям (1) и грубо приближенно равная искомому решению  $y(x)$ . Положим

$$y(x) = y_0(x) + v(x), \quad (5)$$

где  $v(x)$  — поправка к нулевому приближению  $y_0(x)$ . Подставим (5) в уравнение (1) и линеаризуем задачу, используя равенства

$$y''(x) = y_0''(x) + v''(x),$$

$$f(x, y_0 + v, y_0' + v') =$$

$$= f(x, y_0, y_0') + \frac{\partial f(x, y_0, y_0')}{\partial y} v + \frac{\partial f(x, y_0, y_0')}{\partial y'} v' + O(v^2 + |v'|^2).$$

Отбрасывая остаточный член  $O(v^2 + |v'|^2)$ , получим линейную задачу для поправки  $\bar{v}(x)$ :

$$\left. \begin{aligned} \bar{v}'' &= p(x)\bar{v}' + q(x)\bar{v} + \varphi(x), \\ \bar{v}(0) &= \bar{v}(1) = 0, \end{aligned} \right\} \quad (6)$$

где

$$p(x) = \frac{\partial f(x, y_0, y'_0)}{\partial y'}, \quad q(x) = \frac{\partial f(x, y_0, y'_0)}{\partial y},$$
$$\varphi(x) = f(x, y_0, y'_0) - y''_0.$$

Решая линейную задачу (6) аналитически или каким-либо численным методом, найдем приближенно поправку  $\bar{v}$  и примем

$$y_1 \equiv y_0(x) + \bar{v}$$

за следующее приближение.

Описанная процедура может применяться к нелинейной разностной краевой задаче, возникшей при аппроксимации задачи (1).

---

**РАЗНОСТНЫЕ СХЕМЫ ДЛЯ УРАВНЕНИЙ  
С ЧАСТНЫМИ ПРОИЗВОДНЫМИ.  
ОСНОВНЫЕ ПОНЯТИЯ**

Выше, в связи с разностными схемами для обыкновенных дифференциальных уравнений, мы определили понятия сходимости, аппроксимации и устойчивости. Мы доказали теорему о том, что если разностная краевая задача аппроксимирует дифференциальную задачу и устойчива, то при измельчении сетки решение разностной задачи сходится к решению дифференциальной. В этой теореме содержится указание на способы построения сходящихся разностных схем для численного решения дифференциальных краевых задач: надо строить аппроксимирующие разностные схемы и выбирать среди них устойчивые.

Определения сходимости, аппроксимации и устойчивости и теорема о связи между этими понятиями носят общий характер. Они одинаково имеют смысл для любых функциональных уравнений. Мы иллюстрировали их примерами разностных схем для обыкновенных дифференциальных уравнений и для интегрального уравнения. Здесь мы проиллюстрируем некоторые основные способы построения разностных схем и проверки их устойчивости примерами разностных схем для уравнений с частными производными. При этом обнаружится много важных и существенно новых по сравнению со случаем обыкновенных дифференциальных уравнений обстоятельств. Главные из них: разнообразие сеток и способов аппроксимации, неустойчивость большинства взятых наудачу аппроксимирующих схем, сложность исследования устойчивости и трудности вычисления решений разностных краевых задач, требующие специальных усилий для их преодоления.

ГЛАВА 7

**ПРОСТЕЙШИЕ ПРИЕМЫ ПОСТРОЕНИЯ  
И ИССЛЕДОВАНИЯ РАЗНОСТНЫХ СХЕМ**

**§ 21. Напоминание и иллюстрация основных определений**

**1. Определение сходимости.** Пусть требуется приближенно вычислить решение  $u$  дифференциальной краевой задачи

$$Lu = f, \tag{1}$$

поставленной в некоторой области  $D$  с границей  $\Gamma$ . Для этого следует выбрать дискретное множество точек  $D_h$  — сетку, — принадлежащее  $D + \Gamma$ , ввести линейное нормированное пространство  $U_h$  функций, определенных на сетке  $D_h$ , установить соответствие между решением  $u$  и функцией  $[u]_h \in U_h$ , которую будем считать искомой таблицей решения  $u$ . Для приближенного отыскания таблицы  $[u]_h$ , которую мы условились считать точным решением задачи (1), надо на основе задачи (1) составить такую систему уравнений

$$L_h u^{(h)} = f^{(h)} \quad (2)$$

относительно функции  $u^{(h)}$  из  $U_h$ , чтобы имела место сходимость

$$\| [u]_h - u^{(h)} \|_{U_h} \rightarrow 0 \quad \text{при} \quad h \rightarrow 0. \quad (3)$$

Если для решения разностной краевой задачи (2) выполнено неравенство

$$\| [u]_h - u^{(h)} \|_{U_h} \leq Ch^k,$$

то говорят, что сходимость имеет порядок  $k$  относительно  $h$ .

Задача построения сходящейся разностной схемы (2) разбивают на две — на построение разностной схемы (2), аппроксимирующей задачу (1) на решении  $u$  последней, и на проверку устойчивости схемы (2).

**2. Определение аппроксимации.** Напомним определение аппроксимации. Чтобы это понятие имело смысл, надо ввести норму в пространстве  $F_h$ , которому принадлежит правая часть  $f^{(h)}$  уравнения (2). По определению, разностная задача (2) аппроксимирует задачу (1) на решении  $u$ , если в равенстве

$$L_h [u]_h = f^{(h)} + \delta f^{(h)}$$

невязка  $\delta f^{(h)}$ , возникающая при подстановке  $[u]_h$  в разностную краевую задачу (2), стремится к нулю при  $h \rightarrow 0$ :

$$\| \delta f^{(h)} \|_{F_h} = \| L_h [u]_h - f^{(h)} \|_{F_h} \rightarrow 0.$$

Если

$$\| \delta f^{(h)} \|_{F_h} \leq Ch^k,$$

где  $C$  не зависит от  $h$ , то аппроксимация имеет порядок  $k$  относительно  $h$ .

Построим, например, для задачи Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= \varphi(x, t), & -\infty < x < \infty, & \quad 0 \leq t \leq T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty, & \end{aligned} \right\} \quad (4)$$

одну из аппроксимирующих ее разностных схем. Задача (4) записывается в форме (1), если положить

$$Lu \equiv \begin{cases} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x}, & -\infty < x < \infty, & 0 \leq t \leq T, \\ u(x, 0), & -\infty < x < \infty, & \end{cases}$$

$$f = \begin{cases} \varphi(x, t), & -\infty < x < \infty, & 0 \leq t \leq T. \\ \psi(x), & -\infty < x < \infty. & \end{cases}$$

В качестве сетки  $D_h$  (рис. 8) используем совокупность точек пересечения прямых

$$x = mh, \quad t = n\tau, \quad m = 0, \pm 1, \dots; \quad n = 0, 1, \dots, [T/\tau],$$

где  $h > 0$ ,  $\tau > 0$  — некоторые числа, а  $[T/\tau]$  — целая часть дроби  $T/\tau$ . Будем считать, что шаг  $\tau$  связан с шагом  $h$  зависимостью

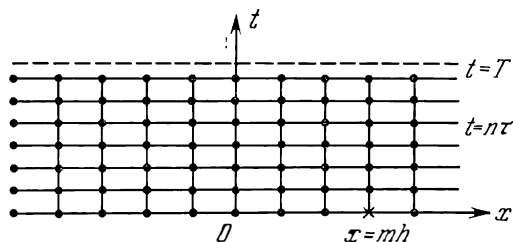


Рис. 8.

$\tau = rh$ , где  $r = \text{const}$ , так что сетка  $D_h$  зависит только от одного параметра  $h$ . Искомой сеточной функцией является таблица  $[u]_h = \{u(mh, n\tau)\}$  значений решения  $u(x, t)$  задачи (4) в точках сетки  $D_h$ .

Перейдем к построению аппроксимирующей задачу (4) разностной схемы (2). Значения сеточной функции  $u^h$  в точке  $(x_m, t_n) = (mh, n\tau)$  сетки  $D_h$  будем обозначать  $u_m^n$ . Схему (2) получим, приблизив производные  $\partial u/\partial t$  и  $\partial u/\partial x$  разностными отношениями

$$\left. \begin{aligned} \frac{\partial u}{\partial t} \Big|_{x, t} &\approx \frac{u(x, t + \tau) - u(x, t)}{\tau}, \\ \frac{\partial u}{\partial x} \Big|_{x, t} &\approx \frac{u(x + h, t) - u(x, t)}{h} \end{aligned} \right\} \quad (4')$$

Эта схема имеет вид

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_m^n}{h} &= \varphi(mh, n\tau), \\ m = 0, \pm 1, \dots; \quad n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi(mh), \quad m = 0, \pm 1, \dots \end{aligned} \right\} \quad (5)$$

Оператор  $L_h$  и правая часть  $f^{(h)}$  для схемы (5) задаются соответственно равенствами

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_m^n}{h}, & m = 0, \pm 1, \dots; n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0, & m = 0, \pm 1, \dots, \end{cases}$$

$$f^{(h)} = \begin{cases} \varphi(mh, n\tau), & m = 0, \pm 1, \dots; n = 0, 1, \dots, [T/\tau] - 1, \\ \psi(mh), & m = 0, \pm 1, \dots \end{cases}$$

Таким образом,  $f^{(h)}$  — это пара сеточных функций  $\varphi(mh, n\tau)$  и  $\psi(mh)$ , одна из которых задана на двумерной сетке

$$(x_m, t_n) = (mh, n\tau), \quad m = 0, \pm 1, \dots; n = 0, 1, \dots, [T/\tau] - 1$$

(см. рис. 8), а другая — на одномерной

$$(x_m, 0) = (mh, 0), \quad m = 0, 1, \dots$$

Разностное уравнение (4) можно разрешить относительно  $u_m^{n+1}$ , получив

$$u_m^{n+1} = (1 - r)u_m^n + ru_{m+1}^n + \tau\varphi(mh, n\tau). \quad (6)$$

Итак, зная значения  $u_m^n$ ,  $m = 0, \pm 1, \dots$ , решения  $u^{(h)}$  в точках сетки при  $t = n\tau$ , можно вычислить значения  $u_m^{n+1}$  в точках сетки при  $t = (n+1)\tau$ . Поскольку значения  $u_m^0$  при  $t = 0$  заданы равенствами  $u_m^0 = \psi(mh)$ , мы можем шаг за шагом вычислить значения решения  $u^{(h)}$  в точках сетки на прямых  $t = \tau$ ,  $t = 2\tau$  и т. д., т. е. всюду на  $D_h$ .

Перейдем к выяснению порядка аппроксимации, которым обладает схема (5). За  $F_h$  можно принять линейное пространство всех пар ограниченных функций  $g^{(h)} = \left( \begin{matrix} \varphi_m^n \\ \psi_m \end{matrix} \right)$ , положив

$$\|g^{(h)}\|_{F_h} = \max_{m, n} |\varphi_m^n| + \max_m |\psi_m|.*$$

Как уже отмечалось в § 13, норма, в которой рассматривается аппроксимация, может быть выбрана многими способами и выбор этот безразличен. Пока нам будет достаточно в каче-

\*) Если  $\max |\varphi_m^n|$  или  $\max |\psi_m|$  не достигается, то имеется в виду точная верхняя грань  $\sup |\varphi_m^n|$  или  $\sup |\psi_m|$ .

стве нормы брать верхнюю грань модулей всех компонент, образующих элемент  $g^{(h)}$  пространства  $F_h$ . Будем иметь в виду всюду в этом параграфе именно такую норму.

Предположим, что решение  $u(x, t)$  задачи (4) имеет ограниченные вторые производные. Тогда по формуле Тейлора

$$\left. \begin{aligned} \frac{u(x_m + h, t_n) - u(x_m, t_n)}{h} &= \frac{\partial u(x_m, t_n)}{\partial x} + \frac{h}{2} \frac{\partial^2 u(x_m + \xi, t_n)}{\partial x^2}, \\ \frac{u(x_m, t_n + \tau) - u(x_m, t_n)}{\tau} &= \frac{\partial u(x_m, t_n)}{\partial t} + \frac{\tau}{2} \frac{\partial^2 u(x_m, t_n + \eta)}{\partial t^2}, \end{aligned} \right\} \quad (7)$$

где  $\xi$  и  $\eta$  — некоторые числа, зависящие от  $m$ ,  $n$  и  $h$  и удовлетворяющие неравенствам  $0 < \xi < h$ ,  $0 < \eta < \tau$ .

С помощью формул (7) выражение

$$L_h[u]_h \equiv \left\{ \begin{aligned} &\frac{u(x_m, t_n + \tau) - u(x_m, t_n)}{\tau} - \frac{u(x_m + h, t_n) - u(x_m, t_n)}{h}, \\ &u(x_m, 0) \end{aligned} \right.$$

можно переписать в виде

$$L_h[u]_h = \left\{ \begin{aligned} &\left( \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} \right)_{x_m, t_n} + \frac{\tau}{2} \frac{\partial^2 u(x_m, t_n + \eta)}{\partial t^2} - \frac{h}{2} \frac{\partial^2 u(x_m + \xi, t_n)}{\partial x^2}, \\ &u(x_m, 0) + 0 \end{aligned} \right.$$

или

$$L_h[u]_h = f^{(h)} + \delta f^{(h)},$$

где

$$\delta f^{(h)} = \left\{ \begin{aligned} &\frac{\tau}{2} \frac{\partial^2 u(x_m, t_n + \eta)}{\partial t^2} - \frac{h}{2} \frac{\partial^2 u(x_m + \xi, t_n)}{\partial x^2}, \\ &0. \end{aligned} \right.$$

Следовательно,

$$\|\delta f^{(h)}\|_{F_h} \leq \left( \sup \left| \frac{\partial^2 u}{\partial t^2} \right| \cdot \frac{\tau}{2} + \sup \left| \frac{\partial^2 u}{\partial x^2} \right| \cdot \frac{1}{2} \right) h.$$

Таким образом, рассматриваемая разностная схема (5) имеет первый порядок аппроксимации относительно  $h$  на решении  $u(x, t)$ , обладающем ограниченными вторыми производными.

**3. Определение устойчивости.** Напомним и проиллюстрируем теперь определение устойчивости. Разностная краевая задача (2), по определению, устойчива, если существуют числа  $\delta > 0$  и  $h_0 > 0$  такие, что при любом  $h < h_0$  и любом  $\delta f^{(h)}$  из  $F_h$ , удовлетворяющем неравенству  $\|\delta f^{(h)}\|_{F_h} < \delta$ , разностная краевая задача

$$L_h z^{(h)} = f^{(h)} + \delta f^{(h)}$$

имеет одно и только одно решение, причем выполняется условие

$$\|z^{(h)} - u^{(h)}\|_{U_h} \leq C \|\delta f^{(h)}\|_{F_h},$$

где  $C$  — некоторая постоянная, не зависящая от  $h$ .

В § 12, где введено понятие устойчивости, показано, что в случае линейного оператора  $L_h$  сформулированное определение равносильно следующему.

**Определение.** *Разностная краевая задача (2) устойчива, если существует  $h_0 > 0$  такое, что при  $h < h_0$  и любом  $f^{(h)} \in F_h$  она однозначно разрешима, причем*

$$\|u^{(h)}\|_{U_h} \leq C \|f^{(h)}\|_{F_h}, \quad (8)$$

где  $C$  — некоторая постоянная, не зависящая от  $h$  и от  $f^{(h)}$ .

Свойство устойчивости можно трактовать как равномерную относительно  $h$  чувствительность решения разностной краевой задачи (2) к возмущениям  $\delta f^{(h)}$  правой части.

Подчеркнем, что в силу приведенного определения устойчивость есть некоторое *внутреннее* свойство разностной краевой задачи. Оно формулируется независимо от какой-либо связи с дифференциальной краевой задачей, в частности независимо от аппроксимации или сходимости.

Однако *если разностная краевая задача аппроксимирует на решении и дифференциальную и устойчива, то имеет место сходимость (3). При этом порядок относительно  $h$  скорости сходимости совпадает с порядком аппроксимации.*

Доказательство этой важной теоремы проведено в § 12.

Покажем, что разностная схема (5) при  $r \leq 1$  устойчива. При этом норму  $\|\cdot\|_{U_h}$  определим равенством

$$\|u^{(h)}\|_{U_h} = \sup_{m, n} \|u_m^n\| = \max_n \sup_m |u_m^n|.$$

Норму  $\|\cdot\|_{F_h}$  будем понимать, как выше: если  $g^{(h)} \in F_h$ ,

$$g^{(h)} = \begin{cases} \varphi_m^n, & m = 0, \pm 1, \dots; \quad n = 0, 1, \dots, [T/\tau], \\ \psi_m, & m = 0, \pm 1, \dots, \end{cases}$$

то

$$\|g^{(h)}\|_{F_h} = \max_{m, n} |\varphi_m^n| + \max_m |\psi_m| = \max_n [\max_m |\varphi_m^n| + \max_m |\psi_m|].$$

Разностную задачу

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_m^n}{h} &= \varphi_m^n, & m = 0, \pm 1, \dots; \\ & & n = 0, 1, \dots, [T/\tau], \\ u_m^0 &= \psi_m, & m = 0, \pm 1, \dots, \end{aligned} \right\} \quad (5')$$





Отсюда непосредственно следует

$$\max_m |u_m^{n+1}| \leq \max_m |\psi_m| + T \max_{m, n} |\varphi_m^n| \leq \\ \leq \|f^{(h)}\|_{F_h} + T \|f^{(h)}\|_{F_h} = (1 + T) \|f^{(h)}\|_{F_h}.$$

Доказанное неравенство

$$\max_m |u_m^{n+1}| \leq (1 + T) \|f^{(h)}\|_{F_h}$$

имеет место для всех  $n$ , так что оно останется справедливым, если вместо  $\max_m |u_m^{n+1}|$  написать  $\max_n \max_m |u_m^n| = \|u^{(h)}\|_{U_h}$ :

$$\|u^{(h)}\|_{U_h} \leq (1 + T) \|f^{(h)}\|_{F_h}. \quad (9)$$

Неравенство (9) означает устойчивость линейной задачи (5), поскольку существование и единственность решения задачи (6') при произвольных ограниченных  $\varphi_m^n$  и  $\psi_m$ , очевидно, имеют место. Роль постоянной  $C$  в неравенстве (8) играет здесь число  $1 + T$ .

Не следует думать, что одна только аппроксимация дифференциальной краевой задачи (1) разностной краевой задачей (2) обеспечивает устойчивость и, следовательно, сходимость (3). Мы убедились в этом в § 9 с помощью специально сконструированного примера аппроксимирующей, но расходящейся разностной схемы.

В случае уравнений с частными производными непригодность наудачу взятой аппроксимирующей разностной схемы является правилом, а выбор устойчивой (и, следовательно, сходящейся) разностной схемы — постоянной заботой вычислителя.

Напомним, например, что доказательство устойчивости разностной схемы (5) мы провели в предположении, что  $\tau/h = r \leq 1$ . В случае  $r > 1$  разностная задача (5) по-прежнему аппроксимирует задачу (4), но наше доказательство устойчивости не проходит. Покажем, что в этом случае нет сходимости решения  $u^{(h)}$  разностной задачи (5) к решению  $u(x, t)$  дифференциальной задачи (4), а значит, не может быть и устойчивости, так как устойчивость влечет за собою сходимость.

Пусть, для определенности,  $\varphi(x, t) \equiv 0$ , так что также  $\varphi(mh, n\tau) \equiv 0$ ; пусть, далее,  $T = 1$ . Шаг  $h$  будем выбирать так, чтобы точка  $(0, 1)$  на плоскости  $Oxt$  принадлежала сетке, т. е. чтобы число

$$N = \frac{1}{\tau} = \frac{1}{rh}$$

было целым (рис. 9). В силу разностного уравнения имеем

$$u_m^{n+1} = (1 - r) u_m^n + r u_{m+1}^n.$$

Значение  $u_0^{n+1} = u_0^N$  решения  $u^{(h)}$  в точке  $(0, 1)$  сетки выражается через значения  $u_0^n$  и  $u_1^n$  решения в точках  $(0, 1-\tau)$  и  $(h, 1-\tau)$  сетки. Два значения  $u_0^n$  и  $u_1^n$  выражаются через значения  $u_0^{n-1}$ ,  $u_1^{n-1}$  и  $u_2^{n-1}$  решения в трех точках сетки  $(0, 1-2\tau)$ ,  $(h, 1-2\tau)$  и  $(2h, 1-2\tau)$ . Значения решения  $u_0^{n-1}$ ,  $u_1^{n-1}$ ,  $u_2^{n-1}$  в свою очередь выражаются через значения решения в четырех точках  $(0, 1-3\tau)$ ,  $(h, 1-3\tau)$ ,  $(2h, 1-3\tau)$ ,  $(3h, 1-3\tau)$  и т. д.

В конечном счете значение  $u_0^{n+1}$  выражается через значения  $u_m^0$  решения в точках сетки  $(0, 0)$ ,  $(h, 0)$ ,  $(2h, 0)$ , ...,  $(h/\tau, 0) = (Nh, 0)$ . Все эти точки лежат на отрезке

$$0 \leq x \leq \frac{h}{\tau} = \frac{1}{r}$$

прямой  $t = 0$  (см. рис. 9), где задано начальное условие

$$u(x, 0) = \psi(x)$$

для дифференциального уравнения. Таким образом, решение разностного уравнения в точке  $(0, 1)$  сетки не зависит от значений функции  $\psi(x)$  в точках  $x$ , лежащих вне отрезка

$$0 \leq x \leq \frac{1}{r}.$$

Далее, решением задачи

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= 0, & -\infty < x < \infty, & t > 0, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty, & \end{aligned} \right\}$$

как легко проверить, является функция

$$u(x, t) \equiv \psi(x + t).$$

Она постоянна на каждой характеристике  $x + t = \text{const}$  и, в частности, на прямой  $x + t = 1$ , которая проходит через точки  $(0, 1)$  и  $(1, 0)$  (см. рис. 9), и в точке  $(0, 1)$  принимает значение  $\psi(1)$ . Отсюда видно, что в случае  $r > 1$  сходимости, вообще говоря, быть не может. Действительно, в этом случае отрезок оси абсцисс

$$0 \leq x \leq \frac{1}{r} < 1$$

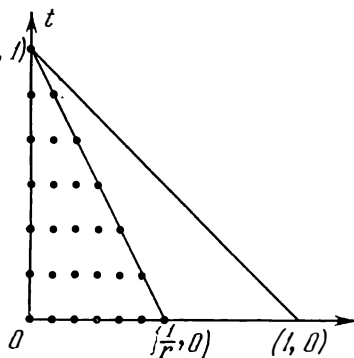


Рис. 9.

не содержит точку  $(1, 0)$ . Если бы при какой-нибудь функции  $\psi(x)$  сходимость, случайно, имела место, то, не меняя значения  $\psi(x)$  на отрезке

$$0 \leq x \leq \frac{1}{r}$$

и не меняя, таким образом, значения решения разностного уравнения в точке  $(0, 1)$ , мы могли бы нарушить сходимость, изменив  $\psi(x)$  в точке  $x = 1$  и ее окрестности, что отразилось бы на значении  $u(0, 1) = \psi(1)$  решения дифференциального уравнения. Изменение  $\psi(x)$  в точке  $x = 1$  и ее окрестности можно внести так, чтобы не нарушить существования вторых производных функции  $\psi(x)$  и решения  $u(x, t) = \psi(x + t)$ , так что аппроксимация на решении  $u(x, t)$  будет иметь место. В этих условиях из устойчивости схемы (5) вытекала бы сходимость. Но поскольку при  $r > 1$  нет сходимости, то нет и устойчивости.

Проведенное доказательство неустойчивости разностной схемы (5) носит косвенный характер. Интересно проследить непосредственно, как сказывается неустойчивость при  $r > 1$  разностной схемы (5) на чувствительности решения  $u^{(h)}$  к ошибкам при задании  $f^{(h)}$ . Ведь именно равномерная относительно  $h$  чувствительность решения к ошибкам при задании  $f^{(h)}$  и определена выше как устойчивость.

Допустим, что при всех  $h$  выполняются тождества  $\varphi(mh, n\tau) \equiv 0$  и  $\psi(mh) \equiv 0$ , так что

$$f^{(h)} = \begin{Bmatrix} \varphi_m^n \\ \psi_m \end{Bmatrix} = 0$$

и решение  $u^{(h)} = \{u_m^n\}$  задачи (5) есть тождественный нуль,  $u_m^n \equiv 0$ . Допустим, далее, что при задании начальных данных допущена ошибка и вместо  $\psi_m = 0$  задано  $\tilde{\psi}_m = (-1)^m \varepsilon$ ,  $\varepsilon = \text{const}$ , так что вместо

$$f^{(h)} = \begin{Bmatrix} \varphi_m^n \\ \psi_m \end{Bmatrix} = 0$$

задано

$$\tilde{f}^{(h)} = \begin{Bmatrix} 0 \\ \tilde{\psi}_m \end{Bmatrix}, \quad \| \tilde{f}^{(h)} \|_{F_h} = \varepsilon.$$

Будем обозначать получающееся при этом решение через  $\tilde{u}^{(h)}$ . В силу уравнений

$$\begin{aligned} \tilde{u}_m^{n+1} &= (1-r)\tilde{u}_m^n + r\tilde{u}_{m+1}^n, \\ \tilde{u}_m^0 &= (-1)^m \varepsilon \end{aligned}$$

для  $\tilde{u}_m^1$  получим

$$\begin{aligned}\tilde{u}_m^1 &= (1-r)\tilde{u}_m^0 + r\tilde{u}_{m+1}^0 = \\ &= (1-r)(-1)^m \varepsilon + r(-1)^{m+1} \varepsilon = (1-2r)(-1)^m \varepsilon = (1-2r)\tilde{u}_m^0.\end{aligned}$$

Мы видим, что допущенная при  $n=0$  ошибка умножилась на число  $(1-2r)$ . При переходе к  $\tilde{u}_m^2$  получим

$$\tilde{u}_m^2 = (1-r)\tilde{u}_m^1 + r\tilde{u}_{m+1}^1 = (1-2r)\tilde{u}_m^1 = (1-2r)^2\tilde{u}_m^0.$$

Вообще

$$\tilde{u}_m^n = (1-2r)^n \tilde{u}_m^0 = (1-2r)^n (-1)^m \varepsilon.$$

При  $r > 1$  будет  $1-2r < -1$ , так что ошибка

$$\tilde{u}_m^0 = (-1)^m \varepsilon$$

при переходе от одного слоя  $t = n\tau$  сетки к следующему умножается на отрицательное число, превосходящее единицу по модулю. При  $n = [T/\tau]$  будет

$$|\tilde{u}_m^n| = |1-2r|^{[T/\tau]} |\tilde{u}_m^0|.$$

Отсюда

$$\begin{aligned}\|\tilde{u}^{(h)}\|_{U_h} &= |1-2r|^{[T/(r h)]} |\tilde{u}_m^0| = |1-2r|^{[T/(r h)]} \max |\tilde{\psi}_m| = \\ &= |1-2r|^{[T/(r h)]} \|\tilde{f}^{(h)}\|_{F_h}.\end{aligned}$$

При фиксированном  $T$  первоначально допущенная в начальных данных ошибка  $(-1)^m \varepsilon$  увеличивается в очень быстро возрастающее при  $h \rightarrow 0$  число раз, равное  $|1-2r|^{[T/(r h)]}$ .

Остановимся теперь кратко на критике принятого нами способа оценки качества аппроксимации сравнением величины нормы невязки  $\|\delta f^{(h)}\|_{F_h}$  с той или иной степенью  $h$ . Как мы знаем, для устойчивых схем порядок аппроксимации совпадает с порядком погрешности  $[u]_h - u^{(h)}$  в решении. Качество схемы естественно оценивать по количеству вычислительной работы, необходимой для получения заданной точности. Количество же работы, вообще говоря, пропорционально числу точек  $N$  использованной разностной сетки. Для обыкновенных дифференциальных уравнений  $N$  пропорционально шагу  $h$ . Поэтому, когда мы говорим, что погрешность  $\varepsilon \approx h^p$ , мы тем самым утверждаем, что  $\varepsilon \approx 1/N^p$ , т. е. что уменьшение погрешности вдвое требует увеличения работы в  $\sqrt[p]{2}$  раз. Таким образом, в случае обыкновенных разностных уравнений порядок аппроксимации относительно  $h$  характеризует объем работы.

Для уравнений с частными производными дело обстоит уже не так. В рассмотренном нами примере задачи с двумя

переменными  $x$  и  $t$  сетка задается двумя шагами  $\tau$  и  $h$ . Число  $N$  точек сетки, помещающихся в ограниченной области на плоскости  $Oxt$ , имеет порядок  $1/(\tau h)$ . Это число также может применяться для оценки количества работы, затрачиваемой при решении разностных уравнений. Пусть  $\tau = rh$ . В этом случае  $N \approx 1/h^2$  и утверждение, что  $\varepsilon \approx h^p$ , эквивалентно утверждению  $\varepsilon \approx 1/N^{p/2}$ . Если  $\tau = rh^2$ , то  $N \approx 1/h^3$  и утверждение  $\varepsilon \approx h^p$  эквивалентно тому, что  $\varepsilon \approx 1/N^{p/3}$ .

Мы видим, что в случае уравнений с частными производными порядок погрешности естественнее было бы измерять не в степенях  $h$ , а в степенях  $1/N$ . Мы все же остановимся на описанном выше способе оценки аппроксимации степенями  $h$ , так как это удобнее при проведении выкладок. Читатель, однако, должен при оценке качества разностных схем иметь в виду отмеченное обстоятельство.

Надо еще заметить, что утверждение о пропорциональности вычислительной работы числу  $N$  точек сетки тоже не всегда является верным. Можно привести примеры разностных схем, для вычисления решения по которым требуется произвести  $\approx N^{1+q}$  арифметических операций, где  $q = 1/2, 1$  или даже  $2$ . С этим приходится встречаться при решении разностных краевых задач, аппроксимирующих эллиптические уравнения, или при решении задач в случае трех и более независимых переменных (например,  $u = u(t, x, y)$ ). В многомерном случае построение разностных схем, для вычисления решения по которым требуется  $\approx N$  арифметических операций, является непростой задачей, о которой будет идти речь в §§ 31, 32.

При реальных расчетах на вычислительной машине для сравнительной оценки используемых алгоритмов за меру качества схемы обычно естественно принять машинное время. Машинное время не обязательно пропорционально числу арифметических действий.

Играют роль, иногда превалирующую, затраты времени на пересылку информации из одного блока машинной памяти в другой. Может играть роль время, расходуемое на логические операции.

### ЗАДАЧИ

1. Для задачи Коши (4) исследовать следующую разностную схему:

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_m^n - u_{m-1}^n}{h} &= \varphi(mh, n\tau), \\ m = 0, \pm 1, \dots; \quad n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi(mh), \quad m = 0, \pm 1, \dots, \end{aligned} \right\}$$

где  $\tau = rh$ ,  $r = \text{const}$ . Именно:

а) Выписать оператор  $L_h$  и правую часть  $f^{(h)}$ , возникающие при записи этой схемы в виде  $L_h u^{(h)} = f^{(h)}$ .

б) Изобразить взаимное расположение трех точек сетки, значения  $u^{(h)}$  в которых связывает разностное уравнение при фиксированных  $m$  и  $n$ .

в) Показать, что разностная схема аппроксимирует дифференциальную задачу с первым относительно  $h$  порядком на решении  $u(x, t)$ , имеющем ограниченные вторые производные.

г) Выяснить, устойчива ли исследуемая разностная схема при каком-либо выборе,  $r, \tau = r h$ .

2. Для задачи Коши  $u_t + u_x = \varphi(x, t)$ ,  $u(x, 0) = \psi(x)$ ,  $-\infty < x < \infty$ ,  $0 \leq t \leq T$  исследовать по предложенному в задаче 1 плану каждую из следующих разностных схем:

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} + \frac{u_m^n - u_{m-1}^n}{h} &= \varphi(mh, n\tau), \\ m = 0, \pm 1, \dots; \quad n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi(mh), \quad m = 0, \pm 1, \dots \end{aligned} \right\}$$

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} + \frac{u_{m+1}^n - u_m^n}{h} &= \varphi(mh, n\tau), \\ m = 0, \pm 1, \dots; \quad n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi(mh), \quad m = 0, \pm 1, \dots \end{aligned} \right\}$$

## § 22. Простейшие приемы построения аппроксимирующих разностных схем

**1. Замена производных разностными отношениями.** Простейший прием построения разностных краевых задач, аппроксимирующих дифференциальные, состоит в замене производных соответствующими разностными отношениями. Приведем несколько примеров разностных схем, полученных таким способом. В этих примерах будут использованы приближенные формулы

$$\left. \begin{aligned} \frac{df(z)}{dz} &\approx \frac{f(z + \Delta z) - f(z)}{\Delta z}, \\ \frac{df(z)}{dz} &\approx \frac{f(z) - f(z - \Delta z)}{\Delta z}, \\ \frac{df(z)}{dz} &\approx \frac{f(z + \Delta z) - f(z - \Delta z)}{2\Delta z}, \\ \frac{d^2f(z)}{dz^2} &\approx \frac{f(z + \Delta z) - 2f(z) + f(z - \Delta z)}{\Delta z^2}. \end{aligned} \right\} \quad (1)$$

Предполагая функцию  $f(z)$  имеющей достаточное число ограниченных производных, можно выписать выражения для

остаточных членов этих формул. По формуле Тейлора

$$\left. \begin{aligned} f(z + \Delta z) &= f(z) + \Delta z f'(z) + \frac{(\Delta z)^2}{2!} f''(z) + \\ &\quad + \frac{(\Delta z)^3}{3!} f'''(z) + \frac{(\Delta z)^4}{4!} f^{(4)}(z) + o[(\Delta z)^4], \\ f(z - \Delta z) &= f(z) - \Delta z f'(z) + \frac{(\Delta z)^2}{2!} f''(z) - \\ &\quad - \frac{(\Delta z)^3}{3!} f'''(z) + \frac{(\Delta z)^4}{4!} f^{(4)}(z) + o[(\Delta z)^4]. \end{aligned} \right\} \quad (2)$$

Используя разложения (2), можно получить выражения для остаточных членов приближенных формул (1). Именно, справедливы равенства

$$\left. \begin{aligned} \frac{f(z + \Delta z) - f(z)}{\Delta z} &= f'(z) + \left[ \frac{\Delta z}{2} f''(z) + o(\Delta z) \right], \\ \frac{f(z) - f(z - \Delta z)}{\Delta z} &= f'(z) + \left[ -\frac{\Delta z}{2} f''(z) + o(\Delta z) \right], \\ \frac{f(z + \Delta z) - f(z - \Delta z)}{2\Delta z} &= f'(z) + \left[ \frac{(\Delta z)^2}{3} f'''(z) + o((\Delta z)^2) \right], \\ \frac{f(z + \Delta z) - 2f(z) + f(z - \Delta z)}{\Delta z^2} &= f''(z) + \left[ \frac{(\Delta z)^2}{12} f^{(4)}(z) + o((\Delta z)^2) \right]. \end{aligned} \right\} \quad (3)$$

Остаточные члены приближенных формул (1) входят в соответствующие равенства (3) в виде выражений в квадратных скобках.

Очевидно, что формулы (1) и выражения остаточных членов, выписанные в формулах (3), можно использовать и при замене частных производных разностными отношениями. Например,

$$\frac{\partial u(x, t)}{\partial t} \approx \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t},$$

причем

$$\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = \frac{\partial u(x, t)}{\partial t} + \left[ \frac{\Delta t}{2} \frac{\partial^2 u(x, t)}{\partial t^2} + o(\Delta t) \right].$$

Точно так же справедливы формулы

$$\frac{\partial u(x, t)}{\partial x} \approx \frac{u(x + \Delta x, t) - u(x, t)}{\Delta x}$$

и при этом

$$\frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} = \frac{\partial u(x, t)}{\partial x} + \left[ \frac{\Delta x}{2} \frac{\partial^2 u(x, t)}{\partial x^2} + o(\Delta x) \right]$$

и т. д.



Пример 1. Вернемся к задаче Коши (4) из § 21:

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= \varphi(x, t), & -\infty < x < \infty, & 0 \leq t \leq T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty. \end{aligned} \right\} \quad (4)$$

Для аппроксимации этой задачи Коши построим три схемы. Во всех этих схемах используем сетку  $D_h$ , образованную точками пересечения прямых  $x = mh$ ,  $t = n\tau$ , попавшими в полосу  $0 \leq t \leq T$ . Значения  $\tau$  и  $h$  будем считать связанными соотношением  $\tau = rh$ , где  $r$  — некоторая положительная постоянная. Простейшая из этих схем имеет вид (5) из § 21:

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_m^n}{h} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh), \end{cases} \quad (5)$$

и получается при замене производных  $u_t = du/dt$  и  $u_x = du/dx$  по приближенным формулам

$$\begin{aligned} u_t(x, t) &\approx \frac{u(x, t + \tau) - u(x, t)}{\tau}, \\ u_x(x, t) &\approx \frac{u(x + h, t) - u(x, t)}{h}. \end{aligned}$$

Мы подробно исследовали эту схему в § 21. Для нее невязка  $\delta f^{(h)}$ , возникающая при подстановке решения  $[u]_h$  дифференциальной задачи в левую часть разностной задачи

$$L_h [u]_h = f^{(h)} + \delta f^{(h)},$$

выражается формулой

$$\delta f^{(h)} = \begin{cases} \left( \frac{\tau}{2} u_{tt} - \frac{h}{2} u_{xx} \right)_m^n + o(\tau + h), \\ 0. \end{cases}$$

За норму элемента  $f^{(h)}$  пространства  $F_h$  примем в этом параграфе максимум всех компонент элемента  $f^{(h)} \in F_h$ . Тогда, очевидно,

$$\|\delta f^{(h)}\|_{F_h} = O(\tau + h) = O(rh + h) = O(h),$$

и порядок аппроксимации получается первый.

Вторая схема получается при использовании другой формулы для замены  $du/dx$ :

$$\frac{\partial u(x, t)}{\partial x} \approx \frac{u(x, t) - u(x - h, t)}{h},$$

Она имеет вид

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_m^n - u_{m-1}^n}{h} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh). \end{cases}$$

Здесь

$$\delta f^{(h)} \equiv \begin{cases} \left( \frac{\tau}{2} u_{tt} + \frac{h}{2} u_{xx} \right)_m + o(\tau + h), \\ 0, \end{cases} \\ \|\delta f^{(h)}\|_{F_h} = O(h)$$

и порядок аппроксимации снова получается первый.

Вторая схема, казалось бы, совсем несущественно отличается от первой. В дальнейшем мы увидим, однако, что вторая схема непригодна для счета: она неустойчива при любом  $\tau/h = r = \text{const}$ .

Третья схема,

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{n+1} - \frac{u_{m+1}^n + u_{m-1}^n}{2}}{\tau} - \frac{u_{m+1}^n - u_m^n}{h} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh), \end{cases}$$

получается при замене производных разностными отношениями по приближенным формулам

$$\frac{\partial u(x, t)}{\partial t} \approx \frac{u(x, t + \tau) - \frac{u(x + h, t) + u(x - h, t)}{2}}{\tau}, \\ \frac{\partial u(x, t)}{\partial x} \approx \frac{u(x + h, t) - u(x, t)}{h}.$$

С помощью тейлоровских разложений (2) для достаточно гладкого решения  $u(x, t)$  задачи (1) получаем

$$\frac{u(x, t + \tau) - \frac{u(x + h, t) + u(x - h, t)}{2}}{\tau} - \frac{u(x + h, t) - u(x, t)}{h} = \\ = \left[ \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} - \frac{h^2}{2\tau} \frac{\partial^2 u}{\partial x^2} + \frac{\tau}{2} \frac{\partial^2 u}{\partial t^2} \right]_{x, t} + O\left(\tau^2 + h^2 + \frac{h^4}{\tau}\right) = \\ = \varphi(x, t) + \left[ -\frac{h}{2r} u_{xx} + \frac{\tau}{2} u_{tt} + O(h^2) \right]_{x, t}.$$

Поэтому

$$L_h [u]_h = \begin{cases} \varphi(mh, nh) + \left[ -\frac{h}{2r} u_{xx} + \frac{\tau}{2} u_{tt} + O(h^2) \right], \\ \psi(mh) + 0, \end{cases}$$

так что  $\delta f^{(h)}$  из формулы

$$L_h[u]_h = f^{(h)} + \delta f^{(h)}$$

имеет вид

$$\delta f^{(h)} \equiv \begin{cases} -\frac{h}{2r} u_{xx} + \frac{\tau}{2} u_{tt} + O(h^2), \\ 0. \end{cases}$$

Следовательно,  $\|\delta f^{(h)}\|_{F_h} = O(h)$  и имеет место аппроксимация первого порядка, как и в двух первых примерах.

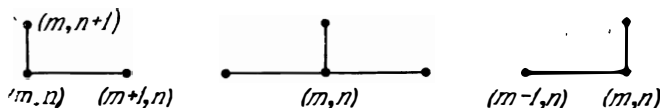


Рис. 10.

Рассмотрим теперь случай, когда связь между шагами сетки задается не формулой  $\tau = rh$ , как выше, а формулой

$$\tau = rh^2, \quad r = \text{const},$$

предписывающей ускоренное измельчение шага  $\tau$  по сравнению с шагом  $h$ . В этом случае

$$L_h[u]_h = \begin{cases} \left[ \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} - \frac{1}{2r} \frac{\partial^2 u}{\partial x^2} \right]_{x_m, t_n} + O(h^2), \\ u(mh, n\tau). \end{cases}$$

Отсюда видно, что рассматриваемая разностная схема аппроксимирует задачу

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} - \frac{1}{2r} \frac{\partial^2 u}{\partial x^2} &= \varphi(x, t), \\ u(x, 0) &= \psi(x), \end{aligned}$$

а вовсе не задачу Коши (4), которую мы хотели аппроксимировать.

Мы столкнулись с тем фактом, что одна и та же разностная схема может в случае различной связи  $\tau = \tau(h)$  аппроксимировать при  $h \rightarrow 0$  различные дифференциальные задачи. Такого рода разностные схемы называют *негибкими*.

Для облегчения запоминания разностной схемы ее обычно принято сопоставлять с картинкой, на которой изображено взаимное расположение точек сетки («шаблон»), значения решения в которых связывает разностное уравнение при некоторых фиксированных значениях  $m$  и  $n$ .

Для трех рассмотренных схем эти картинки изображены на рис. 10.

Пример 2. Приведем две разностные схемы, аппроксимирующие задачу Коши для уравнения теплопроводности

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = \varphi(x, t), \quad -\infty < x < \infty, \quad 0 < t < T,$$

$$u(x, 0) = \psi(x), \quad -\infty < x < \infty.$$

Простейшая из них,

$$L_h^{(1)} u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh), \end{cases}$$

$$f^{(h)} \equiv \begin{cases} \varphi(mh, n\tau), \\ \psi(mh), \end{cases}$$

получается при замене производных  $u_t$  и  $u_{xx}$  разностными отношениями по формулам

$$u_t(x, t) \approx \frac{u(x, t + \tau) - u(x, t)}{\tau},$$

$$u_{xx}(x, t) \approx \frac{u(x + h, t) - 2u(x, t) + u(x - h, t)}{h^2}.$$

Если для замены  $u_{xx}(x, t)$  использовать другую формулу:

$$u_{xx}(x, t) \approx \frac{u(x + h, t + \tau) - 2u(x, t + \tau) + u(x - h, t + \tau)}{h^2},$$

мы приходим к другой схеме для того же уравнения:

$$L_h^{(2)} u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh). \end{cases}$$

Чтобы различать операторы  $L_{h_i}$  этих двух схем, мы снабдили их номерами и написали  $L_h^{(1)} u^{(h)} = f^{(h)}$  и  $L_h^{(2)} u^{(h)} = \tilde{f}^{(h)}$ . Шаблоны, соответствующие этим разностным схемам, изображены на рис. 11.

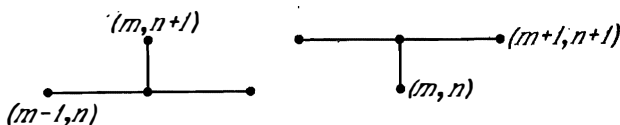


Рис. 11.

Эти схемы существенно отличаются. Вычисление решения  $u^{(h)}$  по первой из них не представляет труда и проводится по явной формуле

$$u_m^{n+1} = (1 - 2r)u_m^n + r(u_m^n + u_{m+1}^n) + \tau\varphi(mh, n\tau),$$

где  $r = \tau/h^2$ . Эта формула получена из разностного уравнения в результате решения его относительно  $u_m^{n+1}$ . Зная значения решения  $u_m^n$ ,  $m = 0, \pm 1, \dots$  на слое  $t = t_n (= n\tau)$  сетки, мы можем вычислить его значения  $u_m^{n+1}$  на следующем слое  $t = t_{n+1}$ .

Вторая схема  $L_h^{(2)} u^{(h)} = f^{(h)}$  лишена этого удобного свойства. Поэтому ее называют *невязной*. В этом случае разностное уравнение, выписанное при фиксированных  $m$  и  $n$ , нельзя разрешить относительно  $u_m^{n+1}$ , выразив это значение через известные значения  $u_{m+1}^n$ ,  $u_m^n$ ,  $u_{m-1}^n$  с предыдущего слоя. Дело в том, что в это уравнение входит не только неизвестное значение  $u_m^{n+1}$ , но также и неизвестные  $u_{m-1}^{n+1}$  и  $u_{m+1}^{n+1}$ . Поэтому для определения  $u_m^{n+1}$ ,  $m = 0, \pm 1, \dots$ , придется решать разностное уравнение относительно сеточной функции  $u_m^{n+1}$  аргумента  $m$ . Тем не менее, в дальнейшем будет показано, что схема  $L_h^{(2)} u^{(h)} = f^{(h)}$ , как правило, удобнее схемы  $L_h^{(1)} u^{(h)} = f^{(h)}$ .

При  $\tau = rh^2$ ,  $r = \text{const}$ , обе схемы имеют второй порядок аппроксимации относительно  $h$ . Вычислим невязку  $\delta f^{(h)}$  и оценим порядок аппроксимации для второй из этих схем. Пользуясь формулами (3), можно написать

$$L_h^{(2)} [u]_h \equiv \begin{cases} (u_t - u_{xx})_{x=mh, t=(n+1)\tau} - \frac{\tau}{2} u_{tt}(x, t_{n+1}) - \frac{h^2}{12} u_{xxxx}(x, t_{n+1}) + o(\tau + h^2), \\ u(mh, 0). \end{cases}$$

Отсюда, с учетом  $\tau = rh^2$ , можно написать

$$L_h^{(2)} [u]_h \equiv \begin{cases} \varphi(x_m, t_{n+1}) + O(h^2), \\ \psi(mh) + 0, \end{cases}$$

$$\delta f^{(h)} = \begin{cases} \varphi(x_m, t_n) - \varphi(x_m, t_{n+1}) + O(h^2), \\ 0. \end{cases}$$

Но

$$\begin{aligned} \varphi(x_m, t_{n+1}) &= \varphi(x_m, t_n) + [\varphi(x_m, t_{n+1}) - \varphi(x_m, t_n)] = \\ &= \varphi(x_m, t_n) + O(\tau) = \varphi(x_m, t_n) + O(h^2). \end{aligned}$$

Поэтому

$$\|\delta f^{(h)}\|_{F_h} = O(h^2).$$

Пример 3. Рассмотрим простейшую разностную схему, аппроксимирующую задачу Дирихле для уравнения Пуассона в квадрате  $D$  ( $0 < x < 1$ ,  $0 < y < 1$ ) с границей  $\Gamma$  (рис. 12, а):

$$\begin{aligned} u_{xx} + u_{yy} &= \varphi(x, y), & (x, y) \in D, \\ u|_{\Gamma} &= \psi(x, y), & (x, y) \in \Gamma. \end{aligned}$$

Построим сетку  $D_h$ , отнеся к ней те точки  $(x_m, t_n) = (mh, nh)$ , которые попали внутрь квадрата или на его границу. Шаг  $h$

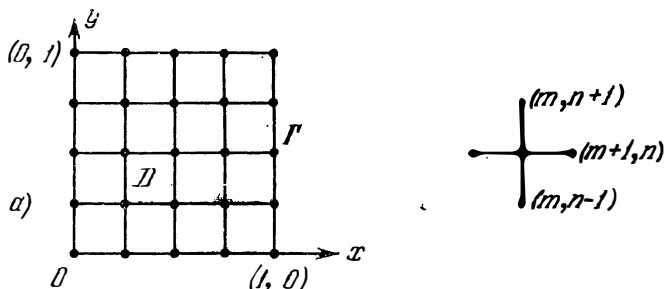


Рис. 12.

будем считать выбранным так, чтобы число  $1/h$  было целым. Разностную схему  $L_h u^{(h)} = f^{(h)}$  зададим равенствами

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{m+1, n} - 2u_{mn} + u_{m-1, n}}{h^2} + \frac{u_{m, n+1} - 2u_{mn} + u_{m, n-1}}{h^2} = \\ = \varphi(mh, nh), & (mh, nh) \in D, \\ u_{mn} = \psi(mh, nh), & (mh, nh) \in \Gamma \end{cases}$$

$$f^{(h)} = \begin{cases} \varphi(mh, nh), & \text{если } (mh, nh) \in D, \\ \psi(mh, nh), & \text{если } (mh, nh) \in \Gamma. \end{cases}$$

Невязка  $\delta f^{(h)}$ ,  $L_h[u]_h = f^{(h)} + \delta f^{(h)}$ , имеет в силу формул (3) вид

$$\delta f^{(h)} = \begin{cases} \frac{h^2}{12} (u_{xxxx} + u_{yyyy})|_{x_m, y_n} + o(h^2), \\ 0, \end{cases}$$

так что аппроксимация имеет второй порядок. Пятиточечный шаблон, отвечающий использованному разностному уравнению, изображен на рис. 12, б.

Разностные схемы, построенные выше, получались путем замены каждой производной в дифференциальном уравнении тем или иным разностным отношением.

**2. Метод неопределенных коэффициентов.** Более общий способ построения разностных схем состоит в том, что приближается не каждая производная в отдельности, а сразу весь дифференциальный оператор. Разъясним этот способ на примерах разностных схем для задачи Коши (4). Сначала рассмотрим схему первого порядка аппроксимации (5). Она связывает значения искомой функции в трех точках, изображенных на рис. 10 слева. Разностное уравнение

$$\Lambda_h u^{(h)} \equiv \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_m^n}{h} = \varphi(mh, n\tau),$$

используемое в этой схеме, имеет вид

$$\Lambda_h u^{(h)} \equiv a^0 u_m^{n+1} + a_0 u_m^n + a_1 u_{m+1}^n = \varphi(mh, n\tau).$$

Забудем на время, что нам уже известна разностная схема (5), для которой

$$a^0 = \frac{1}{\tau}, \quad a_0 = \frac{1}{h} - \frac{1}{\tau}, \quad a_1 = -\frac{1}{h},$$

и, считая эти коэффициенты неопределенными, постараемся подобрать их так, чтобы имело место равенство

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = \left( \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} \right) \Big|_{\substack{x=mh, \\ t=n\tau}} + O(h)$$

или

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = \Lambda u \Big|_{\substack{x=mh, \\ t=n\tau}} + O(h), \quad (6)$$

где

$$\Lambda u \equiv \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x}. \quad (7)$$

Воспользуемся формулой Тейлора:

$$u[mh, (n+1)\tau] = u(mh, n\tau) + \tau u'_t(mh, n\tau) + O(\tau^2),$$

$$u[(m+1)h, n\tau] = u(mh, n\tau) + h u'_x(mh, n\tau) + O(h^2).$$

Подставив эти выражения в правую часть равенства

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} \equiv a^0 u[mh, (n+1)\tau] + a_0 u(mh, n\tau) + a_1 u[(m+1)h, n\tau]$$

получим

$$\begin{aligned} \Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} &= (a^0 + a_0 + a_1) u(mh, n\tau) + \\ &+ a^0 \tau \frac{\partial u(mh, n\tau)}{\partial t} + a_1 h \frac{\partial u(mh, n\tau)}{\partial x} + O(a^0 \tau^2, a_1 h^2). \end{aligned} \quad (8)$$

Поскольку нашей целью является такой подбор коэффициентов  $a^0$ ,  $a_0$ ,  $a_1$ , чтобы выполнялось условие аппроксимации (6), то естественно предварительно так сгруппировать слагаемые в правой части равенства (8), чтобы выделился член (7). Тогда остальные слагаемые образуют остаточный член аппроксимации, который должен быть мал. Чтобы выделить член  $\Lambda u$ , можно заменить в правой части равенства (8) производные  $\partial u/\partial t$  или  $\partial u/\partial x$  соответственно по одной из формул:

$$\frac{\partial u}{\partial t} \equiv \Lambda u + \frac{\partial u}{\partial x} \quad \text{или} \quad \frac{\partial u}{\partial x} = \frac{\partial u}{\partial t} - \Lambda u.$$

Для определенности воспользуемся первой из них.

Кроме того, подчиним шаги  $\tau$  и  $h$  связи  $\tau = rh$ , где  $r$  — какая-нибудь постоянная. После этого равенство (8) примет следующий вид:

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = a^0 r h \Lambda u \Big|_{\substack{x=mh, \\ t=n\tau}} + (a^0 + a_0 + a_1) u(mh, n\tau) + \\ + (a^0 r + a_1) h u_x(mh, n\tau) + O(a^0 r^2 h^2, a_1 h^2). \quad (9)$$

Среди всех гладких функций  $u(x, t)$  можно указать такие, для которых  $u$ ,  $\partial u/\partial x$  и  $\partial u/\partial t$  в любой заранее заданной фиксированной точке принимают любые независимые друг от друга значения. Следовательно, и значения

$$u, \quad \frac{\partial u}{\partial x} \quad \text{и} \quad \Lambda u = \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = \varphi(x, t)$$

также можно считать независимыми друг от друга. Ввиду этого из равенства (9) следует, что для выполнения при любой правой части  $\varphi(x, t)$  задачи (4) условия аппроксимации

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = (\Lambda u) \Big|_{\substack{x=mh, \\ t=n\tau}} + O(h)$$

необходимо, чтобы выполнялись равенства

$$\begin{aligned} a^0 r h &= 1 + O_1(h), \\ a^0 + a_0 + a_1 &= 0 + O_2(h), \\ (a^0 r + a_1) h &= 0 + O_3(h), \end{aligned}$$

где  $O_1(h)$ ,  $O_2(h)$ ,  $O_3(h)$  — какие-нибудь произвольные величины порядка  $h$ . Положим  $O_1(h) = O_2(h) = O_3(h) = 0$ . Получающаяся при этом система

$$\begin{aligned} a^0 r h &= 1, \\ a^0 + a_0 + a_1 &= 0, \\ a^0 r + a_1 &= 0 \end{aligned}$$



имеет единственное решение

$$\begin{aligned} a^0 &= \frac{1}{rh} = \frac{1}{\tau}, \\ a_0 &= \frac{r-1}{rh} = \frac{1}{h} - \frac{1}{\tau}, \\ a_1 &= -\frac{1}{h}, \end{aligned}$$

которое приводит к уже известной схеме (5).

Теперь мы, однако, дополнительно узнали, что среди разностных схем вида

$$L_h u^{(h)} \equiv \begin{cases} a^0 u_m^{n+1} + a_0 u_m^n + a_1 u_{m+1}^n = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh) \end{cases}$$

она является единственной, аппроксимирующей рассматриваемую задачу Коши. Говоря о единственности, мы пренебрегаем тем произволом, который вносит свобода выбора функций  $O_1(h)$ ,  $O_2(h)$ ,  $O_3(h)$ . Всюду в дальнейших примерах мы также будем пренебрегать подобного рода очевидным произволом и даже не всегда будем вводить произвольные величины, аналогичные величинам  $O_1(h)$ ,  $O_2(h)$ ,  $O_3(h)$ , с самого начала полагая их равными нулю.

Читатель без труда убедится, что в рассмотренном сейчас примере учет этих величин привел бы к следующему несущественному изменению результата:

$$\begin{aligned} a^0 &= \frac{1}{h} \left[ \frac{1}{r} + O(h) \right], \\ a_0 &= \frac{1}{h} \left[ \frac{r-1}{r} + O(h) \right], \\ a_1 &= \frac{1}{h} [-1 + O(h)]. \end{aligned}$$

Аналогично будет обстоять дело и в других примерах, которые нам встретятся.

Посмотрим теперь, как можно строить для задачи (4) разностные схемы

$$L_h u^{(h)} \equiv \begin{cases} a^0 u_m^{n+1} + a_0 u_m^n + a_{-1} u_{m-1}^n + a_1 u_{m+1}^n = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh) \end{cases} \quad (10)$$

более общего вида, связывающие значения искомой функции в четырех точках, изображенных на рис. 13.

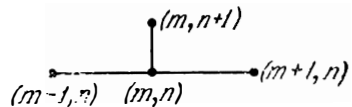


Рис. 13.

Шаги сетки снова свяжем равенством  $\tau = rh$ ,  $r = \text{const}$ , и введем обозначение  $\Lambda_h$ , положив

$$\Lambda_h u^{(h)} \equiv a^0 u_m^{n+1} + a_0 u_m^n + a_{-1} u_{m-1}^n + a_1 u_{m+1}^n. \quad (11)$$

Для всякой достаточно гладкой функции  $u(x, t)$  с помощью формулы Тейлора можно написать

$$\begin{aligned} \Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} &= (a^0 + a_0 + a_1 + a_{-1}) u(mh, n\tau) + \\ &+ a^0 r h u_t(mh, n\tau) + (a_1 - a_{-1}) h u_x(mh, n\tau) + \frac{1}{2} a^0 r^2 h^2 u_{tt}(mh, n\tau) + \\ &+ \frac{1}{2} (a_1 + a_{-1}) h^2 u_{xx}(mh, n\tau) + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3). \end{aligned} \quad (12)$$

Выделим в правой части этого равенства член  $\Lambda u \equiv \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x}$ , воспользовавшись для этого тождеством  $u_t = u_x + \Lambda u$ . Имеем

$$\begin{aligned} \Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} &= a^0 r h \Lambda u \Big|_{\substack{x=mh, \\ t=n\tau}} + (a^0 + a_0 + a_1 + a_{-1}) u(mh, n\tau) + \\ &+ (a^0 r + a_1 - a_{-1}) h u_x(mh, n\tau) + \frac{1}{2} a^0 r^2 h^2 u_{tt}(mh, n\tau) + \\ &+ \frac{1}{2} (a_1 + a_{-1}) h^2 u_{xx}(mh, n\tau) + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3). \end{aligned}$$

Если предполагать, что величина  $O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3)$  достаточно мала, — это предположение подтвердится в дальнейшем, — то для выполнения условия аппроксимации

$$(\Lambda_h [u]_h)_{\substack{x=mh, \\ t=n\tau}} = (\Lambda u)_{\substack{x=mh, \\ t=n\tau}} + O(h)$$

необходимо, чтобы четыре числа  $a^0$ ,  $a_0$ ,  $a_1$ ,  $a_{-1}$  удовлетворяли следующим трем равенствам:

$$\begin{aligned} a^0 r h &= 1 + O_1(h), \\ a^0 + a_0 + a_1 + a_{-1} &= 0 + O_2(h), \\ (a^0 r + a_1 - a_{-1}) h &= 0 + O_3(h). \end{aligned}$$

Положим, как условились, произвольные величины  $O_1(h)$ ,  $O_2(h)$ ,  $O_3(h)$  порядка  $h$  равными нулю. Получим систему уравнений

$$\left. \begin{aligned} a^0 r h &= 1, \\ a^0 + a_0 + a_1 + a_{-1} &= 0, \\ a^0 r + a_1 - a_{-1} &= 0. \end{aligned} \right\} \quad (13)$$

Если условия 13 выполнены, то

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = \Lambda u \Big|_{\substack{x=mh, \\ t=n\tau}} + \frac{1}{2} a^0 r^2 h^2 u_{tt}(mh, n\tau) + \\ + \frac{1}{2} (a_1 + a_{-1}) h^2 u_{xx}(mh, n\tau) + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3).$$

Система (13) имеет много решений — семейство решений, зависящее от одного параметра. Одно из этих решений:

$$a^0 = \frac{1}{rh}, \quad a_{-1} = 0, \quad a_0 = \frac{r-1}{h}, \quad a_1 = -\frac{1}{h},$$

дает уже рассмотренную схему (5). Решению

$$a^0 = \frac{1}{rh}, \quad a_0 = -\frac{1}{rh}, \quad a_{-1} = \frac{1}{2h}, \quad a_1 = -\frac{1}{2h}$$

соответствует схема

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh). \end{cases}$$

Выбрав какое-либо решение системы (13), надо его подставить в остаточный член и убедиться, что он мал. Для двух сейчас приведенных решений подстановка чисел  $a^0$ ,  $a_0$ ,  $a_1$ ,  $a_{-1}$  дает остаточные члены

$$\frac{a^0 r^2 h^2}{2} \frac{\partial^2 u}{\partial t^2} + \frac{a_1 + a_{-1}}{2} h^2 \frac{\partial^2 u}{\partial x^2} + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3)$$

порядка  $O(h)$ .

Среди гладких функций  $u(x, t)$  есть многочлены второй степени, для которых  $\partial^2 u / \partial t^2$  и  $\partial^2 u / \partial x^2$  принимают в любой фиксированной точке любые независимые наперед заданные значения. При этом член  $O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3)$ , в который входят третьи производные многочлена  $u(x, t)$ , обращается в нуль. Поэтому для того, чтобы остаточный член был порядка  $h$ , необходимо, чтобы коэффициенты при  $\partial^2 u / \partial t^2$  и  $\partial^2 u / \partial x^2$  каждый в отдельности были порядка  $h$ . Поскольку из первого уравнения (13) имеем  $a^0 = 1/(rh)$ , то коэффициент при  $\partial^2 u / \partial t^2$  есть  $rh/2$  и порядок остаточного члена всегда не выше первого.

Мы установили, что нельзя построить разностную схему вида (10), которая аппроксимирует задачу

$$Lu = \begin{cases} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = \varphi(x, t), \\ u(x, 0) = \psi(x) \end{cases}$$

с порядком  $h^2$ . Для увеличения порядка аппроксимации пришлось бы увеличить число точек разностной сетки, используемых в конструируемой схеме.

Укажем некоторый способ, позволяющий все же построить разностную схему с аппроксимацией порядка  $h^2$ , использующую только четыре указанные точки разностной сетки. Способ повышения порядка аппроксимации, который мы сейчас изложим с помощью примера, носит общий характер. Оказывается, что можно подобрать коэффициенты  $a^0, a_{-1}, a_0, a_1$  так, чтобы выполнялось равенство

$$\begin{aligned} \Lambda_h [u]_h &\equiv a^0 u(mh, (n+1)\tau) + a_{-1} u((m-1)h, n\tau) + \\ &\quad + a_0 u(mh, n\tau) + a_1 u((m+1)h, n\tau) = \\ &= \Lambda u + \frac{rh}{2} [(\Lambda u)_t + (\Lambda u)_x] \Big|_{x=mh, t=n\tau} + O(h^2) = P_h \Lambda u \Big|_{(x_m, t_n)} + O(h^2), \end{aligned}$$

где

$$P_h = E + \frac{rh}{2} \left( \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right),$$

$E$  — оператор умножения на единицу. Тогда ввиду  $\Lambda u = u_t - u_x = \varphi(x, t)$  разностная схема

$$L_h u^{(h)} = \begin{cases} a^0 u_m^{n+1} + a_{-1} u_{m-1}^n + a_0 u_m^n + a_1 u_{m+1}^n = \varphi_m^n, \\ u_m^0 = \psi(mh), \end{cases}$$

где

$$\varphi_m^n = (P_h \varphi) \Big|_{x=mh, t=n\tau} = \varphi + \frac{rh}{2} (\varphi_t + \varphi_x) \Big|_{x=mh, t=n\tau},$$

будет аппроксимировать рассматриваемую дифференциальную задачу на решении  $u(x, t)$  со вторым порядком относительно  $h$ .

Коэффициенты  $a^0, a_{-1}, a_0, a_1$  снова могут быть подобраны методом неопределенных коэффициентов. Они оказываются следующими:

$$a^0 = \frac{1}{rh}, \quad a_0 = -\frac{1}{rh} + \frac{r}{h}, \quad a_{-1} = \frac{1-r}{2h}, \quad a_1 = -\frac{1+r}{2h}.$$

Оператор  $\Lambda_h$  при этом получается таким:

$$\Lambda_h u^{(h)} = \frac{1}{\tau} (u_m^{n+1} - u_m^n) - \frac{1}{2h} (u_{m+1}^n - u_{m-1}^n) - \frac{r}{2h} (u_{m+1}^n - 2u_m^n + u_{m-1}^n).$$

Методом неопределенных коэффициентов можно не только подобрать коэффициенты  $a^0, a_{-1}, a_0, a_1$ , при которых

$$\Lambda_h [u]_h = a^0 u(x, t + \tau) + a_{-1} u(x - h, t) + \\ + a_0 u(x, t) + a_1 u(x + h, t) = P_h \Lambda u + O(h^2)$$

при выписанном выше операторе  $P_h$ , но и построить все такие операторы.

Покажем, как это делается.  
Считая, что

$$\Lambda_h u^{(h)} \equiv a^0 u_m^{n+1} + a_{-1} u_{m-1}^n + a_0 u_m^n + a_1 u_{m+1}^n,$$

и пользуясь формулой Тейлора, получим

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = (a^0 + a_0 + a_1 + a_{-1}) u(mh, n\tau) + \\ + a^0 r h u_t(mh, n\tau) + (a_1 + a_{-1}) h u_x(mh, n\tau) + \frac{1}{2} a^0 r^2 h^2 u_{tt}(mh, n\tau) + \\ + \frac{1}{2} (a_1 + a_{-1}) h^2 u_{xx}(mh, n\tau) + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3). \quad (14)$$

Это выражение мы сейчас преобразуем. Начнем с вывода тождества

$$\frac{\partial^2 u}{\partial t^2} \equiv \frac{\partial^2 u}{\partial x^2} + (\Delta u)_t + (\Delta u)_x$$

которое вытекает из определения  $\Delta u$ :

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} + \Delta u.$$

Доказательство содержится в цепочке очевидных тождеств:

$$\frac{\partial^2 u}{\partial t^2} = \left( \frac{\partial u}{\partial x} + \Delta u \right)_t = \frac{\partial^2 u}{\partial x \partial t} + (\Delta u)_t \equiv \frac{\partial}{\partial x} u_t + (\Delta u)_t \equiv \\ \equiv \frac{\partial}{\partial x} (u_x + \Delta u) + (\Delta u)_t \equiv u_{xx} + (\Delta u)_t + (\Delta u)_x.$$

Используя эти тождества, можно выражение (14) переписать в следующем эквивалентном виде:

$$\Lambda_h [u]_h \Big|_{\substack{x=mh, \\ t=n\tau}} = a^0 r h (\Delta u)_m^n + \frac{1}{2} a^0 r^2 h^2 [(\Delta u)_t + (\Delta u)_x]_m^n + \\ + (a_0 + a^0 + a_1 + a_{-1}) u(mh, n\tau) + (a^0 r + a_1 - a_{-1}) h u_x(mh, n\tau) + \\ + \left[ \frac{1}{2} a^0 r^2 + \frac{1}{2} (a_1 + a_{-1}) \right] h^2 u_{xx}(mh, n\tau) + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3). \quad (15)$$

Построим оператор  $\Lambda_h$ , удовлетворяющий условию  $\Lambda_h u = P_h \Lambda u + O(h^2)$ . Члены, содержащие  $\Delta u, (\Delta u)_x, (\Delta u)_t$ , включим в выражение  $P_h \Lambda u$ , поскольку

выбор оператора  $P_h \Lambda u$  в наших руках. Все остальные члены

$$\begin{aligned} & (a_0 + a^0 + a_{-1} + a_1) u(mh, n\tau), \\ & (a^0 r + a_1 - a_{-1}) h u_x(mh, n\tau), \\ & \frac{a^0 r^2 + a_1 + a_{-1}}{2} h^2 u_{xx}(mh, n\tau). \\ & O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3), \end{aligned}$$

обязательно войдут слагаемыми в остаточный член равенства

$$\Lambda_h [u]_h = P_h \Lambda u + \text{ост. член},$$

как бы мы ни выбрали оператор  $P_h$ . Справедливость последнего утверждения доказывается тем, что среди функций  $u(x, t)$  существуют такие, для которых  $u, u_x, u_{xx}, \Lambda u, (\Lambda u)_x, (\Lambda u)_t$  принимают в любой фиксированной точке  $(x_0, t_0)$  любые, независимые друг от друга, наперед заданные значения  $u^0, u_x^0, u_{xx}^0, (\Lambda u)^0, (\Lambda u)_x^0, (\Lambda u)_t^0$ . Такой функцией является, например, многочлен

$$\begin{aligned} P(x, t) = & u^0 + u_x^0(x - x_0) + [(\Lambda u)^0 + u_x^0](t - t_0) + \frac{1}{2} u_{xx}^0(x - x_0)^2 + \\ & + \frac{1}{2} [u_{xx}^0 + (\Lambda u)_x^0 + (\Lambda u)_t^0](t - t_0)^2 + [u_{xx}^0 + (\Lambda u)_x^0](x - x_0)(t - t_0). \end{aligned}$$

Ввиду независимости значений  $u, u_x, u_{xx}, \Lambda u, (\Lambda u)_t, (\Lambda u)_x$  при любом выборе оператора  $P_h$  для аппроксимации второго порядка необходимо, чтобы каждое в отдельности слагаемое, входящее в остаточный член, было порядка  $h^2$ . Это требование можно записать такими равенствами:

$$\left. \begin{aligned} a_0 + a^0 + a_1 + a_{-1} &= 0, \\ (a_0 r + a_1 + a_{-1}) h &= 0, \\ (a_0 r^2 + a_1 + a_{-1}) h^2 &= 0. \end{aligned} \right\} \quad (16)$$

Решение системы (16) определено с точностью до множителя. Дополним эту систему уравнением

$$a_0 r h = 1, \quad (17)$$

которое выражает естественное, хотя и не необходимое ограничение на выбор оператора  $P_h$ : коэффициент при  $(\Lambda u)$  в выражении  $P_h \Lambda u$  равен единице.

В правые части равенств (6), (7) можно было бы добавить произвольные слагаемые  $O_1(h^2), O_2(h^2), O_3(h^2), O_4(h^2)$ , но мы считаем их равными нулю, как условились.

Решая систему уравнений (16), (17), получаем коэффициенты  $a^0, a_{-1}, a_0, a_1$ , которые были уже нами приведены:

$$a^0 = \frac{1}{rh}, \quad a_0 = -\frac{1}{rh} + \frac{r}{h}, \quad a_{-1} = \frac{1-r}{2h}, \quad a_1 = -\frac{1+r}{2h}.$$

При этих значениях коэффициентов остаточный член равенства (15)

$$\begin{aligned} \Lambda_h [u]_h &= \Lambda u + \frac{rh}{2} (\Lambda u)_t + \frac{rh}{2} (\Lambda u)_x + O(a^0 r^3 h^3, a_1 h^3, a_{-1} h^3) \equiv \\ &\equiv P_h \Lambda u + O(n^0 r^3 h^3, a_1 h^3, a_{-1} h^3) \end{aligned}$$

удовлетворяет оценке

$$|O(a_0 r^3 h^3, a_1 h^3, a_{-1} h^3)| \leq A(r^2 h^2 + h^2).$$

где  $A$  — некоторая постоянная, зависящая только от максимума абсолютных величин производных третьего порядка функции  $u(x, t)$ . Это можно записать также в форме

$$|\Lambda_h [u]_h - P_h \varphi|_m^n \leq A(r^2 + 1) h^2.$$

Итак, мы установили, что с точностью до несущественных изменений только одна разностная схема

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} - \frac{\tau}{2} \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = \\ \quad = \left[ \varphi + \frac{r h}{2} (\varphi_t + \varphi_x) \right]_{\substack{x=mh, \\ t=n\tau}} \\ u_m^0 = \psi(mh) \end{cases} \quad (17')$$

среди всех разностных схем вида

$$L_h u^{(h)} = \begin{cases} a^0 u_m^{n+1} + a_{-1} u_{m-1}^n + a_0 u_m^n + a_1 u_{m+1}^n = P_h \varphi|_m^n, \\ u_m^0 = \psi(mh) \end{cases}$$

аппроксимирует дифференциальную краевую задачу (4) на решении  $u(x, t)$  последней со вторым порядком относительно  $h$ .

Во всех рассмотренных до сих пор в этой главе примерах разностных схем  $L_h u^{(h)} = f^{(h)}$  оператор  $L_h$ , который отображает пространство  $U_h$  в пространство  $F_h$ , задается явными формулами. Но часто оказываются полезными разностные схемы, в которых оператор  $L_h$  описывается тем или иным более сложным образом. В дальнейшем мы еще встретимся с задачами, где такие схемы возникают естественным образом.

Изложенные приемы построения разностных схем остаются применимыми и в случае задач с переменными коэффициентами, в случае нелинейных задач, в случае сеток с переменным шагом. Например, в случае неравномерной сетки, изображенной на рис. 14, для замены производных, входящих в дифференциальное уравнение  $u_{xx} + u_{yy} = \varphi(x, y)$ , разностными отношениями с целью построения разностной схемы можно воспользоваться

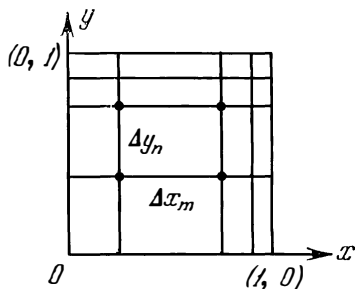


Рис. 14.

формулами

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} \Big|_{(x_n, y_n)} &= \frac{\frac{u(x_{m+1}, y_n) - u(x_m, y_n)}{\Delta x_m} - \frac{u(x_m, y_n) - u(x_{m-1}, y_n)}{\Delta x_{m-1}}}{\frac{\Delta x_m + \Delta x_{m-1}}{2}} + \\ &+ \frac{1}{3} (\Delta x_m - \Delta x_{m-1}) u_{xxx} + O[(\Delta x_m + \Delta x_{m-1})^2], \\ \frac{\partial^2 u}{\partial y^2} \Big|_{(x_n, y_n)} &= \frac{\frac{u(x_m, y_{n+1}) - u(x_m, y_n)}{\Delta y_n} - \frac{u(x_m, y_n) - u(x_m, y_{n-1})}{\Delta y_{n-1}}}{\frac{\Delta y_n + \Delta y_{n-1}}{2}} + \\ &+ \frac{1}{3} (\Delta y_n - \Delta y_{n-1}) + O((\Delta y_n + \Delta y_{n-1})^2), \end{aligned}$$

отбросив в них остаточные члены.

Указанные формулы проверяются с помощью разложений Тейлора (2). Методом неопределенных коэффициентов можно убедиться в единственности этих формул: с точностью до несущественного произвола есть только один набор коэффициентов  $a_{-1}$ ,  $a_0$ ,  $a_1$ , при котором для любой достаточно гладкой функции  $u(x, t)$  имеет место формула

$$\begin{aligned} \frac{\partial^2 u(x_m, y_n)}{\partial x^2} &= a_{-1} u(x_{m-1}, y_n) + a_0 u(x_m, y_n) + \\ &+ a_1 u(x_{m+1}, y_n) + O[\max(\Delta x_{m-1}, \Delta x_m)] \end{aligned}$$

с остаточным членом первого порядка малости относительно  $\max[\Delta x_{m-1}, \Delta x_m]$ .

Формулы вида

$$\begin{aligned} \frac{\partial^2 u(x_m, y_n)}{\partial x^2} &= a_{-1} u(x_{m-1}, y_n) + a_0 u(x_m, y_n) + \\ &+ a_1 u(x_{m+1}, y_n) + O([\max(\Delta x_{m-1}, \Delta x_m)]^2) \end{aligned}$$

с остаточным членом второго порядка малости при  $\Delta x_{m-1} \neq \Delta x_m$  не существует.

Для более точной замены производной разностным отношением здесь необходимо привлечь более трех точек сетки.

**3. Схемы с пересчетом, или схемы предиктор-корректор.** При построении разностных схем, аппроксимирующих нестационарные задачи, может быть использована та же идея, которая лежит в основе конструкции схем Рунге — Кутты для обыкновенных дифференциальных уравнений, — идея пересчета. Пересчет позволяет повысить порядок аппроксимации, получаемый по исходной схеме, не использующей пересчета. Кроме того, в случае квазилинейных дифференциальных уравнений пересчет дает дополнительную возможность получения так называемых дивергентных схем, о которых будет идти речь в § 30.



Напомним идею пересчета на примере простейшей из схем Рунге — Кутты численного решения задачи Коши для уравнения

$$\frac{dy}{dt} = f(t, y), \quad y(0) = \psi, \quad 0 < t < T. \quad (18)$$

Если значение  $y_p$  в точке  $t_p = p\tau$  уже вычислено, то для вычисления  $y_{p+1}$  находим предварительно вспомогательную величину  $\tilde{y}_{p+1/2}$ , пользуясь простейшей схемой Эйлера (схема «предиктор»)

$$\frac{\tilde{y}_{p+1/2} - y_p}{\tau/2} = f(t_p, y_p), \quad (19)$$

а затем осуществляем корректирующий пересчет по схеме

$$\frac{y_{p+1} - y_p}{\tau} = f(t_{p+1/2}, \tilde{y}_{p+1/2}). \quad (20)$$

Вспомогательная величина  $\tilde{y}_{p+1/2}$ , найденная по схеме первого порядка точности, позволяет приближенно найти угловой коэффициент интегральной кривой в середине отрезка  $[t_p, t_{p+1}]$  и получить  $y_{p+1}$  по формуле (20) с большей точностью, чем это было бы по схеме Эйлера (19).

Мы уже отмечали в п. 4 § 19, что все соображения остаются в силе, если  $y$ ,  $y_p$ ,  $\tilde{y}_{p+1/2}$  будут конечномерными векторами, а  $f$  — вектор-функцией. Но можно пойти и дальше, а именно считать  $y$ ,  $y_p$ ,  $\tilde{y}_{p+1/2}$  элементами функционального пространства, а  $f$  оператором в этом пространстве. Например, задачу Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} + A \frac{\partial u}{\partial x} &= 0, & -\infty < x < \infty, & 0 < t < T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty, \end{aligned} \right\} \quad (21)$$

$A = \text{const}$ , можно считать задачей вида (18), если положить  $y(t) = u(x, t)$ , так что при каждом  $t$  под  $y$  надо понимать функцию аргумента  $x$ , а под операцией  $f$  понимать оператор  $-A \frac{\partial}{\partial x}$ . Приведем пример разностной схемы с пересчетом для задачи (21).

Пр и м е р. Пусть сеточная функция  $u^p = \{u_m^p\}$ ,  $m = 0, \pm 1, \dots$ , при данном  $p$  уже вычислена. Найдем вспомогательную сеточную функцию  $\tilde{u}^{p+1/2} = \{\tilde{u}_{m+1/2}^{p+1/2}\}$ ,  $m = 0, \pm 1, \dots$ , отнесенную к моменту времени  $t_{p+1/2} = (p + 1/2)\tau$  и к точкам  $x_{m+1/2} = (m + 1/2)h$ , воспользовавшись следующей схемой первого порядка точности:

$$\frac{\tilde{u}_{m+1/2}^{p+1/2} - \frac{u_{m+1}^p + u_m^p}{2}}{\tau/2} + A \frac{u_{m+1}^p - u_m^p}{h} = 0, \quad m = 0, \pm 1, \dots \quad (22)$$

Затем осуществим коррекцию и найдем  $u^{p+1}$  с помощью схемы

$$\frac{u_m^{p+1} - u_m^p}{\tau} + A \frac{\bar{u}_{m+1/2}^{p+1/2} - \bar{u}_{m-1/2}^{p+1/2}}{h} = 0, \quad m = 0, \pm 1, \dots \quad (23)$$

Исключая  $\bar{u}^{p+1/2}$  из уравнений (22), (23), получим схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} + A \frac{u_{m+1}^p - u_{m-1}^p}{2h} - A^2 \frac{\tau}{2} \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0, \\ u_m^0 = \psi(x_m), \quad m = 0, \pm 1, \dots; \quad p = 0, 1, \dots, [T/\tau] - 1. \end{aligned} \right\} \quad (24)$$

Эта схема при  $A = -1$  совпадает со схемой (17'). Случай  $A \neq 1$  несущественно отличается от разобранного. Схема (24), а значит, и схема с пересчетом (22), (23) имеют второй порядок аппроксимации по  $h$ ;  $\tau = rh$ ,  $r = \text{const}$ .

**4. О других приемах.** Назовем еще два весьма важных и широко применяемых приема построения разностных схем. Первый из них основан на использовании записи дифференциального уравнения, для которого строится разностная схема, в форме «интегрального закона сохранения». Необходимость в использовании этого приема естественным образом возникает при расчете так называемых обобщенных решений, не обладающих достаточным числом производных или даже вообще разрывных. Возникающие разностные схемы носят название дивергентных или консервативных. Способ построения дивергентных схем излагается в главе 9.

Второй прием основан на использовании той или иной вариационной постановки дифференциальной краевой задачи, решение которой надо вычислить. Этот прием часто называют методом конечных элементов, а возникающие разностные схемы — вариационно-разностными или проекционно-разностными. Этот прием позволяет строить разностные схемы на нерегулярных сетках, более мелких там, где решение меняется более быстро. Вариационно-разностным схемам посвящена гл. 12.

## ЗАДАЧИ

1. Для решения задачи Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = \varphi(x, t), \quad -\infty < x < \infty, \quad 0 < t < T, \\ u(x, 0) = \psi(x), \quad -\infty < x < \infty \end{aligned} \right\}$$

воспользоваться сеткой  $x_m = mh$ ,  $t_n = n\tau$ ,  $h = \tau$  и построить разностную схему вида

$$L_h u^{(h)} = \begin{cases} a^0 u_m^{n+1} + a^1 u_{m+1}^{n+1} + a_0 u_m^n + a_1 u_{m+1}^n = \varphi_m^n \\ u_m^0 = \psi(mh). \end{cases}$$

Как надо определить  $a^0, a^1, a_0, a_1$  и  $\Phi_m$ , чтобы имела место аппроксимация порядка  $h^2$ ?

2. Для задачи Коши

$$\frac{\partial u}{\partial t} - \left( \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} \right) = \Phi(x, y, t), \quad -\infty < x, y < \infty, \quad 0 < t < T,$$

$$u(x, y, 0) = \psi(x, y), \quad -\infty < x, y < \infty$$

воспользоваться сеткой  $x_m = mh, y_n = nh, t_p = p\tau$  и построить какую-либо аппроксимирующую ее разностную схему.

3. Для задачи о теплопроводности

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2}, & -\infty < x < \infty, & \quad 0 \leq t \leq T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty \end{aligned} \right\} \quad (25)$$

рассмотреть разностную схему

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} &= \sigma \frac{u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}}{h^2} + (1 - \sigma) \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2}, \\ u_m^0 &= \psi(mh), \end{aligned} \right\}$$

где  $\sigma$  — параметр,  $u_m^n$  — значение искомой функции в точке  $(x_m = mh, t_n = n\tau)$  сетки.

а) Показать, что при любом  $\sigma$  имеет место аппроксимация на гладком решении  $u(x, t)$  с порядком  $O(\tau + h^2)$ .

б) Подобрать  $\sigma$  так, чтобы аппроксимация была  $O(\tau^2 + h^2)$ .

в) Связав шаги сетки соотношением  $\tau/h^2 \equiv r = \text{const}$ , подобрать затем  $\sigma$  так, чтобы получить аппроксимацию порядка  $h^4$ .

г) При  $\sigma = 0$  подобрать число  $r = \tau/h^2$  так, чтобы аппроксимация имела порядок  $h^4$ .

д) Можно ли за счет выбора  $\sigma$  при фиксированном  $r = \tau/h^2$  добиться того, чтобы аппроксимация на любом гладком решении была порядка выше четвертого?

4. Для задачи о теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left[ a(x, t) \frac{\partial u}{\partial x} \right], \quad -\infty < x < \infty, \quad 0 < t < T,$$

$$u(x, 0) = \psi(x), \quad -\infty < x < \infty,$$

пользуясь сеткой  $x_m = mh, t_n = n\tau$ , построить аппроксимирующую ее разностную схему.

5. Для нелинейной задачи о теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left[ a(u) \frac{\partial u}{\partial x} \right], \quad -\infty < x < \infty, \quad 0 < t < T,$$

$$u(x, 0) = \psi(x), \quad -\infty < x < \infty,$$

пользуясь сеткой  $x_m = mh, t_n = n\tau$ , построить аппроксимирующую ее явную разностную схему. Выписать формулы для вычисления  $u^{(h)}$  по этой схеме.

6. Доказать, что при ограниченной сеточной функции  $u^p = \{u_m^p\}$  существует и единственна ограниченная сеточная функция  $u^{p+1} = \{u_m^{p+1}\}$ , определяемая разностной схемой

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^{p+1} - u_{m-1}^{p+1}}{2h} = 0, \quad m = 0, \pm 1, \dots$$

7. Доказать, что схема с пересчетом для задачи (25), в которой значения решения  $\{\tilde{u}_m^{p+1/2}\}$  на промежуточном слое определяются по неявной схеме порядка аппроксимации  $O(\tau + h^2)$

$$\frac{\tilde{u}_m^{p+1/2} - u_m^p}{\tau/2} - \frac{\tilde{u}_{m+1}^{p+1/2} - 2\tilde{u}_m^{p+1/2} + \tilde{u}_{m-1}^{p+1/2}}{h^2} = 0, \quad m = 0, \pm 1, \dots,$$

а решение  $\{u_m^{p+1}\}$  определяется по схеме

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{\tilde{u}_{m+1}^{p+1/2} - 2\tilde{u}_m^{p+1/2} + u_{m-1}^{p+1/2}}{h^2} = 0, \quad u_m^0 = \psi(x_m), \quad m = 0, \pm 1, \dots$$

обладает аппроксимацией порядка  $O(\tau^2 + h^2)$  на гладком решении  $u$

### § 23. Примеры конструирования граничных условий при построении разностных схем

Рассмотренные в § 22 примеры были подобраны так, чтобы не возникало вопросов относительно построения разностных краевых условий. Их без труда удавалось получить из дифференциальных граничных условий так, чтобы разностные условия при подстановке в них  $[u]_h$  выполнялись точно. Здесь мы рассмотрим более сложные в этом смысле примеры.

Пример 1. Для задачи

$$Lu = \begin{cases} u_t - u_x = \varphi(x, t), \\ u(x, 0) = \psi(x) \end{cases} \quad (1)$$

при построении разностной схемы воспользуемся разностным уравнением

$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} = \varphi(mh, n\tau), \quad (2)$$

$$n = 1, 2, \dots; \quad m = 0, \pm 1, \dots; \quad \tau = rh.$$

Чтобы вычислить решение уравнения (2), надо задать не только  $u_m^0$ ,

$$u_m^0 = \psi(mh), \quad m = 0, \pm 1, \dots, \quad (3)$$

но также  $u_m^1$ ,  $m = 0, \pm 1, \dots$ . Тогда из разностного уравнения (2) при  $n = 1, 2, \dots$  можно последовательно вычислить  $u_m^2$ ,  $m = 0, \pm 1, \dots$ , затем  $u_m^3$ ,  $m = 0, \pm 1, \dots$ , и так далее. Значения  $u_m^1$  должны быть заданы близкими к

$$u(mh, \tau) = u(mh, 0) + \tau u_t(mh, 0) + O(\tau^2).$$

Поскольку  $u_t = u_x + \Lambda u$ ,  $\Lambda u \equiv u_t - u_x = \varphi(x, t)$ ,  $u(x, 0) = \psi(x)$ , то

$$u(mh, \tau) = u(mh, 0) + \tau [u_x + \Lambda u]_{x=mh, t=0} + O(\tau^2) = \\ = \psi(mh) + \tau [\psi'(mh) + \varphi(mh, 0)] + O(\tau^2).$$

Таким образом, отбрасывая член  $O(\tau^2)$ , можно положить

$$u_m^1 = \psi(mh) + \tau [\psi'(mh) + \varphi(mh, 0)]. \quad (4)$$

Ясно, что разностная схема

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{n+1} - u_m^{n-1}}{2\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh), \\ u_m^1 = \psi(mh) + \tau [\psi'(mh) + \varphi(mh, 0)] \end{cases} \quad (5)$$

аппроксимирует дифференциальную краевую задачу (1) с порядком  $h^2$ . Нетривиальность этой схемы состоит в том, что разностное уравнение (2) имеет второй порядок по  $t$ , в то время как дифференциальное (1) — первый. Поэтому потребовалось конструировать второе разностное краевое условие (4), не возникающее непосредственно из заданного краевого условия для дифференциальной задачи.

Приведем другой пример, в котором построение разностных граничных условий неочевидно.

**Пример 2.** Рассмотрим дифференциальную краевую задачу

$$Lu = \begin{cases} u_t - u_x = \varphi(x, t), & 0 < x < 1, \quad 0 < t < T, \\ u(0, x) = \psi_0(x), & 0 < x < 1, \\ u(t, 1) = \psi_1(t), & 0 < t < T. \end{cases} \quad (6)$$

Любое решение дифференциального уравнения задачи (6) однозначно определяется, если известно его значение в одной точке на каждой из прямых  $x + t = \text{const}$ . Действительно, вдоль такой прямой

$$\frac{du}{dt} = u_t + u_x \frac{dx}{dt} = u_t - u_x = \varphi(x, t),$$

так что  $u(x, t)$  является интегралом вдоль прямой

$$x + t = \text{const}$$

от  $\varphi(x, t)$ . Значение постоянной интегрирования определяется по величине  $u$  в заданной точке.

На рис. 15 изображен прямоугольник  $0 \leq x \leq 1$ ,  $0 \leq t \leq T$ , в котором мы собираемся искать решение, и нанесено семейство параллельных прямых  $x + t = \text{const}$ . Каждая из этих прямых пересекается в одной точке либо с отрезком  $0 \leq x \leq 1$  оси  $Ox$ , либо с отрезком  $0 \leq t \leq T$  прямой  $x = 1$ , где задано  $u(x, t)$ . Таким образом, задача (6) имеет единственное решение.

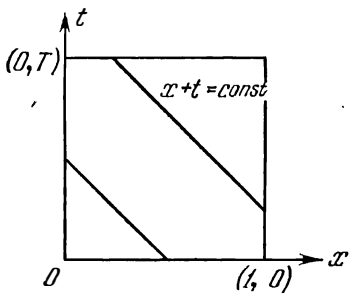


Рис. 15.

Приступим к конструированию разностной схемы для вычисления решения задачи (6). Зададим  $h$  так, чтобы  $Mh = 1$ , и положим  $\tau = rh$ , где  $M$  — целое число,  $r = \text{const}$ . В качестве сетки  $D_h$  используем точки  $(mh, n\tau)$ ,  $m = 0, 1, \dots, M$ ;  $n = 0, 1, \dots, [T/\tau]$ .

Точкам  $D_h$ , не лежащим на верхней и боковых границах прямоугольника  $D$ , поставим в соответствие уравнение

$$\frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} - \frac{\tau}{2} \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = \varphi_m^n, \quad (7)$$

где

$$\varphi_m^n = \left[ \varphi(x, t) + \frac{r\tau}{2} (\varphi_t + \varphi_x) \right]_{\substack{x=mh, \\ t=n\tau}}. \quad (8)$$

Получение этого уравнения было подробно описано в § 22.

Значения  $u_m^0$  и  $u_M^n$  зададим равенствами

$$\left. \begin{aligned} u_m^0 &= \psi_0(mh), & m &= 0, 1, \dots, M-1, \\ u_M^n &= \psi_1(n\tau), & n &= 0, 1, \dots, N, \quad N \equiv [T/\tau], \end{aligned} \right\} \quad (9)$$

которые аналогичны граничным условиям для рассматриваемой дифференциальной задачи. Но равенств (9) недостаточно, чтобы определить решение  $u_m^n$  всюду на  $D_h$ . Не удастся определить значения  $u_0^{n+1}$  на левой границе прямоугольника. Поэтому дополним разностные граничные условия следующим:

$$\frac{u_0^{n+1} - u_0^n}{\tau} - \frac{u_1^n - u_0^n}{h} = \varphi(0, n\tau), \quad n = 0, 1, \dots, N-1. \quad (10)$$

Это условие возникает при замене в равенстве

$$\frac{\partial u(0, t)}{\partial t} - \frac{\partial u(x, 0)}{\partial x} = \varphi(0, t),$$

являющемся следствием заданного дифференциального уравнения (6), производных соответствующими разностными отношениями.

Итак, мы построили разностную схему  $L_h u^{(h)} = f^{(h)}$ :

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} - \frac{\tau}{2} \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2}, \\ \quad m = 1, 2, \dots, M-1; \quad n = 0, 1, \dots, N-1, \\ u_m^0, \quad m = 0, 1, \dots, M-1, \\ u_M^n, \quad m = 0, 1, \dots, N, \\ \frac{u_0^{n+1} - u_0^n}{\tau} - \frac{u_1^n - u_0^n}{h}, \quad n = 0, 1, \dots, N-1, \end{cases}$$

$$f^{(h)} = \begin{cases} \left[ \Phi + \frac{r h}{2} (\Phi_t + \Phi_x) \right]_{\substack{x=mh, \\ t=n\tau}}, \\ \psi_0(mh), \quad m = 0, 1, \dots, M-1, \\ \psi_1(n\tau), \quad n = 0, 1, \dots, N, \\ \varphi(0, n\tau), \quad n = 0, 1, \dots, N-1. \end{cases}$$

Выясним порядок аппроксимации. С учетом рассмотрений § 22 ясно, что невязка  $\delta f^{(h)}$ , возникающая при подстановке  $[u]_h$  в разностную схему,  $L_h[u]_h = f^{(h)} + \delta f^{(h)}$ , в предположении достаточной гладкости решения  $u(x, t)$  имеет вид

$$\delta f^{(h)} \equiv \begin{cases} O_{mn}(h^2) = O(h^2), \quad m = 1, 2, \dots, M-1; \quad n = 0, 1, \dots, N-1, \\ \quad 0, \quad m = 0, 1, \dots, M-1, \\ \quad 0, \quad n = 0, 1, \dots, N, \\ \frac{\tau}{2} u_{tt}(0, n\tau + \xi_1\tau) + \frac{h}{2} u_{xx}(\xi_2 h, n\tau), \quad n = 0, 1, \dots, N-1, \\ \quad 0 < \xi_1 < 1, \quad 0 < \xi_2 < 1. \end{cases}$$

Если ввести в  $F_h$  норму, положив для произвольного элемента  $g^{(h)} \in F_h$

$$g^{(h)} = \begin{cases} A_m^n, \\ a_m, \quad \|g^{(h)}\|_{F_h} = \max_{m, n} |A_m^n| + \max_m |a_m| + \max_n |b^n| + \max_n |c^n|, \\ b^n, \\ c^n, \end{cases}$$

то  $\|\delta f^{(h)}\|_{F_h} = O(h)$  и аппроксимация окажется имеющей лишь первый порядок относительно  $h$ . Из выражения для  $\delta f^{(h)}$  видно,

что первый порядок определяется невязкой  $\frac{\tau}{2} u_{tt} + \frac{h}{2} u_{xx} = O(h)$ , возникающей при подстановке  $[u]_h$  в дополнительное искусственно сконструированное нами граничное условие на левой боковой границе.

Остаточный член в используемой сейчас норме  $\|\cdot\|_{F_h}$  оценивается только через вторые производные решения, т. е. эта норма не позволяет воспользоваться при исследовании граничных условий той гладкостью решения, которая была нужна для получения второго порядка аппроксимации во внутренних точках.

Приведем норму  $\|\cdot\|_{F_h}$ , при которой построенная выше разностная схема имеет второй порядок аппроксимации на достаточно гладком решении  $u(x, t)$ :

$$\|g^{(h)}\|_{F_h} = h \max_n |c^n| + h \max_n \left| \frac{c^{n+1} - c^n}{\tau} \right| + \\ + \left( h \sum_{m=0}^N |a_m|^2 \right)^{1/2} + \max_n |b^n| + \max_n \left| \frac{b^{n+1} - b^n}{\tau} \right| + \max_{m, n} |A_m^n|.$$

Для этой схемы, как легко видеть,

$$\|\delta f^{(h)}\|_{F_h} \leq A(rh^2 + h^2), \quad r = \tau/h.$$

При этом постоянная  $A$  оценивается через производные  $u(x, t)$  до третьего порядка включительно.

Учет гладкости в этой норме осуществляется членами

$$\left| \frac{c^{n+1} - c^n}{\tau} \right|, \quad \left| \frac{b^{n+1} - b^n}{\tau} \right|.$$

Читатель, вероятно, заметил, что часть слагаемых в формуле, задающей новую норму в  $F_h$ , отличается от соответствующих слагаемых в первой норме множителем  $h$ . Ясно, что если делать такие умножения на  $h$  и на различные степени  $h$  произвольно, то можно добиться любого порядка аппроксимации. Однако мы уже обсуждали в § 13 вопрос о выборе норм в связи с обыкновенными дифференциальными уравнениями и знаем, что разумны только такие нормы, в которых разностная схема одновременно аппроксимирует дифференциальную краевую задачу и устойчива.

Устойчивость рассматриваемой схемы с использованной нормой, в которой имеется аппроксимация второго порядка, будет доказана в § 42.

Пример 2 очень поучителен. Он показывает, что для проверки аппроксимации в разумном смысле надо правильно выбрать норму. Исследуя ту или иную схему, приходится перепро-



бовать много норм. В каждой из них надо попытаться провести исследование устойчивости, которое само по себе, по крайней мере в настоящее время, часто требует изобретательности и труда.

На практике в большинстве случаев вместо интересующей задачи все исследование проводится на упрощенной, так называемой модельной задаче, после чего проводят экспериментальный счет по разностной схеме для исходной неупрощенной задачи.

### ЗАДАЧИ

1. Для задачи Коши

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} &= \varphi(x, t), & -\infty < x < \infty & \quad 0 \leq t \leq T, \\ u(x, 0) &= \psi_1(x), & -\infty < x < \infty, \\ \frac{\partial u(x, 0)}{\partial t} &= \psi_2(x), & -\infty < x < \infty. \end{aligned}$$

исследовать аппроксимацию, которой обладает на достаточно гладком решении  $u(x, t)$  разностная схема

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{n+1} - 2u_m^n + u_m^{n-1}}{\tau^2} - \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = \varphi(mh, n\tau), \\ u_m^0 = \psi_1(mh), \\ \frac{u_m^1 - u_m^0}{\tau} = (\psi_2)_m, \end{cases}$$

если  $(\psi_2)_m = \psi_2(mh)$ . За норму  $\| \cdot \|_{F_h}^{(h)}$  принять максимум модулей всех компонент элемента

$$j^{(h)} = \begin{cases} \varphi_m^n, \\ \psi_1(mh), \\ (\psi_2)_m. \end{cases}$$

Показать, что аппроксимация имеет первый порядок относительно  $h$ ;  $\tau = rh$ ,  $r = \text{const}$ .

Как следует задать значения  $(\psi_2)_m$ , используя заданные функции  $\varphi(x, t)$ ,  $\psi_1(x)$  и  $\psi_2(x)$ , чтобы порядок аппроксимации оказался вторым?

2. Для задачи о распространении тепла на отрезке

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} &= \varphi(x, t), & 0 < x < 1, & \quad 0 < t < T, \\ u(x, 0) &= \psi_0(x), & 0 < x < 1, \\ \frac{\partial u(0, t)}{\partial x} &= \psi_1(t), & 0 < t < T, \\ u(1, t) &= \psi_2(t) \end{aligned}$$

рассмотреть разностную схему вида

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2} &= \varphi(mh, n\tau), \\ m = 1, 2, \dots, M-1; \quad n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi_0(nh), \quad m = 0, 1, \dots, M, \\ \frac{u_1^n - u_0^n}{h} &= (\psi_1)^n \quad n = 1, 2, \dots, [T/\tau], \\ u_M^n &= \psi_2(n\tau), \quad n = 1, 2, \dots, [T/\tau]. \end{aligned} \right\}$$

За норму  $\|\cdot\|_{F_h}$  принять максимум абсолютных величин правых частей уравнений, составляющих в совокупности рассматриваемую разностную схему. Шаги  $\tau$  и  $h$  считать связанными равенством  $\tau = rh^2$ ,  $r = \text{const}$ . Показать, что, положив  $(\psi_1)^n = \psi_1(nh)$ , получим схему с первым порядком аппроксимации на гладком решении. Какой формулой следует определить  $(\psi_1)^n$ , чтобы получилась аппроксимация второго порядка?

#### § 24. Условие Куранта, Фридрихса и Леви, необходимое для сходимости

В § 21 мы доказали, что разностная схема

$$\left. \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_m^n}{h} &= 0, \\ u_m^0 &= \psi(mh), \end{aligned} \right\} \quad (1)$$

аппроксимирующая задачу Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= 0, \quad 0 < t < T, \\ u(x, 0) &= \psi(x), \end{aligned} \right\} \quad (2)$$

не может оказаться сходящейся при произвольной функции  $\psi(x)$ , если  $\tau/h > 1$  (см. рис. 9 на стр. 179). При этом было использовано соображение общего характера, впервые на несколько другом примере сформулированное Курантом, Фридрихсом и Леви.

Это соображение часто помогает при конструировании и исследовании разностных схем. Оно состоит в следующем.

**1. Условие Куранта, Фридрихса и Леви.** Допустим, что в постановке дифференциальной задачи участвует некоторая функция  $\psi$  (см., например, (2)). Выберем произвольную точку  $P$ , принадлежащую области определения решения  $u$ . Пусть значение решения  $u(P)$  зависит от значений функции  $\psi$  в точках некоторого множества  $G_\psi = G_\psi(P)$ , принадлежащего области определения функции  $\psi$ , т. е., изменяя значения  $\psi$  в малой

окрестности любой точки  $Q$  из области  $G_\psi(P)$ , можно вызвать изменение значения решения  $u(P)$ . Допустим, что для вычисления решения  $u$  используется некоторая разностная схема  $L_h u^{(h)} = f^{(h)}$ , причем значение решения  $u^{(h)}$  в точке сетки, ближайшей к  $P$ , полностью определяется значениями функции  $\psi$  на некотором множестве  $G_\psi^{(h)} = G_\psi^{(h)}(P)$ .

Для того чтобы имела место сходимость  $u^{(h)} \rightarrow u$  при  $h \rightarrow 0$ , разностная схема необходимо должна быть устроена так, чтобы при  $h \rightarrow 0$  в произвольной окрестности любой точки области  $G_\psi(P)$  при достаточно малом  $h$  имелась точка множества  $G_\psi^{(h)} = G_\psi^{(h)}(P)$ .

Объясним, почему в случае невыполнения сформулированного условия Куранта, Фридрихса и Леви сходимости ожидать не приходится. Пусть оно не выполнено, так что в некоторой фиксированной окрестности некоторой точки  $Q$  из области  $G_\psi(P)$  при всех достаточно малых  $h$  нет точек из множества  $G_\psi^{(h)} = G_\psi^{(h)}(P)$ . Если сходимость  $u^{(h)} \rightarrow u$  и при данной функции  $\psi$  имеет (случайно!) место, то изменим  $\psi$  в указанной окрестности точки  $Q$  так, чтобы изменилось значение  $u(P)$ , оставляя вне этой окрестности функцию  $\psi$  неизменной. Сходимость  $u^{(h)} \rightarrow u$  при новой функции  $\psi$  уже не может иметь места: значение  $u(P)$  изменилось, в то время как значения  $u^{(h)}$  в точке сетки, ближайшей к  $P$ , остались при малых  $h$  неизменными, поскольку функция  $\psi$  в точках множества  $G_\psi^{(h)} = G_\psi^{(h)}(P)$  осталась неизменной.

Условию Куранта, Фридрихса и Леви нетрудно придать форму теоремы, а проведенные рассуждения превратить в ее доказательство, однако мы не будем этого делать.

Рассмотрим несколько примеров, где изложенное нами соображение позволяет установить расходимость и непригодность разностной схемы и нащупать устойчивую и сходящуюся разностную схему. Конечно, доказательство сходимости приходится проводить отдельно, так как выполнение условия Куранта, Фридрихса и Леви лишь необходимо, но недостаточно для сходимости. Заметим также, что при наличии аппроксимации условие Куранта, Фридрихса и Леви необходимо и для устойчивости, поскольку из аппроксимации и устойчивости следует сходимость.

**2. Примеры разностных схем для задачи Коши.** Используем условие Куранта, Фридрихса и Леви для анализа нескольких разностных схем, аппроксимирующих задачу Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} + a(t) \frac{\partial u}{\partial x} &= \psi_0(x, t), & -\infty < x < \infty, & 0 < t < \tau, \\ u(x, 0) &= \psi_1(x), & -\infty < x < \infty, \end{aligned} \right\} \quad (3)$$

где  $\psi_0(x, t)$  и  $\psi_1(x)$  — заданные «входные данные» задачи (3) и

$$a(t) \equiv -1 - 2t.$$

Решение задачи (3) в какой-либо точке  $(x_P, t_P)$  зависит от значений функций  $\psi_0(x, t)$  и  $\psi_1(x)$  во всех тех точках, через

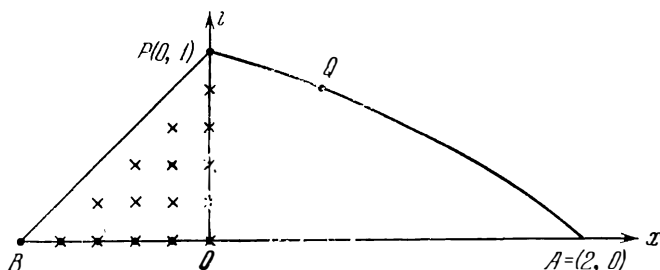


Рис. 16.

которые проходит характеристика дифференциального уравнения (3), выходящая из некоторой точки  $A$  оси  $Ox$  и входящая в точку  $P$ .

Действительно, характеристики здесь — интегральные кривые дифференциального уравнения

$$\frac{dx}{dt} = a(t),$$

т. е. параболы  $x = -t^2 - t + C$ . Вдоль каждой характеристики

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} \frac{dx}{dt} = \frac{\partial u}{\partial t} + a(t) \frac{\partial u}{\partial x} = \psi_0(x, t).$$

Поэтому значение решения  $u(x_P, t_P)$  в какой-либо точке  $P = (x_P, t_P)$  выражается формулой

$$u(x_P, t_P) = \psi_1(A) + \int_0^{t_P} \psi_0[x(t), t] dt = \psi_1(A) + \int_{AQP} \psi_0(x, t) dt,$$

где  $A$  есть точка на оси  $Ox$ , а  $AQP$  — отрезок характеристики.

На рис. 16 изображена характеристика  $x = 2 - t - t^2$ , выходящая из точки  $A = (2, 0)$  и входящая в точку  $P = (0, 1)$ . Мы видим, что значение  $u(P) \equiv u(x_P, t_P)$  решения задачи (3) зависит от значения функции  $\psi_1(x)$  в точке  $A$ , так что  $A = G_{\psi_1}(P)$ . Далее,  $u(P)$  зависит от значений  $\psi_0(x, t)$  на отрезке характеристики  $AQP$ . Этот отрезок  $AQP$  и есть  $G_{\psi_0}(P)$ .

Рассмотрим разностную схему

$$L_h u^{(h)} \equiv \left\{ \begin{array}{l} \frac{u_m^{n+1} - u_m^n}{\tau} + a(t_n) \frac{u_m^n - u_{m-1}^n}{h} = \psi_0(x_m, t_n) \\ u_m^0 = \psi_1(x_m), \quad m = 0, \pm 1, \dots; \\ n = 0, 1, \dots, [1/\tau] - 1, \end{array} \right. \quad (4)$$

или

$$\left. \begin{array}{l} u_m^{n+1} = [1 + a(t_n)r]u_m^n - a(t_n)ru_{m-1}^n + \tau\psi_0(x_m, t_n), \\ u_m^1 = \psi_1(x_m), \end{array} \right\} \quad (5)$$

где  $x_m = mh$ ,  $t_n = n\tau$ ,  $r = \tau/h$ ,  $a(t) \equiv -1 - 2t$ . Покажем, что эта схема не может быть сходящейся ни при каком соотношении шагов  $r$ , так как ни при каком  $r$  она не удовлетворяет условию Куранта, Фридрихса и Леви.

Возьмем в качестве точки  $P$  точку  $(0, 1)$ . Сетку выберем так, чтобы  $N\tau = 1$ . Значение решения  $u^{(h)} = u^{(h)}(P)$  в точке  $P = (0, 1)$ , т. е.  $u_0^N$ , в силу разностной формулы (5) выразится через значение  $\psi_0(0, 1 - \tau)$  и через значения  $u_{-1}^{N-1}$ ,  $u_0^{N-1}$ . Эти два значения в свою очередь выразятся через  $\psi_0(-h, 1 - 2\tau)$ ,  $\psi_0(0, 1 - 2\tau)$  и через три значения  $u_{-2}^{N-2}$ ,  $u_{-1}^{N-2}$ ,  $u_0^{N-2}$  и т. д. В конечном счете значение  $u_0^N$  выразится через значения функции  $\psi_0(x, t)$  в точках сетки, отмеченных на рис. 16 крестиками, и через значения  $u_{-N}^1 = \psi_1(x_{-N})$ ,  $u_{-N+1}^1 = \psi_1(x_{-N+1})$ ,  $\dots$ ,  $u_0^1 = \psi_1(x_0)$  функции  $\psi_1(x)$  в точках  $x_{-N}$ ,  $x_{-N+1}$ ,  $\dots$ ,  $x_0$  на оси  $Ox$ . Таким образом, множество  $G_{\psi_0}^{(h)}(P)$  состоит из точек сетки, отмеченных крестиками, а множество  $G_{\psi_1}^{(h)}(P)$  — из точек  $x_{-N}$ ,  $x_{-N+1}$ ,  $\dots$ ,  $x_0$  на оси  $Ox$  (эти множества имеют общие точки на оси  $Ox$ ). Очевидно, что любая точка  $Q$  множества  $G_{\psi_1}^{(h)}(P)$  имеет окрестность, в которую не попадают точки множества  $G_{\psi_0}^{(h)}(P)$ , как бы мало ни было  $h$ . Разностная схема (4) не удовлетворяет условию Куранта, Фридрихса и Леви, необходимому для сходимости.

Рассмотрим теперь для задачи (3) разностную схему (рис. 17)

$$L_h u^{(h)} \equiv \left\{ \begin{array}{l} \frac{u_m^{n+1} - u_m^n}{\tau} + a(t_n) \frac{u_{m+1}^n - u_m^n}{h} = \psi_0(x_m, t_n), \\ u_m^0 = \psi_1(x_m), \quad m = 0, \pm 1, \dots; \\ n = 0, 1, \dots, 1/\tau - 1, \end{array} \right. \quad (6)$$

ИЛИ

$$\left. \begin{aligned} u_m^{n+1} &= [1 + a(t_n)r]u_m^n - a(t_n)ru_{m+1}^n + \tau\psi_0(x_m, t_n), \\ u_m^0 &= \psi_1(x_m), \end{aligned} \right\} \quad (7)$$

где  $r = \tau/h$ .

Шаг  $\tau$  сетки выберем из условия  $N\tau = 1$ ,  $N$  — целое, так что точка  $P = (0, 1)$  будет принадлежать сетке. Значение решения  $u^{(h)}$  в этой точке, т. е.  $u_0^N$ , в силу формул (7) выразится через  $\psi_0(0, 1 - \tau)$  и два значения  $u_0^{N-1}$  и  $u_1^{N-1}$ . Эти два значения в свою очередь в силу (7) выражаются через  $\psi_0(0, 1 - 2\tau)$ ,  $\psi_0(h, 1 - 2\tau)$  и через три значения  $u_0^{N-2}$ ,  $u_1^{N-2}$ ,  $u_2^{N-2}$  и т. д. В конечном счете  $u_0^N$  выражается через значения  $\psi_0(x_m, t_n)$  в точках

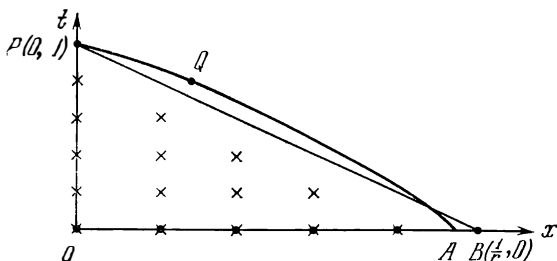


Рис. 17.

сетки, отмеченных на рис. 17 крестиками, и через значения  $u_0^0 = \psi_1(0)$ ,  $u_1^0 = \psi_1(x_1)$ , ...,  $u_N^0 = \psi_1(x_N)$  функции  $\psi_1(x)$  в точках  $x_0, x_1, \dots, x_N$  на оси  $Ox$ . Таким образом,  $G_{\psi_0}^{(h)}(P)$  в этом случае — это множество точек, отмеченных крестиками, а  $G_{\psi_1}^{(h)}(P)$  — это множество точек  $x_0, x_1, \dots, x_N$  на оси  $Ox$ . Ясно, что в случае  $r = \tau/h > 1/2$  (этот случай не изображен на рисунке) точка  $B = (1/r, 0)$  лежит левее точки  $A = G_{\psi_0}(P)$ . Поэтому существует окрестность точки  $A$ , в которую не попадают при  $h \rightarrow 0$  точки  $G_{\psi_1}^{(h)}(P)$ . Условие Куранта, Фридрихса и Леви нарушено, и сходимости ожидать нельзя.

Для того чтобы схема (6) могла оказаться сходящейся, необходимо, чтобы  $r \leq 1/2$ . Но этого мало. Допустим, что  $r < 1$ , но некоторая точка  $Q$  характеристики  $AQP$  лежит над прямой  $BP$ , как на рис. 17. Тогда тоже нельзя ожидать сходимости. Значение функции  $\psi_0(x, t)$  в точке  $Q$  оказывает влияние на значение  $u(0, 1)$  решения дифференциальной задачи, т. е.  $Q$  принадлежит множеству  $G_{\psi_0}(P)$ . Но значение  $\psi_0(x, t)$  в точке  $Q$  (как и значения  $\psi_0(x, t)$  на всем участке  $QP$  характеристики) не оказывает влияния на значение  $u^{(h)}(P)$  решения разностного

уравнения в точке  $P$ : существует окрестность точки  $Q$ , куда при  $h \rightarrow 0$  не попадают точки множества  $G_{\psi}^{(h)}(P)$ . Условие Куранта, Фридрихса и Леви не выполнено.

Выбрав  $r$  настолько малым, чтобы треугольник  $OPB$  содержал не только точку  $A = (2, 0)$ , но и всю характеристику  $AQP$ , уже можно доказать устойчивость (и сходимую) разностной схемы (6). Для такого выбора числа  $r$  учтем, что (в силу дифференциального уравнения характеристики  $dx/dt = a(t)$ ) величина  $-1/a(t)$  есть тангенс угла наклона касательной к характеристике к оси  $Ox$ , а  $-r = -\tau/h$  есть тангенс угла наклона прямой  $BP$  к оси  $Ox$ . Легко понять, что характеристика  $AQP$  будет лежать в треугольнике  $BOP$ , если

$$r \leq \frac{1}{\max_{0 \leq t \leq 1} |a(t)|} = \frac{1}{3}, \quad \tau \leq \frac{h}{3}, \quad (8)$$

и тогда условие Куранта, Фридрихса и Леви выполнено.

Покажем, что при условии (8) разностная схема (6), аппроксимирующая задачу Коши (3), устойчива, и следовательно, сходится. При этом нормы определим равенствами

$$\begin{aligned} \|u^{(h)}\|_{U_h} &= \max_{m, n} |u_m^n|, \\ \|f^{(h)}\|_{F_h} &= \max_{m, n} |\psi_0(x_m, t_n)| + \max_m |\psi_1(x_m)|. \end{aligned}$$

Учитывая, что при условии (8)

$$1 + a(t_n)r \geq 1 - \frac{2t_n + 1}{3} \geq 0, \quad 0 \leq t_n \leq 1,$$

из равенства (7) получим

$$\begin{aligned} |u_m^{n+1}| &\leq \left[ 1 - \frac{2t_n + 1}{3} r + \frac{2t_n + 1}{3} r \right] \max_m |u_m^n| + \tau \max_{m, n} |\psi_0(x_m, t_n)| \leq \\ &\leq \max_m |u_m^n| + \tau \max_{m, n} |\psi_0(x_m, t_n)| \leq \\ &\leq \max_m |u_m^{n-1}| + 2\tau \max_{m, n} |\psi_0(x_m, t_n)| \leq \\ &\dots \dots \dots \\ &\leq \max_m |u_m^0| + (n+1)\tau \max_{m, n} |\psi_0(x_m, t_n)| \leq \\ &\leq \max_m |\psi_1(x_m)| + 1 \cdot \max_{m, n} |\psi_0(x_m, t_n)| = \|f^{(h)}\|_{F_h}. \end{aligned}$$

Поскольку полученное неравенство

$$|u_m^{n+1}| \leq \|f^{(h)}\|_{F_h}$$

справедливо при любых  $m = 0, \pm 1, \dots$  и любых  $n$ ,  $(n+1)\tau \leq \leq 1$ , то

$$\|u^{(h)}\|_{U_h} \leq \|f^{(h)}\|_{F_h},$$

и устойчивость разностной схемы (6) при условии (8) доказана. Ограничение (8) на шаг  $\tau$  при заданном шаге  $h$ ,  $\tau \leq 1/3h$ , можно ослабить, не нарушая условия Куранта, Фридрихса и Леви, если

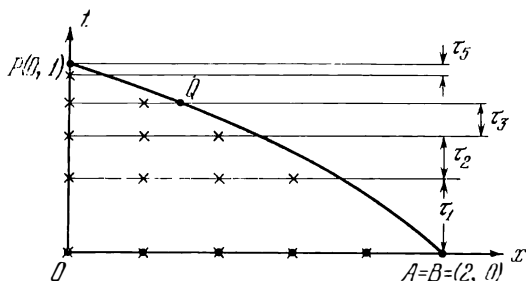


Рис. 18

сделать шаг  $\tau$  переменным,  $t_{n+1} = t_n + \tau_n$ , и выбирать его при переходе от  $t_n$  к  $t_{n+1}$  с учетом наклона характеристики вблизи точки  $t = t_n$ , а именно из условия

$$r_n \equiv \frac{\tau_n}{h} \leq \frac{1}{|a(t_n)|} = \frac{1}{2t_n + 1}, \quad n = 0, 1, \dots \quad (9)$$

Измененная таким образом схема (6) имеет вид

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau_n} + a(t_n) \frac{u_{m+1}^n - u_m^n}{h} = \psi_0(x_m, t_n), \\ u_m^0 = \psi_1(x_m) \end{cases} \quad (10)$$

или

$$\left. \begin{aligned} u_m^{n+1} &= [1 + a(t_n)r_n]u_m^n - a(t_n)r_n u_{m+1}^n + \tau_n \psi_0(x_m, t_n), \\ u_m^0 &= \psi_1(x_m). \end{aligned} \right\} \quad (11)$$

В соответствии с формулой (9) ограничение на шаг  $\tau_n$  менее жесткое, чем при использовании схемы (6) с постоянным шагом. При малых  $n$  используется шаг  $\tau_n \approx h$ , и лишь при приближении  $t_n$  к  $t = 1$  приходится выбирать  $\tau_n = 1/3h$  (рис. 18). Доказательство устойчивости схемы (10) при условии (9) лишь несущественно отличается от доказательства устойчивости схемы (6) при условии (8): используя неравенство  $1 + a(t_n)r_n \geq 0$ ,



получим в силу (11)

$$\begin{aligned}
 |u_m^{n+1}| &\leq \max_m |u_m^n| + \tau_n \max_{m,n} |\psi_0(x_m, t_n)| \leq \\
 &\leq \max_m |u_m^{n-1}| + (\tau_{n-1} + \tau_n) \max_{m,n} |\psi_0(x_m, t_n)| \leq \\
 &\leq \max_m |u_m^0| + t_{n+1} \max_{m,n} |\psi_0(x_m, t_n)| \leq \|f^{(h)}\|_{F_h}.
 \end{aligned}$$

Отсюда следует неравенство

$$\|u^{(h)}\|_{U_h} \leq \|f^{(h)}\|_{F_h},$$

означающее устойчивость.

**3. Примеры разностных схем для задачи Дирихле.** Воспользуемся условием Куранта, Фридрихса и Леви для анализа двух

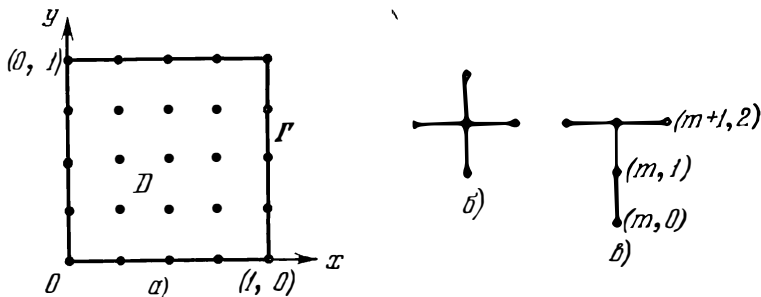


Рис. 19.

разностных схем, аппроксимирующих следующую задачу Дирихле для уравнения Пуассона:

$$\left. \begin{aligned}
 \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= \varphi(x, y), & 0 \leq x, y \leq 1, \\
 u|_{\Gamma} &= \psi(x, y), & (x, y) \in \Gamma,
 \end{aligned} \right\} \quad (12)$$

в квадратной области  $D = (0 \leq x, y \leq 1)$  с границей  $\Gamma$ . Построим сетку  $x_m = mh, y_n = nh$ , где  $h = 1/M, M$  — целое число (рис. 19, a). К сетке  $D_h$  отнесем те точки  $(x_m, y_n)$ , которые попали внутрь квадрата  $D$  или на его границу. Рассмотрим разностную схему, аппроксимирующую задачу (12):

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{m+1,n} - 2u_{mn} + u_{m-1,n}}{h^2} + \frac{u_{m,n+1} - 2u_{mn} + u_{m,n-1}}{h^2} = \\ = \varphi(mh, nh), & \text{если } (mh, nh) \in D, \\ u_{mn} = \psi(mh, nh), & \text{если } (mh, nh) \in \Gamma. \end{cases} \quad (13)$$

Схема (13) получена путем замены производных  $u_{xx}$  и  $u_{yy}$  разностными отношениями, и аппроксимация не вызывает сомнения. Мы докажем ее устойчивость в § 34 и изложим способы вычисления решения  $u^{(h)}$  в §§ 35—37. Но обратим внимание на то, что вычислить ее решение не просто, так как система уравнений  $L_h u^{(h)} = f^{(h)}$  для определения значений сеточной функции  $u^{(h)}$  при малых  $h$  достаточно сложна. Сама эта сложность побуждает выяснить, нельзя ли построить схему, решения которой вычисляются просто. На первый взгляд, можно воспользоваться схемой

$$L_h u^{(h)} \equiv \left\{ \begin{array}{l} \frac{u_{m-1, n} - 2u_{mn} + u_{m+1, n}}{h^2} + \frac{u_{m, n+1} - 2u_{mn} + u_{m, n-1}}{h^2} = \\ \phantom{\frac{u_{m-1, n} - 2u_{mn} + u_{m+1, n}}{h^2} + \frac{u_{m, n+1} - 2u_{mn} + u_{m, n-1}}{h^2}} = \varphi(mh, nh), \\ m = 1, 2, \dots, M-1; n = 1, 2, \dots, M-2, \\ \frac{u_{m-1, 2} - 2u_{m2} + u_{m+1, 2}}{h^2} + \frac{u_{m2} - 2u_{m1} + u_{m0}}{h^2} = \\ \phantom{\frac{u_{m-1, 2} - 2u_{m2} + u_{m+1, 2}}{h^2} + \frac{u_{m2} - 2u_{m1} + u_{m0}}{h^2}} = \varphi(mh, h), \quad m = 1, 2, \dots, M-1, \\ u_{mn} = \psi(mh, nh), \quad (x, y) \in \Gamma. \end{array} \right. \quad (14)$$

Аппроксимация, очевидно, в этом случае есть, так как схема получена путем замены производных разностными отношениями, а граничное условие аппроксимировано точно. Каждое уравнение из первой группы уравнений связывает значения решения в пяти точках сетки, изображенных на рис. 19, б. Вторая группа уравнений при фиксированном  $m$  связывает значения решения в пяти точках сетки, изображенных на рис. 19, в.

Рассмотрим совокупность уравнений первой группы, отвечающих фиксированному значению  $n$ , а именно  $n = 1$ , и всю вторую группу уравнений совместно. Полученная система уравнений связывает значения  $u_{m1}$ ,  $u_{m2}$  и  $u_{m0}$ , причем  $u_{m0}$ ,  $u_{01}$ ,  $u_{02}$ ,  $u_{M1}$ ,  $u_{M2}$  заданы граничными условиями. Эту систему можно решить, определив  $u_{m1}$  и  $u_{m2}$ ,  $m = 1, 2, \dots, M-1$ . Затем используем разностное уравнение из первой группы уравнений при  $n = 2$  и определим  $u_{m3}$  по явной формуле, разрешая это уравнение относительно единственной входящей в него неизвестной величины  $u_{m3}$ . Продвигаясь слой за слоем от  $u_{mn}$  к  $u_{m, n+1}$ , мы вычислим в силу уравнений первой группы решение  $u^{(h)}$  во всех внутренних точках сетки. Значения же в граничных точках сетки заданы с самого начала.

Однако эта на первый взгляд удобная схема совершенно непригодна. Известно, что решение задачи Дирихле для уравнения Лапласа зависит в каждой точке от значений  $\psi(x, y)|_{\Gamma}$  всюду на границе. А в построенной нами разностной схеме вычисление решения  $u^{(h)}$  во всех внутренних точках происходит без ис-

пользования значения  $\psi(x, y)$  на верхней стороне квадрата. Эта разностная схема не может оказаться сходящейся. Сложность, присущая схеме (13), связана с существом дела.

В заключение подчеркнем еще раз, что условие Куранта, Фридрихса и Леви не является достаточным условием устойчивости. В § 25 будет, в частности, показано, что разностная схема

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} = \varphi(mh, n\tau), \\ u_m^0 = \psi(mh) \end{cases}$$

неустойчива при любом  $r = \tau/h = \text{const}$ . Эта схема аппроксимирует задачу Коши

$$u_t - u_x = \varphi(x, t),$$

$$u(x, 0) = \psi(x),$$

для которой мы уже рассмотрели несколько других схем. Легко проверить в то же время, что эта схема при  $r \leq 1$  удовлетворяет необходимому условию устойчивости.

Чтобы сделать это, возьмем опять, для определенности, точку  $(0, 1)$  на плоскости  $Oxt$  и будем считать, что она принадлежит сетке  $D_h$  при всех  $h$ , так что  $1 = N\tau$ , где  $N$  — целое. Значение  $u_0^N$  вычисляется через значения  $u_{-1}^{N-1}$ ,  $u_0^{N-1}$ ,  $u_1^{N-1}$ . Эти три значения вычисляются затем через пять значений на предыдущем слое  $t = (N-2)\tau$  и т. д. В конечном счете  $u_0^N$  вычисляется, очевидно, через значения  $u_m^0 = \psi(mh)$ ,  $m = -N, -N+1, \dots, -1, 0, 1, \dots, N$ , в точках сетки, принадлежащих отрезку  $-1/r \leq x \leq 1/r$  на оси  $Ox$ . Если  $r = \tau/h < 1$ , то этот отрезок содержит точку  $x = 1$ , значение в которой определяет  $u(0, 1)$ ,  $u(0, 1) = \psi(1)$ . Условие Куранта, Фридрихса и Леви при  $r \leq 1$  выполнено.

## ЗАДАЧИ

1. Решение уравнения теплопроводности  $u_t = u_{xx}$ ,  $-\infty < x < \infty$ ,  $t > 0$ , имеет вид

$$u(x, t) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{t}} u(\xi, 0) e^{-\frac{(x-\xi)^2}{4t}} d\xi.$$

Может ли существовать сходящаяся разностная схема, аппроксимирующая эту задачу и имеющая вид

$$\frac{u_m^{p+1} - u_m^p}{\tau} = \frac{1}{h^2} (\alpha_{-2} u_{m-2}^p + \alpha_{-1} u_{m-1}^p + \alpha_0 u_m^p + \alpha_1 u_{m+1}^p + \alpha_2 u_{m+2}^p),$$

где  $\alpha_j$  — некоторые постоянные, если  $\tau = h^2$

2. Система уравнений акустики

$$\left. \begin{aligned} \frac{\partial v}{\partial t} + \frac{\partial w}{\partial x} &= 0, \\ \frac{\partial w}{\partial t} + \frac{\partial v}{\partial x} &= 0, \\ v(x, 0) &= \varphi(x), \quad w(x, 0) = \psi(x) \end{aligned} \right\} \quad t > 0, \quad -\infty < x < \infty,$$

имеет решение вида

$$\left. \begin{aligned} v(x, t) &= \frac{\varphi(x-t) + \psi(x-t) + \varphi(x+t) - \psi(x+t)}{2}, \\ w(x, t) &= \frac{\varphi(x-t) + \psi(x-t) - \varphi(x+t) + \psi(x+t)}{2}. \end{aligned} \right\}$$

Может ли оказаться сходящейся разностная схема вида

$$\left. \begin{aligned} \frac{v_m^{p+1} - v_m^p}{\tau} + \frac{w_{m+1}^p - w_m^p}{h} &= 0, & p \geq 0, \\ \frac{w_m^{p+1} - w_m^p}{\tau} + \frac{v_{m+1}^p - v_m^p}{h} &= 0, & m = 0, \pm 1, \dots, \\ v_m^0 &= \varphi(x_m), & w_m^0 = \psi(x_m) \end{aligned} \right\}$$

Сопоставьте области влияния начальных данных для разностной и дифференциальной задач.

### 3. Задача Коши

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= 0, & t > 0, & -\infty < x < \infty, \\ u(x, 0) &= e^{iax}, & -\infty < x < \infty, \end{aligned}$$

имеет решение

$$u(x, t) = e^{iat} e^{iax}.$$

Соответствующая разностная схема

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= 0, & p = 0, 1, \dots, \\ u_m^0 &= e^{iahm}, & m = 0, \pm 1, \dots, \end{aligned} \right\}$$

имеет решение

$$u_m^p = [1 - r + r e^{iah}]^p e^{iahm},$$

которое при  $p = t/\tau$ ,  $m = x/h$  стремится к решению дифференциальной задачи при  $h \rightarrow 0$ , каково бы ни было фиксированное  $r = \tau/h$ . Между тем при  $r > 1$  разностная схема не удовлетворяет условию Куранта, Фридрикса и Леви, необходимому для сходимости. Объясните кажущийся парадокс.

НЕКОТОРЫЕ ОСНОВНЫЕ ПРИЕМЫ  
ИССЛЕДОВАНИЯ УСТОЙЧИВОСТИ

§ 25. Спектральный анализ разностной задачи Коши

Мы изложим широко применяемый способ Неймана исследования разностных задач с начальными данными. В этом параграфе ограничимся случаем разностной задачи Коши с постоянными коэффициентами, а в § 26 частично распространим результаты на случай переменных коэффициентов.

**1. Устойчивость по начальным данным.** Простейшим примером разностной задачи Коши может служить неоднократно встречавшаяся выше задача

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = \varphi_m^p, & p = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 = \psi_m, & m = 0, \pm 1, \dots \end{cases} \quad (1)$$

Положив

$$f^{(h)} = \begin{cases} \varphi_m^p, & p = 0, 1, \dots, [T/\tau] - 1, \\ \psi_m, & m = 0, \pm 1, \dots \end{cases}$$

запишем задачу (1) в форме

$$L_h u^{(h)} = f^{(h)}. \quad (2)$$

Определим нормы  $\|u^{(h)}\|_{U_h}$  и  $\|f^{(h)}\|_{F_h}$  равенствами

$$\|u^{(h)}\|_{U_h} = \max_p \max_m |u_m^p|, \quad \|f^{(h)}\|_{F_h} = \max_m |\psi_m| + \max_{m,p} |\varphi_m^p|.$$

Тогда условие устойчивости задачи (2)

$$\|u^{(h)}\|_{U_h} \leq c \|f^{(h)}\|_{F_h} \quad (3)$$

примет вид

$$\max_m |u_m^p| \leq c [\max_m |\psi_m| + \max_{m,k} |\varphi_m^k|], \quad p = 0, 1, \dots, [T/\tau], \quad (4)$$

где  $c$  не зависит от  $h$  (и от  $\tau = rh$ ,  $r = \text{const}$ ). Условие (4) должно выполняться при произвольных  $\{\psi_m\}$  и  $\{\varphi_m^p\}$ . В частности, для устойчивости необходимо, чтобы оно выполнялось при произвольных  $\{\psi_m\}$  и  $\varphi_m^p \equiv 0$ , т. е. чтобы решение задачи

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \quad p = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 = \psi_m, \quad m = 0, \pm 1, \dots, \end{aligned} \right\} \quad (5)$$

удовлетворяло условию

$$\max_m |u_m^p| \leq c \max_m |u_m^0|, \quad p = 0, 1, \dots, [T/\tau], \quad (6)$$

при произвольной ограниченной функции  $u_m^0 = \psi_m$ .

Свойство (6), необходимое для устойчивости (4) задачи (1), называют *устойчивостью задачи (1) относительно возмущения начальных данных*. Оно означает, что возмущение  $\{u_m^0\}$ , внесенное в начальные данные задачи (1), вызовет возмущение  $\{u_m^p\}$  решения задачи (1), которое в силу (6) не более чем в  $c$  раз превосходит возмущение начальных данных, причем  $c$  не зависит от  $h$ .

**2. Необходимое спектральное условие устойчивости.** Для устойчивости задачи Коши (1) по начальным данным необходимо, чтобы условие (6) выполнялось, в частности, если  $\{u_m^0\}$  есть какая-нибудь гармоника

$$u_m^0 = e^{iam}, \quad m = 0, \pm 1, \dots, \quad (7)$$

где  $\alpha$  — вещественный параметр. Но решение задачи (5) при начальном условии (7) имеет вид

$$u_m^p = \lambda^p e^{iam}, \quad (8)$$

где  $\lambda = \lambda(\alpha)$  определяется путем подстановки выражения (8) в однородное разностное уравнение задачи (5):

$$\lambda(\alpha) = 1 - r + r e^{i\alpha}, \quad r = \frac{\tau}{h} = \text{const}. \quad (9)$$

Для решения (8) справедливо равенство

$$\max_m |u_m^p| = |\lambda(\alpha)|^p \max_m |u_m^0|.$$

Поэтому для выполнения условия (6) необходимо, чтобы при всех вещественных  $\alpha$  выполнялось неравенство

$$|\lambda(\alpha)|^p \leq c, \quad p = 0, 1, \dots, [T/\tau],$$

или

$$|\lambda(\alpha)| \leq 1 + c_1 \tau, \quad (10)$$

где  $c_1$  — некоторая постоянная, не зависящая от  $\alpha$  и  $\tau$ . Это и есть *необходимое спектральное условие Неймана* применительно к рассматриваемому примеру. Спектральным оно называется по следующей причине. Существование решения вида (8) показывает, что гармоника  $\{e^{iam}\}$  является собственной функцией оператора перехода

$$u_m^{p+1} = (1-r)u_m^p + ru_{m+1}^p, \quad m = 0, \pm 1, \dots,$$

который в силу разностного уравнения (5) ставит в соответствие сеточной функции  $\{u_m^p\}$ ,  $m = 0, \pm 1, \dots$ , определенной на слое  $t_p = p\tau$  сетки, сеточную функцию  $\{u_m^{p+1}\}$ ,  $m = 0, \pm 1, \dots$ , определенную на слое  $t_{p+1} = (p+1)\tau$ . Число  $\lambda(\alpha) = 1 - r + re^{ia}$  является соответствующим этой гармонике  $\{e^{iam}\}$  собственным числом оператора перехода. Линия, которую пробегает точка  $\lambda(\alpha)$  на комплексной плоскости, когда  $\alpha$  пробегает вещественную ось, вся состоит из собственных значений и является спектром оператора перехода.

Таким образом, необходимое условие устойчивости (10) можно сформулировать так: спектр оператора перехода, соответствующего разностному уравнению задачи (5), должен лежать в круге радиуса  $1 + c_1\tau$  на комплексной плоскости. В нашем примере спектр (9) не зависит от  $\tau$ . Поэтому условие (10) равносильно требованию, чтобы спектр  $\lambda(\alpha)$  лежал в единичном круге

$$|\lambda(\alpha)| \leq 1. \quad (11)$$

Воспользуемся сформулированным признаком для анализа устойчивости задачи (1). Спектр (9) представляет собою окружность с центром в точке  $1-r$  и радиусом  $r$  на комплексной плоскости. В случае  $r < 1$  эта окружность лежит внутри единичного круга (касаясь его в точке  $\lambda = 1$ ), при  $r = 1$  совпадает с единичной окружностью, а при  $r > 1$  лежит вне единичного круга (рис. 20). Соответственно

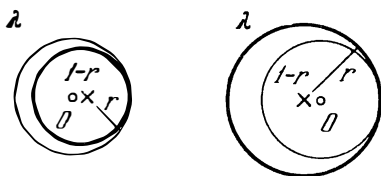


Рис. 20.

необходимое условие устойчивости (11) выполнено при  $r \leq 1$  и не выполнено при  $r > 1$ . В п. 3 § 21 мы исследовали рассматриваемую разностную схему и показали, что при  $r \leq 1$  она устойчива, а при  $r > 1$  неустойчива. Таким образом, необходимое условие устойчивости Неймана оказалось в данном случае достаточно чувствительным, чтобы в точности отделить случай устойчивости от случая неустойчивости.

В общем случае задачи Коши для разностных уравнений и систем разностных уравнений необходимый спектральный признак устойчивости Неймана состоит в том, что спектр  $\lambda = \lambda(\alpha, h)$  разностной задачи при всех достаточно малых  $h$  должен лежать в круге

$$|\lambda| \leq 1 + \varepsilon \quad (12)$$

на комплексной плоскости, как бы мало ни было заранее выбранное положительное число  $\varepsilon$ .

Заметим, что если для рассматриваемой разностной задачи спектр окажется не зависящим от  $h$  (и от  $\tau$ ), то условие (12) равносильно требованию, чтобы спектр  $\lambda = \lambda(\alpha, h) = \lambda(\alpha)$  лежал в единичном круге

$$|\lambda(\alpha)| \leq 1. \quad (12')$$

Под спектром разностной задачи в условии (12) понимается совокупность всех  $\lambda = \lambda(\alpha, h)$ , при которых соответствующее однородное разностное уравнение (или система уравнений) имеет решение вида

$$u_m^p = [\lambda(\alpha, h)]^p [u^0 e^{i\alpha m}], \quad m = 0, \pm 1, \dots, \quad (13)$$

где  $u^0$  — число (единица), если речь идет о скалярном разностном уравнении, и числовой вектор, если речь идет о векторном разностном уравнении, т. е. о системе скалярных разностных уравнений.

Если необходимое условие Неймана (12) не выполнено, то ни при каком разумном выборе норм нельзя ожидать устойчивости, а в случае его выполнения можно надеяться, что при некотором разумном выборе норм устойчивость имеет место. Аналогичный вопрос о независимости спектрального признака устойчивости от выбора норм мы уже обсуждали в связи с разностными схемами для обыкновенных дифференциальных уравнений в § 15.

**3. Примеры.** Рассмотрим ряд интересных разностных задач Коши и применим для анализа их устойчивости спектральный признак Неймана. Начнем с разностных схем, аппроксимирующих дифференциальную задачу Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= \varphi(x, t), & -\infty < x < \infty, & 0 < t < T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty. \end{aligned} \right\} \quad (14)$$

**Пример 1.** Рассмотрим разностную схему

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_m^p - u_{m-1}^p}{h} = \varphi(x_m, t_p), \quad p = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 = \psi(x_m), \quad m = 0, \pm 1, \dots$$



Подставляя выражение вида (8) в соответствующее однородное разностное уравнение, после простых преобразований получим

$$\lambda(\alpha) = 1 + r - re^{-i\alpha}.$$

В силу этой формулы спектр представляет собою окружность с центром в точке  $1 + r$  и радиусом  $r$  (рис. 21). Ни при каком  $r$  спектр не лежит в единичном круге. Условие устойчивости (12') всегда не выполнено.

В § 24 уже было установлено, что при любом  $r$  не выполнено необходимое условие сходимости (и устойчивости) Куранта, Фридрихса и Леви.

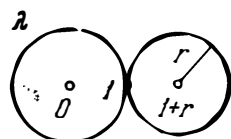


Рис. 21.

Пример 2. Рассмотрим следующую разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \frac{\tau}{2h^2} (u_{m-1}^p - 2u_m^p + u_{m+1}^p) &= \varphi(x_m, t_p), \\ u_m^0 &= \psi_m, \end{aligned} \right\} \quad (15)$$

аппроксимирующую задачу (14) со вторым порядком относительно  $h$  (§ 22). Для нее  $\lambda = \lambda(\alpha)$  определяется из уравнения

$$\frac{\lambda - 1}{\tau} - \frac{e^{i\alpha} - e^{-i\alpha}}{2h} - \frac{\tau}{2h^2} (e^{i\alpha} - 2 + e^{-i\alpha}) = 0.$$

Обозначим по-прежнему  $r = \tau/h$ . Заметив, что

$$\begin{aligned} \frac{e^{i\alpha} - e^{-i\alpha}}{2i} &= \sin \alpha, \\ \frac{e^{i\alpha} - 2 + e^{-i\alpha}}{4} &= - \left( e^{\frac{i\alpha}{2}} - e^{-\frac{i\alpha}{2}} \right)^2 = - \sin^2 \frac{\alpha}{2}, \end{aligned}$$

получим

$$\begin{aligned} \lambda &= 1 + ir \sin \alpha - 2r^2 \sin^2 \frac{\alpha}{2}, \\ |\lambda(\alpha)|^2 &= \left( 1 - 2r^2 \sin^2 \frac{\alpha}{2} \right)^2 + r^2 \sin^2 \alpha. \end{aligned} \quad (16)$$

После простых преобразований найдем

$$1 - |\lambda|^2 = 4r^2 \sin^4 \frac{\alpha}{2} (1 - r^2). \quad (17)$$

Условие Неймана выполнено, если правая часть неотрицательна,  $r \leq 1$ , и не выполнено при  $r > 1$ .

Пример 3. Рассмотрим следующую разностную схему

$$L_h u^{(h)} = \begin{cases} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_{m-1}^p}{2h} = \varphi(x_m, t_p), \\ u_m^0 = \psi(x_m) \end{cases} \quad (18)$$

для той же задачи Коши (14).

Подставляя в уравнение (18) выражение (8), после сокращений получим уравнение для  $\lambda$ :

$$\lambda \begin{array}{c} \circ \\ \text{---} \\ \circ \end{array} \begin{array}{l} /+i.r \\ /-i.r \end{array} \quad \text{или} \quad \frac{\lambda - 1}{\tau} - \frac{e^{i\alpha} - e^{-i\alpha}}{2h} = 0$$

$$\lambda(\alpha) = 1 + i \left( \frac{\tau}{h} \sin \alpha \right).$$

Рис. 22.

Спектр  $\lambda = \lambda(\alpha)$  заполняет вертикальный отрезок длины  $2\tau/h$ , проходящий через точку  $\lambda = 1$  (рис. 22).

Если  $\tau/h = r = \text{const}$ , то условие (12') не выполняется — спектр не лежит в единичном круге. Если при  $h \rightarrow 0$  шаг  $\tau$  изменяется, как  $h^2$ , так что  $\tau = rh^2$ , то самая далекая от точки  $\lambda = 0$  точка  $\lambda(\alpha)$  имеет модуль

$$|\lambda(\alpha)|_{\alpha=\pi/2} = \sqrt{1 + \left(\frac{\tau}{h}\right)^2} = \sqrt{1 + \tau r} \leq 1 + \frac{r}{2} \tau.$$

Условие  $|\lambda(\alpha)| \leq 1 + c\tau$  в этом случае выполнено при  $c = r/2$ .

Ясно, что требование  $\tau = rh^2$  является гораздо более жестким условием на убывание шага по времени  $\tau$  при стремлении шага  $h$  к нулю, чем требование  $\tau = rh$ ,  $r \leq 1$ , которого было достаточно для выполнения признака Неймана для разностных схем (5) и (15), аппроксимирующих ту же задачу Коши (14).

Отметим, что признак Куранта, Фридрихса и Леви, как показано в конце § 24, позволяет утверждать неустойчивость обсуждаемой схемы только при  $\tau/h > 1$ , а при  $\tau/h \leq 1$  суждений об устойчивости не дает и оказывается слабее признака Неймана.

Рассмотрим теперь две построенные в § 22 разностные схемы, аппроксимирующие задачу Коши для уравнения теплопроводности

$$\left. \begin{array}{l} u_t - a^2 u_{xx} = \varphi(x, t), \quad -\infty < x < \infty, \quad 0 < t < T, \\ u(x, 0) = \psi(x), \quad -\infty < x < \infty. \end{array} \right\} \quad (19)$$

Пример 4. Явная разностная схема

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{p+1} - u_m^p}{\tau} - a^2 \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = \varphi(mh, p\tau), \\ u_m^0 = \psi(mh), \quad m = 0, \pm 1, \dots; \quad p = 0, 1, \dots, [T/\tau] - 1, \end{cases}$$

при подстановке  $u_m^p = \lambda^p e^{i\alpha m}$  в соответствующее однородное разностное уравнение приводит к соотношению

$$\frac{\lambda - 1}{\tau} - a^2 \frac{e^{-i\alpha} - 2 + e^{i\alpha}}{h^2} = 0.$$

Заметив, что

$$\frac{e^{-i\alpha} - 2 + e^{i\alpha}}{4} = - \left( \frac{e^{\frac{i\alpha}{2}} - e^{-\frac{i\alpha}{2}}}{2i} \right)^2 = - \sin^2 \frac{\alpha}{2},$$

получим

$$\lambda(\alpha) = 1 - 4ra^2 \sin^2 \frac{\alpha}{2}, \quad r = \frac{\tau}{h^2}.$$

При изменении  $\alpha$  число  $\lambda(\alpha)$  пробегает отрезок  $1 - 4ra^2 \leq \lambda \leq 1$  вещественной оси (рис. 23).

Для устойчивости необходимо, чтобы левый конец этого отрезка лежал в единичном круге  $1 - 4ra^2 \geq -1$  или

$$r \leq \frac{1}{2a^2}. \quad (20)$$

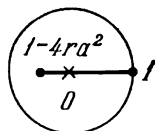


Рис. 23.

В случае, если  $r > \frac{1}{2a^2}$ , точка  $\lambda(\alpha) = 1 - 4ra^2 \sin^2 \frac{\alpha}{2}$ , отвечающая  $\alpha = \pi$ , лежит левее точки  $-1$ . Гармоника  $e^{i\pi m} = (-1)^m$  порождает решение

$$u_m^p = (1 - 4a^2 r)^p (-1)^m,$$

не удовлетворяющее условию (6) ни при какой постоянной  $c$ .

Пример 5. Рассмотрим теперь вторую схему

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{p+1} - u_m^p}{\tau} - a^2 \frac{u_{m+1}^{p+1} - 2u_m^{p+1} + u_{m-1}^{p+1}}{h^2} = \varphi(mh, p\tau), \\ u_m^0 = \psi(mh), \\ m = 0, \pm 1, \dots; p = 0, 1, \dots, [T/\tau] - 1. \end{cases} \quad (21)$$

Аналогичные выкладки приводят к выражению

$$\lambda(\alpha) = \frac{1}{1 + 4ra^2 \sin^2 \frac{\alpha}{2}}, \quad r = \frac{\tau}{h^2}. \quad (22)$$

Спектр этой задачи заполняет отрезок

$$\left[ 1 + 4ra^2 \sin^2 \frac{\alpha}{2} \right]^{-1} \leq \lambda \leq 1$$

вещественной оси, и условие  $|\lambda| \leq 1$  выполнено при любом  $r$ .

Спектральный признак Неймана применим для исследования разностной задачи Коши и в случае, если пространственных переменных две или более.

Пример 6. Для задачи

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \quad t > 0, \\ u(x, y, t) &= \psi(x, y) \end{aligned} \right\}$$

возьмем сетку  $(x_m, y_n, t_p) = (mh, nh, p\tau)$ . Заменяя производные разностными отношениями, построим разностную схему

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} - \frac{u_{m+1,n}^p - 2u_{mn}^p + u_{m-1,n}^p}{h^2} - \\ - \frac{u_{m,n+1}^p - 2u_{mn}^p + u_{m,n-1}^p}{h^2} = 0, \\ u_{mn}^0 = \psi(mh, nh). \end{cases} \quad (23)$$

Задавая  $u_{mn}^0 = e^{i(\alpha m + \beta n)}$ , т. е. в виде двумерной гармоники, зависящей от двух вещественных параметров  $\alpha$  и  $\beta$ , найдем решение вида

$$u_{mn}^p = \lambda^p(\alpha, \beta) e^{i(\alpha m + \beta n)}.$$

Подставляя это выражение в разностное уравнение, после сокращений и тождественных преобразований найдем

$$\lambda(\alpha, \beta) = 1 - 4r \sin^2 \frac{\alpha}{2} - 4r \sin^2 \frac{\beta}{2}.$$

При изменении вещественных  $\alpha$  и  $\beta$  точка  $\lambda = \lambda(\alpha, \beta)$  пройдет отрезок

$$1 - 8r \leq \lambda \leq 1$$

вещественной оси. Условие устойчивости выполняется, если  $1 - 8r \geq -1$ ,  $r \leq 1/4$ .

Приведем пример, иллюстрирующий применение признака Неймана для разностных уравнений, связывающих значения искомой функции не на двух, а на трех временных слоях.

Пример 7. Задачу Коши для волнового уравнения

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} &= 0, \quad -\infty < x < \infty, \quad 0 < t < T, \\ u(x, 0) &= \psi_1(x), \quad \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), \quad -\infty < x < \infty, \end{aligned}$$

аппроксимируем разностной схемой

$$\left. \begin{aligned} \frac{u_m^{p+1} - 2u_m^p + u_m^{p-1}}{\tau^2} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= 0, \\ p &= 1, 2, \dots, [T/\tau] - 1, \\ u_m^0 = \psi_1(x_m), \quad \frac{u_m^1 - u_m^0}{\tau} &= \psi_2(x_m), \quad m = 0, \pm 1, \dots \end{aligned} \right\} \quad (24)$$

Подставляя в разностное уравнение решение вида (8), получим после простых преобразований следующее уравнение для определения  $\lambda$ :

$$\lambda^2 - 2\left(1 - 2r^2 \sin^2 \frac{\alpha}{2}\right) \lambda + 1 = 0, \quad r = \frac{\tau}{h}.$$

Произведение корней этого уравнения равно единице. Если дискриминант

$$d(\alpha) = 4r^2 \sin^2 \alpha \left(r^2 \sin^2 \frac{\alpha}{2} - 1\right)$$

квадратного уравнения отрицателен, то корни  $\lambda_1(\alpha)$  и  $\lambda_2(\alpha)$  комплексно-сопряженные и равные единице по модулю. В случае  $r < 1$  дискриминант остается отрицательным при всех  $\alpha$ . На рис. 24, а изображен спектр в этом случае. Он заполняет часть единичной окружности. В случае  $r = 1$  спектр заполняет всю окружность. При  $r > 1$  по мере увеличения  $\alpha$  от нуля до  $\pi$  корни  $\lambda_1(\alpha)$  и  $\lambda_2(\alpha)$  движутся из точки  $\lambda = 1$  по единичной окружности один по часовой стрелке, а другой против часовой стрелки, пока не сольются в точке  $\lambda = -1$ , а затем один из корней пойдет по вещественной оси из точки  $\lambda = -1$  влево, а другой вправо, так как они вещественны и  $\lambda_1 \lambda_2 = 1$  (рис. 24, б).

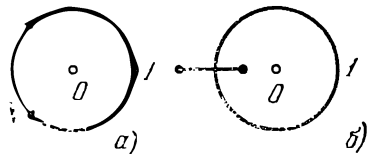


Рис. 24.

Условие устойчивости выполнено при  $r \leq 1$ .

Рассмотрим задачу Коши для следующей гиперболической системы дифференциальных уравнений, описывающей распространение звука:

$$\left. \begin{aligned} \frac{\partial v}{\partial t} &= \frac{\partial w}{\partial x}, \\ \frac{\partial w}{\partial t} &= \frac{\partial v}{\partial x}, \\ v(x, 0) &= \psi_1(x), \quad w(x, 0) = \psi_2(x), \quad -\infty < x < \infty. \end{aligned} \right\} \quad (25)$$

Положим

$$u(x, t) = \begin{pmatrix} v(x, t) \\ w(x, t) \end{pmatrix}, \quad \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix}$$

и запишем (25) в векторной форме:

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - A \frac{\partial u}{\partial x} &= 0, & -\infty < x < \infty, & 0 < t < T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty, \end{aligned} \right\} \quad (25')$$

где

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Исследуем две разностные схемы, аппроксимирующие задачу (25').

Пример 8. Рассмотрим разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - A \frac{u_{m+1}^p - u_m^p}{h} &= 0, & p = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi(x_m), & m = 0, \pm 1, \dots \end{aligned} \right\} \quad (26)$$

Ищем решение векторного однородного разностного уравнения в виде (13):

$$u_m^p = \lambda^p (u^0 e^{iam}) = \lambda^p \begin{pmatrix} v^0 \\ w^0 \end{pmatrix} e^{iam}.$$

Подставляя это выражение в разностное уравнение (26), приходим к равенству

$$\frac{\lambda - 1}{\tau} u^0 - A \frac{e^{ia} - 1}{h} u^0 = 0,$$

или

$$(\lambda - 1) u^0 - r (e^{ia} - 1) A u^0 = 0, \quad r = \frac{\tau}{h}, \quad (27)$$

которое можно рассматривать как векторную запись системы линейных уравнений для определения компонент вектора  $u^0$ .

Запишем систему (27) в развернутой форме:

$$\begin{pmatrix} \lambda - 1 & -r(e^{ia} - 1) \\ -r(e^{ia} - 1) & \lambda - 1 \end{pmatrix} \begin{pmatrix} v^0 \\ w^0 \end{pmatrix} = 0. \quad (28)$$

Система линейных уравнений (28) имеет нетривиальное решение  $u^0 = \begin{pmatrix} v^0 \\ w^0 \end{pmatrix}$  лишь при тех  $\lambda = \lambda(\alpha)$ , при которых определитель системы (28) обращается в нуль:

$$(\lambda - 1)^2 = r^2 (e^{ia} - 1)^2.$$

Отсюда

$$\lambda_1(\alpha) = 1 - r + r e^{ia},$$

$$\lambda_2(\alpha) = 1 + r - r e^{ia}.$$

Корни  $\lambda_1(\alpha)$  и  $\lambda_2(\alpha)$  пробегает окружности радиуса  $r$  с центрами в точках  $1-r$  и  $1+r$  соответственно (рис. 25). Условие устойчивости Неймана не выполнено ни при каком  $r$ .

Пример 9. Рассмотрим разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - A \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \frac{\tau}{2h^2} A^2 (u_{m+1}^p - 2u_m^p + u_{m-1}^p) &= 0, \\ \rho &= 0, 1, \dots, [T/\tau] - 1; \quad m = 0, \pm 1, \dots, \\ u_m^0 &= \psi(x_m), \quad m = 0, \pm 1, \dots, \end{aligned} \right\} \quad (29)$$

аппроксимирующую задачу (25') со вторым порядком и аналогичную схеме (15) для скалярного случая (14). Условие существования нетривиального решения вида (13) у векторного уравнения (25) состоит, как и в примере 8, в том, чтобы обращался в нуль определитель системы, возникающей для определения  $u_0 = \begin{pmatrix} v^0 \\ w^0 \end{pmatrix}$ . Приравняв этот определитель нулю, получим квадратное уравнение относительно  $\lambda = \lambda(\alpha)$ , из которого находим

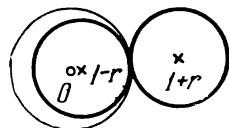


Рис. 25.

$$\left. \begin{aligned} \lambda_1 &= 1 + ir \sin \alpha - 2r^2 \sin^2 \frac{\alpha}{2}, \\ \lambda_2 &= 1 - ir \sin \alpha - 2r^2 \sin^2 \frac{\alpha}{2}. \end{aligned} \right\} \quad (30)$$

Эти формулы аналогичны (16), и, как в (17), получим

$$1 - |\lambda_{1,2}(\alpha)|^2 = 4r^2 \sin^4 \frac{\alpha}{2} (1 - r^2).$$

Спектр, задаваемый формулами (30), лежит в единичном круге при  $r \leq 1$ .

4. Интегральное представление решения\*). Рассмотрим задачу Коши вида

$$\left. \begin{aligned} b_{-1} u_{m-1}^{p+1} + b_0 u_m^{p+1} + b_1 u_{m+1}^{p+1} - \\ - (a_{-1} u_{m-1}^p + a_0 u_m^p + a_1 u_{m+1}^p) &= \tau \varphi_m^p, \\ \rho &= 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi_m, \quad m = 0, \pm 1, \dots, \end{aligned} \right\} \quad (31)$$

с постоянными коэффициентами, предполагая, что

$$b_{-1} e^{-i\alpha} + b_0 + b_1 e^{i\alpha} \neq 0, \quad 0 \leq \alpha \leq 2\pi. \quad (32)$$

\*) Результаты этого пункта в следующих параграфах не используются.

Разностные схемы (1), (15), (18), (21) приводятся к виду (31), если обе части входящих в них разностных уравнений умножить на  $\tau$ .

Заметим прежде всего, что при произвольных ограниченных сеточных функциях  $\{\varphi_m^p\}$  и  $\{\psi_m\}$  задача (31) имеет одно и только одно ограниченное решение. Действительно, если уже известно, что  $\{u_m^p\}$  при данном фиксированном  $p$  существует и ограничено, то уравнение (31) превращается в обыкновенное разностное уравнение второго порядка

$$b_{-1}u_{m-1}^{p+1} + b_0u_m^{p+1} + b_1u_{m+1}^{p+1} = \tau\varphi_m^p + (a_{-1}u_{m-1}^p + a_0u_m^p + a_1u_{m+1}^p) \quad (33)$$

относительно  $\{u_m^{p+1}\}$  с ограниченной правой частью. Соответствующее характеристическое уравнение  $b_{-1} + b_0q + b_1q^2 = 0$  благодаря (32) не имеет корней  $q = e^{i\alpha}$ , по модулю равных единице. Поэтому, как показано в конце п. 2 § 3, оно имеет единственное ограниченное решение  $\{u_m^{p+1}\}$ . Но  $\{u_m^j\} = \{\psi_m\}$  задано и ограничено, поэтому последовательно из (33) однозначно определяются ограниченные функции  $\{u_m^1\}$ ,  $\{u_m^2\}$  и т. д.

Нам понадобятся следующие известные сведения о рядах Фурье.

Каждой последовательности чисел  $c_m$ ,  $m = 0, \pm 1, \dots$ , для которой  $\sum |c_m| < \infty$ , соответствует сходящийся (квадратически в среднем) ряд Фурье

$$C(\alpha) = \frac{1}{\sqrt{2\pi}} \sum c_m e^{-iam}, \quad (34)$$

суммой которого является интегрируемая с квадратом на отрезке  $0 \leq \alpha \leq 2\pi$  функция  $C(\alpha)$ ,  $\int_0^{2\pi} |C(\alpha)|^2 d\alpha < \infty$ .

Обратно, каждая интегрируемая с квадратом на отрезке  $0 \leq \alpha \leq 2\pi$  функция  $C(\alpha)$  разлагается единственным образом в некоторый ряд Фурье (34) с коэффициентами  $c_m$ , вычисляемыми по формулам

$$c_m = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} C(\alpha) e^{iam} d\alpha. \quad (35)$$

При этом выполнено равенство Парсеваля

$$\int_0^{2\pi} |C(\alpha)|^2 d\alpha = \sum |c_m|^2. \quad (36)$$



Теорема 1. Пусть в задаче (31)

$$\max_p \sum_m |\varphi_m^p|^2 < \infty, \quad \sum_m |\psi_m|^2 < \infty.$$

Тогда ограниченное решение этой задачи допускает интегральное представление

$$u_m^p = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} U^p(\alpha) e^{i\alpha m} d\alpha, \quad (37)$$

где интегрируемая с квадратом функция  $U^p(\alpha)$  определяется рекуррентным соотношением

$$U^{p+1}(\alpha) = \lambda(\alpha) U^p(\alpha) + \frac{\tau}{b_{-1}e^{i\alpha} + b_0 + b_1e^{-i\alpha}} \Phi^p(\alpha), \quad p = 0, 1, \dots \quad (38)$$

Здесь

$$\Phi^p(\alpha) = \frac{1}{\sqrt{2\pi}} \sum_m \varphi_m^p e^{-i\alpha m}; \quad U^0(\alpha) \equiv \Psi(\alpha) = \frac{1}{\sqrt{2\pi}} \sum_m \psi_m e^{-i\alpha m},$$

а функция

$$\lambda(\alpha) = \frac{a_1 e^{-i\alpha} + a_0 + a_{-1} e^{i\alpha}}{b_{-1} e^{-i\alpha} + b_0 + b_{-1} e^{i\alpha}}$$

подобрана так, чтобы при каждом  $\alpha$ ,  $0 \leq \alpha \leq 2\pi$ , сеточная функция  $u_m^p = \lambda^p(\alpha) e^{i\alpha m}$  удовлетворяла однородному уравнению, соответствующему уравнению (31).

Доказательство получается прямой подстановкой выражения (37) в уравнение и начальное условие (31) с помощью равенств (34) и (35).

Следствие. Если в (31) функция  $\varphi_m^p \equiv 0$ , то  $\bar{\Phi}^p(\alpha) = 0$ ; в силу (38) имеем  $U^p(\alpha) = \lambda^p(\alpha) \bar{\Psi}(\alpha)$ , а из (37) следует

$$u_m^p = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \lambda^p(\alpha) \bar{\Psi}(\alpha) e^{i\alpha m} d\alpha. \quad (39)$$

Интегральными представлениями (37) и (39) можно пользоваться для анализа свойств разностной схемы (31).

Определим нормы равенствами

$$\left. \begin{aligned} \|u^p\|^2 &= \sum_m |u_m^p|^2; & \|u^{(h)}\|_{U_h} &= \max_p \|u^p\|, \\ \|\varphi^p\|^2 &= \sum_m |\varphi_m^p|^2; & \|\psi\|^2 &= \sum |\psi_m|^2, \\ \|f^{(h)}\|_{F_h} &= \left\| \frac{\varphi_m^p}{\psi_m} \right\|_{F_h} = \|\psi\| + \max_p \|\varphi^p\|. \end{aligned} \right\} \quad (40)$$

Теорема 2. Для устойчивости разностной схемы (31) по начальным данным, т. е. для выполнения неравенств

$$\|u^p\| \leq c \|u^0\|, \quad p = 0, 1, \dots, \left[ \frac{T}{\tau} \right],$$

при произвольном  $u_m^0 = \psi_m$ ,  $\|\psi\| < \infty$ , и при  $\varphi_m^p \equiv 0$  с постоянной  $c$ , не зависящей от  $h$  (и от  $\tau = \tau(h)$ ), необходимо и достаточно, чтобы спектр  $\lambda = \lambda(\alpha)$  лежал в круге (10):

$$|\lambda(\alpha)| \leq 1 + c_1 \tau, \quad (41)$$

где  $c_1$  не зависит от  $h$  (и от  $\tau$ ).

Доказательство. Сначала установим достаточность. При условии (41), очевидно,

$$|\lambda(\alpha)|^p \leq |1 + c_1 \tau|^{T/\tau} \leq e^{c_1 T}. \quad (42)$$

Из представления (39) в силу равенства Парсеваля и неравенства (42) следует

$$\begin{aligned} \|u^p\| &= \left[ \int_0^{2\pi} |\lambda^p(\alpha) U(\alpha)|^2 d\alpha \right]^{1/2} \leq e^{c_1 T} \left[ \int_0^{2\pi} |U(\alpha)|^2 d\alpha \right]^{1/2} = \\ &= e^{c_1 T} \|u^0\| = c \|u^0\|. \end{aligned}$$

Необходимость. Покажем теперь, что из невыполнения (41) при любом фиксированном  $c_1$  следует неустойчивость. Использовать неограниченность при  $\tau \rightarrow 0$  решения

$$u_m^p = \lambda^p(\alpha) e^{i\alpha m}, \quad p = 0, \dots, \left[ \frac{T}{\tau} \right],$$

имеющую место в этом случае, для доказательства неустойчивости при выбранной норме (40) нельзя, так как  $\{e^{i\alpha m}\}$  не принадлежит пространству сеточных функций, у которых сумма квадратов модулей их значений ограничена.

Для доказательства неустойчивости заметим сначала, что всегда можно выбрать интегрируемую с квадратом функцию  $U(\alpha)$  так, чтобы выполнялось неравенство

$$\frac{1}{2\pi} \int_0^{2\pi} |\lambda(\alpha)|^{2p} |U(\alpha)|^2 d\alpha \geq \max_{\alpha} [|\lambda(\alpha)|^{2p} - \varepsilon] \cdot \frac{1}{2\pi} \int_0^{2\pi} |U(\alpha)|^2 d\alpha, \quad (43)$$

где  $\varepsilon > 0$  — произвольное. В самом деле, если  $\max_{\alpha} |\lambda(\alpha)| =$

$= |\lambda(\alpha^*)|$ , то положим

$$U(\alpha) = \begin{cases} 1, & \text{если } \alpha \in [\alpha^* - \delta, \alpha^* + \delta], \\ 0, & \text{если } \alpha \notin [\alpha^* - \delta, \alpha^* + \delta]. \end{cases}$$

Благодаря непрерывности функции  $|\lambda(\alpha)|^{2p}$  при достаточно малом  $\delta = \delta(\varepsilon)$  будет выполнено (43). Если (42) не выполнено, то найдется последовательность  $h_k$  и соответствующая последовательность  $\tau_k = \tau(h_k)$ , при которых

$$c_k \equiv [\max_{\alpha} |\lambda(\alpha, h_k)|]^{[T/\tau_k]} \rightarrow \infty \quad \text{при } k \rightarrow \infty.$$

Положим  $\varepsilon = 1$  и выберем  $U(\alpha)$  так, чтобы выполнялось (43). За последовательность  $\{u_m^j\}$  примем последовательность коэффициентов Фурье функции  $U(\alpha)$ . Тогда (43) при  $p = [T/\tau]$  примет вид

$$\|u^{[T/\tau]}\| \geq (c_k - 1) \|u^0\|, \quad c_k \rightarrow \infty, \quad \text{при } h \rightarrow 0,$$

что и означает отсутствие устойчивости по начальным данным.

**Теорема 3.** *Для устойчивости разностной задачи Коши (31) при сделанном выборе норм (40) необходимо и достаточно, чтобы выполнялся спектральный признак устойчивости (41).*

**Доказательство.** Необходимость очевидна, так как при невыполнении этого признака в силу теоремы 2 нет устойчивости по начальным данным.

Для доказательства достаточности установим, что при каждом  $k \geq 0$  справедливо неравенство

$$\|u^{k+1}\| \leq (1 + c_1\tau) \|u^k\| + c_2\tau \max_n \|\varphi^n\|, \quad (44)$$

из которого, очевидно, следует справедливость при всех  $j$ ,  $p \geq j \geq 0$  неравенств

$$(1 + c_1\tau)^j \|u^{p+1-j}\| \leq (1 + c_1\tau)^{j+1} \|u^{p-j}\| + c_2\tau (1 + c_1\tau)^p \cdot \max_n \|\varphi^n\|.$$

Суммируя левые и правые части этих неравенств по  $j = 0, 1, \dots, p$  почленно и отбрасывая одинаковые слагаемые в левой и правой частях, можно написать

$$\begin{aligned} \|u^{p+1}\| &\leq (1 + c_1\tau)^{p+1} \|u^0\| + c_2\tau p (1 + c_1\tau)^p \max_n \|\varphi^n\| \leq \\ &\leq (1 + c_1\tau) e^{c_1 T} \cdot \|u^0\| + c_2 T e^{c_1 T} \max_n \|\varphi^n\| \leq \text{const} \cdot \|f^{(h)}\|_{Fh}. \end{aligned}$$

Отсюда, ввиду произвольности  $p = 0, 1, \dots, \left[ \frac{T}{\tau} \right] - 1$ , следует устойчивость.

Для доказательства неравенства (44) воспользуемся интегральным представлением решения (37) и рекуррентным соотношением (38), откуда

$$\begin{aligned}
 u_m^{k+1} &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} U^{k+1}(\alpha) e^{i\alpha m} d\alpha = \\
 &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \lambda(\alpha) U^k(\alpha) e^{i\alpha m} d\alpha + \frac{\tau}{\sqrt{2\pi}} \int_0^{2\pi} \frac{\Phi^k(\alpha)}{b_1 e^{-i\alpha} + b_0 + b_{-1} e^{i\alpha}} e^{i\alpha m} d\alpha. \quad (45)
 \end{aligned}$$

Таким образом, сеточная функция  $\{u_m^{k+1}\}$  аргумента  $m$  представлена в виде суммы двух сеточных функций, записанных в виде интегралов по параметру  $\alpha$ .

В силу равенства Парсеваля, для норм этих двух сеточных функций можно написать

$$\begin{aligned}
 \left\| \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \lambda(\alpha) U^k(\alpha) e^{i\alpha m} d\alpha \right\| &= \left[ \int_0^{2\pi} |\lambda(\alpha) U^k(\alpha)|^2 d\alpha \right]^{1/2} \leq \\
 &\leq \max_{\alpha} |\lambda(\alpha)| \left[ \int_0^{2\pi} |U^k(\alpha)|^2 d\alpha \right]^{1/2} \leq (1 + c_1 \tau) \cdot \|u^k\|; \\
 \left\| \frac{\tau}{\sqrt{2\pi}} \int_0^{2\pi} \frac{\Phi^k(\alpha)}{b_1 e^{-i\alpha} + b_0 + b_{-1} e^{i\alpha}} e^{i\alpha m} d\alpha \right\| &= \\
 &= \tau \left[ \int_0^{2\pi} \left| \frac{\Phi^k(\alpha)}{b_1 e^{-i\alpha} + b_0 + b_{-1} e^{i\alpha}} \right|^2 d\alpha \right]^{1/2} \leq \\
 &\leq \frac{\tau}{\min_{\alpha} |b_1 e^{-i\alpha} + b_0 + b_{-1} e^{i\alpha}|} \left[ \int_0^{2\pi} |\Phi^k(\alpha)|^2 d\alpha \right]^{1/2} = \\
 &= \frac{\tau}{\min_{\alpha} |b_1 e^{-i\alpha} + b_0 + b_{-1} e^{i\alpha}|} \cdot \|\Phi^k\| \leq \tau c_2 \max_n \|\Phi^n\|.
 \end{aligned}$$

В силу двух последних оценок для норм слагаемых, входящих в правую часть равенства (45), получаем оценку (44), завершающую доказательство.

Можно показать, что если за норму принять не (40), а равенство  $\|u^p\| = \sup_m |u_m^p|$ , то спектральный признак  $|\lambda(\alpha)| < 1 + c_1\tau$  перестанет быть достаточным признаком устойчивости. Для разностной задачи Коши в случае системы уравнений этот признак также лишь необходимый признак устойчивости.

Интегральным представлением (37) решения разностной задачи Коши можно воспользоваться не только для исследования устойчивости, но и для выяснения других свойств разностной схемы.

Если, например, спектр  $\lambda = \lambda(\alpha)$  при  $\alpha \neq 0$  лежит строго внутри единичного круга, то решения  $u_m^p = \lambda^p(\alpha) e^{iam}$ , отвечающие  $\alpha \neq 0$ , при переходе от слоя к слою гасятся, умножаясь на  $\lambda(\alpha)$ . Из формулы (39) видно, что при  $[T/\tau] = p$  получается сеточная функция, отвечающая функции  $\lambda^p(\alpha)\psi(\alpha)$ , которая сосредоточена на длинных волнах ( $\alpha \approx 0$ ). Разностная схема «выглаживает» начальные данные.

**5. Выглаживание разностного решения как действие аппроксимационной вязкости.** Мы видели, что спектр разностной схемы

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= 0, \quad m = 0, \pm 1, \dots \\ p &= 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi(x_m), \quad m = 0, \pm 1, \dots, \end{aligned} \right\} \quad (46)$$

аппроксимирующей задачу Коши

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= 0, \quad -\infty < x < \infty, \quad 0 < t < T, \\ u(x, 0) &= \psi(x) \quad -\infty < x < \infty, \end{aligned} \right\} \quad (47)$$

есть окружность  $\lambda = 1 - r + re^{i\alpha}$ ,  $0 \leq \alpha < 2\pi$ . При  $r < 1$  каждой точке  $\alpha \neq 0$  соответствует точка спектра  $\lambda(\alpha)$ ,  $|\lambda(\alpha)| < 1$ . Это значит, что каждая гармоника  $u_m^0 = e^{iam}$ , заданная в качестве начальных данных, гасится, умножаясь на  $\lambda(\alpha)$ , при каждом переходе со слоя на слой; решение с течением времени выглаживается, так как при малых  $\alpha h$  (низкочастотные гармоники) погашение слабее. Отметим, что решение дифференциальной задачи (47)  $u(x, t) = \psi(x + t)$  с течением времени не выглаживается — оно получается из начальных данных сдвигом влево. При этом решение задачи (47), отвечающее начальному условию  $u(x, 0) = e^{i\alpha x}$ , есть  $u(x, t) = e^{iat} e^{i\alpha x}$  и множитель  $e^{iat}$  по модулю равен единице. Вычислительный эффект выглаживания решения, имеющий место при использовании разностной схемы (46), можно понимать как проявление *аппроксимационной вязкости*, присущей этой схеме. Объясним, что мы понимаем под аппроксимационной вязкостью. Если уравнение  $\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0$  считать простейшей моделью уравнения движения невязкого газа, то уравнение

$$\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = \mu \frac{\partial^2 u}{\partial x^2} \quad (48)$$

естественно считать моделью уравнений движения вязкого газа с вязкостью, равной  $\mu > 0$ , выглаживающей решение. При начальном условии  $u(x, 0) = e^{i\alpha x}$  решение уравнения (48) имеет вид

$$u(x, t) = e^{-\mu\alpha t + i\alpha t} e^{i\alpha x} \equiv \bar{\lambda}(\alpha, t) e^{i\alpha x}.$$

При  $\mu = O(\tau)$  и  $t = \tau$  гасящий гармонику  $e^{i\alpha x}$  множитель  $\bar{\lambda}(\alpha, t)$  можно записать так:

$$\bar{\lambda}(\alpha, \tau) = 1 - \mu\alpha^2\tau + i\alpha\tau - \frac{\alpha^2\tau^2}{2} + o(\tau^2). \quad (49)$$

Будем предполагать, что решение  $u^{(h)}$  разностной задачи можно доопределить вне сетки так, чтобы полученная при этом гладкая функция  $u^{(h)}(x, t)$  была равномерно по  $h$  ограничена вместе со своими производными до четвертого порядка.

Тогда в точках сетки  $(x, t)$ , пользуясь формулой Тейлора, можно написать

$$\begin{aligned} 0 &= \frac{u^{(h)}(x, t + \tau) - u^{(h)}(x, t)}{\tau} - \frac{u^{(h)}(x + h, t) - u^{(h)}(x, t)}{h} = \\ &= \frac{\partial u^{(h)}(x, t)}{\partial t} - \frac{\partial u^{(h)}(x, t)}{\partial x} + \frac{\tau}{2} \frac{\partial^2 u^{(h)}(x, t)}{\partial t^2} - \\ &\quad - \frac{h}{2} \frac{\partial^2 u^{(h)}(x, t)}{\partial x^2} + h^2 e_1^{(h)}(x, t). \end{aligned} \quad (50)$$

Здесь и далее  $e_1^{(h)}$ ,  $e_2^{(h)}$ ,  $e_3^{(h)}$  — равномерно по  $h$  ограниченные вместе со своими производными функции.

Из равенства (50) следует, в частности,

$$\frac{\partial u^{(h)}}{\partial t} = \frac{\partial u^{(h)}}{\partial x} + h e_2^{(h)}(x, t).$$

Дифференцируя это тождество по  $t$ , получим

$$\frac{\partial^2 u^{(h)}}{\partial t^2} = \frac{\partial}{\partial x} \left( \frac{\partial u^{(h)}}{\partial t} \right) + h \frac{\partial e_2^{(h)}}{\partial t} = \frac{\partial^2 u^{(h)}}{\partial x^2} + h \frac{\partial e_2^{(h)}}{\partial x} + h \frac{\partial e_2^{(h)}}{\partial t} = \frac{\partial^2 u^{(h)}}{\partial x^2} + h e_3^{(h)}.$$

Подставляя выражение для  $\partial^2 u^{(h)}/\partial t^2$  в равенство (50) и отбрасывая члены второго порядка малости, получим дифференциальное уравнение вида (48):

$$\frac{\partial U^{(h)}}{\partial t} - \frac{\partial U^{(h)}}{\partial x} = \frac{h - \tau}{2} \frac{\partial^2 U^{(h)}}{\partial x^2}, \quad (51)$$

которое будем рассматривать не на сетке, а всюду при  $t > 0$ .

Таким образом, разностное уравнение (46) оказалось в «основном совпадающим» с дифференциальным приближением (51), которое есть уравнение вида (48) с малой вязкостью  $\mu = (h - \tau)/2$ . Эта вязкость носит название аппроксимационной, так как появилась в результате аппроксимации задачи (47) разностной задачей (46). Дифференциальное уравнение (51) сглаживает начальные данные в основном так же, как схема (46). Действительно, если  $U(x, 0) = e^{i\alpha x}$ , то к моменту  $t = \tau$  эта гармоника, в соответствии с формулой (49), умножится на

$$\lambda(\alpha, \tau) = 1 - \frac{h - \tau}{2} \alpha^2 \tau + i\alpha\tau - \frac{\alpha^2 \tau^2}{2} + o(\tau^2) = 1 + i\alpha\tau - \frac{h}{2} \alpha^2 \tau + o(\tau^2). \quad (52)$$

При  $u_m^0 = e^{i\alpha x} \Big|_{x=mh} = e^{i\alpha mh}$  по разностной схеме (46) получим в момент  $t = \tau$  ту же гармонику, умноженную на множитель

$$\begin{aligned} \lambda(\alpha) &= 1 - r + r e^{i\alpha h} = 1 - r + r \left( 1 + i\alpha h - \frac{\alpha^2 h^2}{2} \right) + o(h^2) = \\ &= 1 + i\alpha\tau - \frac{h}{2} \alpha^2 \tau + O(\tau^2), \end{aligned}$$

который совпадает с множителем (52) с точностью до бесконечно малых второго относительно  $\tau$  (или  $h$ ) порядка.

### ЗАДАЧИ

1. При каких значениях параметра  $\sigma > 0$  разностная схема, аппроксимирующая задачу Коши для уравнения теплопроводности

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} &= \sigma \frac{u_{m+1}^{p+1} - 2u_m^{p+1} + u_{m-1}^{p+1}}{h^2} + (1 - \sigma) \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2}, \\ u_m^0 &\text{ задано, } m = 0, \pm 1, \dots, \end{aligned} \right\}$$

удовлетворяет спектральному признаку устойчивости Неймана при любом  $r = \tau/h^2$ ?

2. Удовлетворяет ли спектральному признаку устойчивости следующая разностная схема:

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^{p-1}}{2\tau} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= \varphi(x_m, t_p), \quad p \geq 1, \\ u_m^0 &= \psi_m, \\ u_m^1 &= \psi_m^{(1)}, \end{aligned} \right\} m = 0, \pm 1, \dots,$$

где

$$\psi_m^{(1)} = u(x, 0) + \tau \frac{\partial u(x, 0)}{\partial t} = u(x_m, 0) + \tau \frac{\partial^2 u(x, 0)}{\partial x^2} = \psi(x_m) + \tau \psi''(x_m)?$$

Эта разностная схема аппроксимирует задачу Коши (19) для уравнения теплопроводности с порядком  $O(\tau^2 + h^2)$ .

3. Показать, что разностная схема

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} + A \frac{u_{m+1}^{p+1} - u_{m-1}^{p+1}}{2h} &= 0, \quad m = 0, \pm 1, \dots, \\ p &= 0, 1, \dots, \\ u_m^0 &= \psi(x_m), \quad m = 0, \pm 1, \dots, \end{aligned} \right\}$$

аппроксимирующая задачу Коши

$$\begin{aligned} \frac{\partial u}{\partial t} + A \frac{\partial u}{\partial x} &= 0, & -\infty < x < \infty \quad t > 0, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty, \end{aligned}$$

с порядком  $O(\tau + h^2)$ , удовлетворяет спектральному признаку устойчивости при любом  $r = \tau/h$  и любой постоянной  $A$ .

4. Исследовать разностную схему с пересчетом для решения задачи Коши  $u_t + Au_x = 0$ ,  $u(x, 0) = \varphi(x)$ :

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} + A \frac{u_{m+1/2}^{p+1/2} - u_{m-1/2}^{p+1/2}}{h} &= 0, & m = 0, \pm 1, \dots, \\ p &= 0, 1, \dots, \\ u_m^0 &= \varphi_m, & m = 0, \pm 1, \dots \end{aligned} \right\}$$

$A = \text{const}$ , где промежуточная сеточная функция  $u^{p+1/2} = \{u_{m+1/2}^{p+1/2}\}$  определяется по  $u^p = \{u_m^p\}$  в два этапа: сначала вычисляется  $v^p = \{v_m^p\}$  как решение разностной задачи

$$\frac{v_m^{p+1/2} - u_m^p}{\tau/2} + A \frac{v_{m+1}^{p+1/2} - v_{m-1}^{p+1/2}}{2h} = 0, \quad m = 0, \pm 1, \dots,$$

а потом  $u^{p+1/2} = \{u_{m+1/2}^{p+1/2}\}$  по формулам

$$u_{m+1/2}^{p+1/2} = (1 - \alpha) \frac{v_{m+1}^{p+1/2} + v_m^{p+1/2}}{2} + \alpha \frac{v_{m+2}^{p+1/2} + v_{m-1}^{p+1/2}}{2}.$$

Показать, что если параметр интерполяции  $\alpha$  лежит на отрезке  $0 \leq \alpha \leq 0,25$ , то при любом  $r = \tau/h = \text{const}$  выполнен спектральный признак устойчивости. При  $\alpha = 0$  весь спектр лежит на единичной окружности, а при  $0 < \alpha \leq 0,25$  он располагается внутри единичного круга и касается этого круга лишь при  $\lambda = 1$ . Собственному значению  $\lambda = 1$  отвечают собственные функции  $u_m = (\pm 1)^m$ .

## § 26. Принцип замороженных коэффициентов

Здесь мы изложим прием, весьма расширяющий класс нестационарных разностных задач, для исследования которых можно пользоваться спектральным признаком устойчивости. Этот необходимый признак устойчивости, изложенный в § 25 для исследования разностной задачи Коши с постоянными коэффициентами, можно применять и в случае разностной задачи Коши с «непрерывными», но не постоянными коэффициентами, а также для задач в ограниченных областях, когда граничные условия задаются не только при  $t = 0$ , но и на боковых границах. Этим приемом можно пользоваться и для исследования нелинейных задач.

**1. Замораживание коэффициентов во внутренних точках.** Сформулируем принцип замороженных коэффициентов, поль-



зуюсь в качестве примера следующей разностной краевой задачей:

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - a(x_m, t_p) \frac{u_{m-1}^p - 2u_m^p + u_{m+1}^p}{h^2} &= 0, \\ \rho &= 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi_m, \quad m = 0, 1, \dots, M; \quad Mh = 1, \\ l_1 u_0^{p+1} &= 0, \quad l_2 u_M^{p+1} = 0. \end{aligned} \right\} \quad (1)$$

В этой разностной краевой задаче  $l_1 u_0^{p+1} = 0$  и  $l_2 u_M^{p+1} = 0$  — некоторые условия, задаваемые соответственно на левой и правой границах сеточного отрезка  $0 \leq m \leq M$ ;  $a(x, t) > 0$ .

Выберем произвольную внутреннюю точку  $(\bar{x}, \bar{t})$  области  $0 \leq x \leq 1$ ,  $0 \leq t \leq T$ , где рассматривается задача (1), и «заморозим» коэффициенты задачи (1) в этой точке.

Возникающее разностное уравнение с постоянными коэффициентами

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - a(\bar{x}, \bar{t}) \frac{u_{m-1}^p - 2u_m^p + u_{m+1}^p}{h^2} &= 0, \\ \rho &= 0, 1, \dots, [T/\tau] - 1; \quad m = 0, \pm 1, \dots, \end{aligned} \right\} \quad (2)$$

будем рассматривать теперь не при  $0 < m < M$ , а при всех целочисленных  $m$ . Сформулируем теперь

*Принцип замороженных коэффициентов. Для устойчивости задачи (1) необходимо, чтобы задача Коши для разностного уравнения с постоянными коэффициентами (2) удовлетворяла необходимому спектральному признаку устойчивости Неймана.*

В обоснование принципа замороженных коэффициентов приведем следующее рассуждение на эвристическом уровне строгости.

При измельчении сетки коэффициент  $a(x, t)$  в окрестности точки  $(\bar{x}, \bar{t})$  за любое фиксированное число шагов сетки длины  $h$  по пространству и длины  $\tau$  по времени ввиду непрерывности функции  $a(x, t)$  меняется все меньше и все меньше отличается от значения  $a(\bar{x}, \bar{t})$ . Добавим к этому, что расстояние от точки  $(\bar{x}, \bar{t})$  до границ  $x = 0$  и  $x = 1$  отрезка, измеренное числом шагов сетки, стремится к бесконечности. Поэтому при мелкой сетке возмущения, наложенные на решение задачи (1) в момент времени  $t = \bar{t}$ , в окрестности точки  $x = \bar{x}$  развиваются (за малое время) примерно так же, как для задачи (2).

Понятно, что это рассуждение носит общий характер. Оно не зависит от числа пространственных переменных, числа

искомых функций и вида разностного уравнения или системы уравнений.

В § 25 мы рассматривали задачу Коши для уравнения вида (2) и нашли, что для выполнения условия Неймана отношение  $r = \tau/h^2$  шагов сетки должно удовлетворять условию

$$r \leq \frac{1}{2a(\bar{x}, \bar{t})}.$$

Поскольку в силу принципа замороженных коэффициентов для устойчивости задачи (1) это условие должно выполняться при любых  $(\bar{x}, \bar{t})$ , отношение  $r = \tau/h$  шагов сетки должно быть подчинено условию

$$r \leq \frac{1}{2 \max_{\bar{x}, \bar{t}} a(\bar{x}, \bar{t})}. \quad (3)$$

Принцип замороженных коэффициентов позволяет ориентироваться на эвристическом уровне строгости и при исследовании устойчивости нелинейных задач. Поясним это на следующей нелинейной задаче:

$$\begin{aligned} u_t - (1 + u^2) u_{xx} &= 0, & 0 < x < 1, \\ u(x, 0) &= \psi_0(x), & 0 < x < 1, \\ u(0, t) &= \psi_1(t), & u(1, t) = \psi_2(t), & 0 < t < T. \end{aligned}$$

Используем следующую разностную схему:

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{p+1} - u_m^p}{\tau_p} - [1 + (u_m^p)^2] \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0, \\ 0 < m < M, & p = 0, 1, \dots, [T/\tau] - 1, \\ u^0 = \psi_0(mh), & 0 \leq m \leq M, \\ u_0^p = \psi_1(t_p + \tau_p), & p = 1, 2, \dots, [T/\tau], \\ u_M^p = \psi_2(t_p + \tau_p). \end{cases}$$

В ней допускается изменение шага  $\tau_p$  от слоя к слою. Эта схема позволяет последовательно, слой за слоем, вычислить  $u_m^1$ ,  $m = 0, \dots, M$ , затем  $u_m^2$ ,  $m = 0, 1, \dots, M$ , и т. д.

Допустим, что мы уже добрались до слоя  $t = t_p$ , вычислили  $u_m^p$ ,  $m = 0, 1, \dots, M$ , и хотим продолжать счет.

Как выбрать следующий шаг  $\tau = \tau_p$ ? Можно принять, что нам предстоит сосчитать решение линейного разностного уравнения

$$\frac{u_m^{p+1} - u_m^p}{\tau_p} - a(x_m, t_p) \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0$$

с заданным переменным коэффициентом  $a(x_m, t_p) \equiv 1 + (u_m^p)^2$ . Действительно, естественно считать, что значения  $u_m^p$  близки к значениям  $u(x, t)$  гладкого решения  $u(x, t)$  дифференциальной задачи. Коэффициент тогда близок к непрерывной функции  $a(x, t) \equiv 1 + u^2(x, t)$ , которая на протяжении нескольких временных шагов мало изменяется.

Применение признака Неймана к уравнению с переменным коэффициентом  $a(x_m, t_p)$  дает ограничение (3) на соотношение шагов сетки, необходимое для устойчивости:

$$\frac{\tau_p}{h^2} = r_p \leq \frac{1}{2 \max_x |a(x, t_p)|} = \frac{1}{2 \max_m |1 + (u_m^p)^2|}.$$

Отсюда следует рекомендация выбрать очередной шаг  $\tau_p$  из условия

$$\tau_p \leq \frac{1}{2 \max_m |1 + (u_m^p)^2|} h^2.$$

Численный эксперимент на машине подтверждает правильность этих эвристических рассуждений.

Если необходимое условие устойчивости, полученное путем рассмотрения задачи Коши с замороженными в произвольной точке области коэффициентами, окажется нарушенным, то устойчивости нельзя ожидать ни при каком задании граничных условий. Подчеркнем, однако, что принцип замороженных коэффициентов не учитывает влияния граничных условий. В случае выполнения необходимого условия устойчивости, вытекающего из принципа замороженных коэффициентов, устойчивость может иметь место при одних, и не иметь места при других граничных условиях. Теперь изложим признак К. И. Бабенко и И. М. Гельфанда, учитывающий влияние границ в случае задачи на отрезке.

**2. Признак Бабенко и Гельфанда.** При рассмотрении задачи (1) мы полагали, что возмущения, сообщенные решению задачи (1) в окрестности произвольной внутренней точки  $(\bar{x}, \bar{t})$ , при мелкой сетке развиваются примерно так же, как такие же возмущения, сообщенные решению задачи Коши (2) с замороженными в точке  $(\bar{x}, \bar{t})$  коэффициентами. В обоснование этого принципа мы принимали во внимание, что расстояния от внутренней точки  $(\bar{x}, \bar{t})$  до границ, измеренные числом шагов сетки, при измельчении сетки неограниченно возрастают. Но если точка  $(\bar{x}, \bar{t})$  лежит на боковой границе  $\bar{x} = 0$  или  $\bar{x} = 1$ , то это эвристическое рассуждение теряет убедительность. Пусть, например,  $\bar{x} = 0$ . Тогда расстояние от точки  $\bar{x}$  до любой фиксированной точки  $x > 0$  (в частности, до правого конца отрезка  $x = 1$ ),

измеренное числом шагов сетки, при  $h \rightarrow 0$  по-прежнему неограниченно возрастает, но число шагов до левого конца  $x = 0$  не меняется и остается равным нулю.

Поэтому возмущение решения задачи (1) вблизи левой границы  $x = 0$  за малое время должно развиваться подобно возмущению решения задачи

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - a(0, \tilde{t}) \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0, \quad m = 1, 2, \dots, \\ l_1 u_0^{p+1} = 0. \end{aligned} \right\} \quad (4)$$

Эта задача получилась из исходной задачи (1) при замораживании коэффициента  $a(x, \tilde{t})$  в левом конце отрезка  $\tilde{x} = 0$  и одновременном удалении правой границы в  $+\infty$ . Задачу (4) естественно рассматривать только на тех функциях  $u^p = \{u_0^p, u_1^p, u_2^p, \dots\}$ , для которых

$$u_m^p \rightarrow 0 \quad \text{при} \quad m \rightarrow +\infty.$$

Только в этом случае возмущение сосредоточено вблизи границы  $x = 0$ , и только относительно возмущений такого вида задача (1) и задача (4) вблизи левой границы  $x = 0$  сходны между собой.

Точно так же развитие возмущений решения задачи (1) вблизи правой границы  $x = 1$  должно быть похоже на развитие таких же возмущений для задачи

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - a(1, \tilde{t}) \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0, \\ m = \dots, -2, -1, 0, 1, \dots, M-1, \\ l_2 u_M^{p+1} = 0 \end{aligned} \right\} \quad (5)$$

с одной только правой границей. Эта задача возникла из исходной задачи (1) при замораживании коэффициента  $a(x, \tilde{t})$  в правом конце  $\tilde{x} = 1$  и при удалении левой границы в  $-\infty$ . Задачу (5) надо рассматривать на сеточных функциях  $u^p = \{\dots, u_{-2}^p, u_{-1}^p, u_0^p, u_1^p, \dots, u_M^p\}$ , удовлетворяющих условию  $u_m^p \rightarrow 0$  при  $m \rightarrow -\infty$ .

Задачи (2), (4) и (5) проще исходной задачи (1) в том отношении, что при фиксированном  $r = \tau/h^2$  они не зависят от  $h$  и являются задачами с постоянными коэффициентами.

Таким образом, процедура исследования устойчивости, учитывающая влияние границ, применительно к задаче (1) состоит в следующем. Надо составить три вспомогательные задачи (2), (4) и (5). Для каждой из этих трех задач, не зависящих от  $h$ , надо найти все те числа  $\lambda$  (собственные числа

оператора перехода от  $u^p$  к  $u^{p+1}$ ), при которых существуют решения вида

$$u_m^p = \lambda^p \bar{u}_m^0.$$

При этом в случае задачи (2)  $u^0 = \{u_m^0\}$ ,  $m = 0, \pm 1, \dots$ , должно быть ограничено. В случае задачи (4)  $u^0 = \{u_0^0, u_1^0, \dots, u_m^0, \dots\}$ ,  $u_m^0 \rightarrow 0$  при  $m \rightarrow +\infty$ , а в случае задачи (5)

$$u^0 = \{\dots, u_{-2}^0, u_{-1}^0, u_0^0, u_1^0, \dots, u_M^0\},$$

$$u_m^0 \rightarrow 0 \text{ при } m \rightarrow -\infty.$$

Для устойчивости задачи (1) совокупность собственных чисел каждой из трех задач (2), (4) и (5) должна лежать в единичном круге  $|\lambda| \leq 1$ . (Задача (2) рассматривается при каждом фиксированном  $\bar{x}$ ,  $0 < \bar{x} < 1$ .)

Продолжим рассмотрение задачи (1). Будем считать в дальнейшем, что  $a(x, t) \equiv 1$ , и вычислим спектры для всех трех задач (2), (4) и (5) при различных краевых условиях  $l_1 u_0^{p+1} = 0$  и  $l_2 u_M^{p+1} = 0$ .

Подставляя решение  $u_m^p = \lambda^p u_m$  в разностное уравнение (2), получаем

$$(\lambda - 1)u_m - r(u_{m+1} - 2u_m + u_{m-1}) = 0, \quad r = \frac{\tau}{h^2},$$

или

$$u_{m+1} - \left( \frac{-2r + 1 - \lambda}{r} \right) u_m + u_{m-1} = 0. \quad (6)$$

Это — уравнение второго порядка. Подобными уравнениями мы занимались в гл. 1. Чтобы написать общее решение уравнения (6), составим характеристическое уравнение

$$q^2 + \left( 2 + \frac{\lambda - 1}{r} \right) q + 1 = 0. \quad (7)$$

Если  $q$  — корень этого уравнения, то сеточная функция

$$u_m^p = \lambda^p q^m$$

есть одно из решений уравнения

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0.$$

Если  $|q| = 1$ , т. е.  $q = e^{i\alpha}$ , то ограниченная при  $m \rightarrow \pm\infty$  и при  $m \rightarrow -\infty$  сеточная функция

$$u_m^p = \lambda^p e^{iam},$$

как мы видели в § 25, является решением при

$$\lambda = 1 - 4r \sin^2 \frac{\alpha}{2}, \quad 0 \leq \alpha \leq 2\pi.$$

Эти  $\lambda = \lambda(\alpha)$  заполняют отрезок  $1 - 4r \leq \lambda \leq 1$  на вещественной оси. Этот отрезок и есть спектр задачи (2). Собственных значений  $\lambda$ , не лежащих на этом отрезке, задача (2) не имеет, так как в случае отсутствия у характеристического уравнения (7) корня  $q$ , по модулю равного единице, задача (6) не имеет ограниченного при  $m \rightarrow \pm\infty$  решения.

Если  $\lambda$  не лежит на отрезке  $1 - 4r \leq \lambda \leq 1$ , то оба корня характеристического уравнения (7) отличны по модулю от единицы, но их произведение равно свободному члену квадратного уравнения (7), т. е. единице. Поэтому среди корней уравнения (7) один по модулю больше, а другой меньше единицы. Пусть для определенности  $|q_1| < 1$ , а  $|q_2| > 1$ . Тогда общее решение уравнения (6), убывающее по модулю при  $m \rightarrow +\infty$ , имеет вид

$$u_m = c [q_1(\lambda)]^m,$$

а общее решение уравнения (6), стремящееся к нулю при  $m \rightarrow -\infty$ , имеет вид

$$u_m = c [q_2(\lambda)]^m.$$

Для определения собственных значений задачи (4) надо подставить  $u_m = c q_1^m(\lambda)$  в левое граничное условие  $l_1 u = 0$  и найти все те  $\lambda$ , при которых оно выполняется. Это и будут все собственные значения задачи (4). Если, например,

$$l_1 u_0 \equiv u_0 = 0,$$

то условие  $c q_1^0 = 0$  не выполняется ни при каком  $c \neq 0$ , так что собственных значений нет.

Если  $l_1 u_1 = u - u_0 = 0$ , то условие  $c q_1^1 - c q_1^0 = c(q_1 - 1) = 0$  ввиду  $q_1 \neq 1$  приводит к  $c = 0$ , так что собственных значений опять нет.

Если  $l_1 u = 2u - u_0 = 0$ , то условие  $c(2q_1 - 1) = 0$  выполняется при  $c \neq 0$ , если  $q_1 = 1/2$ .

Из уравнения (7) находим, что в случае  $q_1 = 1/2$  число  $\lambda$  есть

$$\lambda = 1 + r \left( q_1 - 2 + \frac{1}{q_1} \right) = 1 + r \frac{1 - 4 + 4}{2} = 1 + \frac{r}{2}.$$

Это и есть единственное собственное значение задачи (4). Оно лежит вне единичного круга, так как  $\lambda = 1 + r/2 > 1$ . Аналогично вычисляются собственные значения задачи (5). Они получаются из уравнения

$$l_2 u_M = 0,$$

при

$$u_m = q_2^m, \quad q_2 = q_2(\lambda), \quad m = M, M-1, M-2, \dots$$

Рассмотрим в качестве еще одного примера разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= 0, & p &= 0, 1, \dots, [T/\tau] - 1, \\ & & m &= 0, 1, \dots, M-1, \\ & & Mh &= 1, \\ u_m^0 &= \psi(x_m), & m &= 0, 1, \dots, M, \\ u_M^{p+1} &= 0, \end{aligned} \right\} (8)$$

аппроксимирующую задачу

$$\left. \begin{aligned} u_t - u_x &= 0, & 0 < x < 1, & 0 < t < T, \\ u(x, 0) &= \psi(x), \\ u(1, t) &= 0. \end{aligned} \right\}$$

Применим для исследования ее устойчивости признак Бабенко — Гельфанда. Сопоставим схеме (8) три задачи: задачу без боковых границ

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \quad m = 0, \pm 1, \dots, \quad (9)$$

задачу с одной только левой боковой границей

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \quad m = 0, 1, \dots, \quad (10)$$

и задачу с одной только правой боковой границей

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= 0, & m &= M-1, M-2, \dots, \\ u_M^{p+1} &= 0, & p &= 0, 1, \dots, [T/\tau] - 1. \end{aligned} \right\} (11)$$

В случае задачи (10) с одной только левой боковой границей граничного условия нет, так как его не было в исходной задаче (8).

Надо найти совокупность собственных чисел всех трех операторов перехода от  $u^p$  к  $u^{p+1}$ , соответствующих каждой из трех вспомогательных задач (9), (10), (11), и выяснить, при каких условиях все они лежат в круге  $|\lambda| \leq 1$ .

Решение вида

$$u_m^p = \lambda^p u_m^0$$

при подстановке в разностное уравнение

$$u_m^{p+1} = (1-r)u_m^p + ru_{m+1}^p, \quad r = \frac{\tau}{h},$$

приводит к следующему обыкновенному разностному уравнению первого порядка для собственной функции:

$$(\lambda - 1 + r)u_m - ru_{m+1} = 0. \quad (12)$$

Соответствующее характеристическое уравнение

$$(\lambda - 1 + r) - rq = 0 \quad (13)$$

дает связь между  $\lambda$  и  $q$ . Общее решение уравнения (12) есть

$$u_m = cq^m = c \left( \frac{\lambda - 1 + r}{r} \right)^m, \quad m = 0, \pm 1, \dots$$

При  $|q| = 1$ ,  $q = e^{i\alpha}$ ,  $0 \leq \alpha \leq 2\pi$

$$\lambda = (1-r) + re^{i\alpha}.$$

Точка  $\lambda = \lambda(\alpha)$  пробегает окружность с центром в точке  $1-r$  и радиусом  $r$ . Это и есть собственные значения задачи (9).

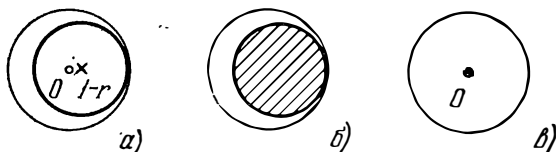


Рис. 26.

(рис. 26, а). Убывающее при  $m \rightarrow +\infty$  нетривиальное решение

$$u_m^p = \lambda^p u_m = c_0 \lambda^p q^m$$

задачи (10) существует при любом  $q$ ,  $|q| < 1$ .

Соответствующие  $\lambda = 1 - r + rq$  заполняют, очевидно, всю внутренность круга, ограниченного окружностью  $\lambda = (1-r) + re^{i\alpha}$  (рис. 26, б).

Наконец, решения задачи (11)  $u_m^p = \lambda^p u_m$ , убывающие при  $m \rightarrow -\infty$ , должны иметь вид

$$u_m^p = c \lambda^p q^m, \quad |q| > 1,$$

где  $\lambda$  и  $q$  связаны равенством (13).

Из граничного условия  $u_M^p = 0$  следует, что нетривиальное решение ( $c \neq 0$ ) существует только при  $\lambda = \lambda(q) = 0$ , т. е.



при  $q = (r-1)/r$ . Эта величина  $q$  по модулю больше единицы в случае выполнения одного из неравенств  $(r-1)/r > 1$  или  $(r-1)/r < -1$ . Первое неравенство решений не имеет. Решение второго:  $r < 1/2$ .

Итак, при  $r < 1/2$  задача (10) имеет собственное значение  $\lambda = 0$  (рис. 26, в). На рис. 27, а, б, в изображены объединения собственных значений всех трех задач соответственно для случаев  $r < 1/2$ ,  $1/2 < r < 1$  и  $r > 1$ .

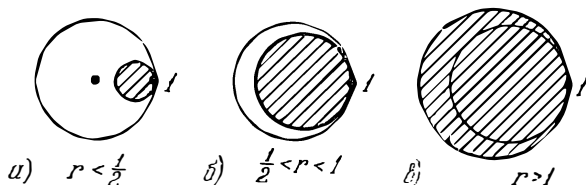


Рис. 27.

Ясно, что объединение собственных значений всех трех задач лежит в круге  $|\lambda| < 1 + c\tau$ , где  $c$  не зависит от  $h$ , в том и только том случае, если  $r \leq 1$ .

Изложенный здесь признак устойчивости нестационарных разностных задач на отрезке, учитывающий влияние граничных условий, применим и в случае краевых задач на отрезке для систем разностных уравнений. В этом случае естественные на первый взгляд схемы, удовлетворяющие признаку Неймана, часто оказываются неустойчивыми из-за неудачной аппроксимации граничных условий, и важно уметь подбирать схемы, свободные от этого недостатка.

В гл. 14 мы еще вернемся к обсуждению спектрального признака Бабенко — Гельфанда с некоторой более общей точки зрения. В частности, будет строго доказано, что его выполнение необходимо для устойчивости и что при его выполнении устойчивость не может «грубо» нарушаться.

### ЗАДАЧИ

1. Выяснить условия выполнения спектрального признака устойчивости для разностной схемы

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \frac{\tau}{2} \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0,$$

$$m = 1, 2, \dots, M-1,$$

$$u_m^0 = \psi(x_m), \quad m = 0, 1, \dots, M,$$

$$\left. \begin{aligned} u_M^{p+1} &= 0, \\ \frac{u_0^{p+1} - u_0^p}{\tau} - \frac{u_1^p - u_0^p}{h} &= 0, \end{aligned} \right\} p = 0, 1, \dots, [T/\tau] - 1,$$

аппроксимирующей дифференциальную задачу

$$\left. \begin{aligned} u_t - u_x &= 0, & 0 < x < 1, & 0 < t < T, \\ u(x, 0) &= \psi(x), \\ u(0, t) &= u(1, t) = 0 \end{aligned} \right\}$$

на гладком решении  $u(x, t)$  со вторым порядком относительно  $h$ .

Ответ:  $\tau/h \leq 1$ .

2. Для построения разностной схемы, аппроксимирующей следующую краевую задачу для гиперболической системы дифференциальных уравнений

$$\left. \begin{aligned} \frac{\partial v}{\partial t} &= \frac{\partial w}{\partial x}, \\ \frac{\partial w}{\partial t} &= \frac{\partial v}{\partial x}, & 0 \leq x \leq 1, & 0 \leq t \leq T, \\ v(x, 0) &= \psi_1(x), & w(x, 0) &= \psi_2(x), \\ v(0, t) &= w(1, t) = 0, \end{aligned} \right\}$$

положим  $u(x, t) = \begin{pmatrix} v(x, t) \\ w(x, t) \end{pmatrix}$  и запишем ее в матричной форме:

$$\left. \begin{aligned} \frac{\partial}{\partial t} u - A \frac{\partial}{\partial x} u &= 0, \\ u(x, 0) &= \psi(x), \\ v(0, t) &= w(1, t) = 0, \end{aligned} \right\}$$

где  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Выберем сетку  $(x_m, t_n) = (mh, n\tau)$ ,  $h = 1/M$ ,  $M$  — натуральное. Положим

$$\begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - A \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \frac{\tau}{2} A^2 \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= 0, \\ u_m^0 &= \psi(x_m), \\ v_0^{p+1} = w_M^{p+1} &= 0, \end{aligned} \quad m = 1, 2, \dots, M-1,$$

Для завершения построения схемы надо задать дополнительные граничные условия на левой и правой боковых границах. Заметив, что при любых  $\alpha$  и  $\beta$  из системы дифференциальных уравнений следуют равенства

$$\left. \begin{aligned} \frac{\partial(v + \alpha w)}{\partial t} - \frac{\partial(w + \alpha v)}{\partial x} \Big|_{x=0} &= 0, \\ \frac{\partial(v + \beta w)}{\partial t} - \frac{\partial(w + \beta v)}{\partial x} \Big|_{x=1} &= 0, \end{aligned} \right\}$$

зададим дополнительные разностные краевые условия, положив

$$\begin{aligned} \frac{(v_0^{p+1} + \alpha w_0^{p+1}) - (v_0^p + \alpha w_0^p)}{\tau} - \frac{(w_1^p + \alpha v_1^p) - (w_0^p + \alpha v_0^p)}{h} &= 0, \\ \frac{(v_M^{p+1} + \beta w_M^{p+1}) - (v_M^p + \beta w_M^p)}{\tau} - \frac{(w_M^p + \beta v_M^p) - (w_{M-1}^p + \beta v_{M-1}^p)}{h} &= 0. \end{aligned}$$

При условии  $r = \tau/h \leq 1$  показать, что:

- а) Если  $\alpha = 1$ ,  $\beta = -1$ , то спектральный признак устойчивости выполнен;  
 б) Если  $\alpha = -1$ , то независимо от выбора числа  $\beta$  спектральный признак устойчивости не выполнен;  
 в) Найти условия, которым должны удовлетворять  $\alpha$  и  $\beta$ , чтобы выполнялся спектральный признак устойчивости, учитывающий влияние граничных условий.

## § 27. Представление решений некоторых модельных задач в виде конечных рядов Фурье

Приведем примеры модельных задач, решения которых удается представить в виде конечных рядов Фурье. Такие представления имеют большую ценность, так как позволяют понять свойства рассматриваемых модельных задач, а тем самым и того класса задач, который они моделируют.

Сначала необходимо объяснить, что такое ряды Фурье для сеточных функций.

**1. Ряды Фурье для сеточных функций.** Рассмотрим множество всех вещественных функций  $v = \{v_m\}$ , определенных в точках  $x_m = mh$ ,  $m = 0, 1, \dots, M$ ;  $Mh = 1$ , обращающихся в нуль при  $m=0$  и  $m=M$ . Совокупность этих функций с обычными операциями сложения и умножения их на вещественные числа образует линейное пространство. Размерность этого пространства есть  $M - 1$ , поскольку система функций

$$\tilde{\psi}_m^{(k)} = \begin{cases} 0, & \text{если } m \neq k, \\ 1, & \text{если } m = k, \end{cases}$$

$k = 1, 2, \dots, M - 1$ , очевидно, образует базис. Действительно, каждую функцию  $v = (v_0, v_1, \dots, v_M)$ ,  $v_0 = v_M = 0$ , можно единственным образом представить в виде линейной комбинации функций  $\tilde{\psi}^{(1)}$ ,  $\tilde{\psi}^{(2)}$ , ...,  $\tilde{\psi}^{(M-1)}$ :

$$v = v_1 \tilde{\psi}^{(1)} + \dots + v_{M-1} \tilde{\psi}^{(M-1)}.$$

Введем в рассматриваемом пространстве скалярное умножение, положив

$$(v, w) = h \sum_{m=0}^M v_m w_m. \quad (1)$$

Покажем, что система функций

$$\psi^{(k)} = \left\{ \sqrt{2} \sin \frac{k\pi m}{M} \right\}, \quad k = 1, 2, \dots, M - 1, \quad (2)$$

образует ортонормальный базис в рассматриваемом пространстве, т. е. что

$$(\psi^{(k)}, \psi^{(r)}) = \begin{cases} 0, & k \neq r, \\ 1, & k = r, \end{cases} \\ k, r = 1, 2, \dots, M-1. \quad (3)$$

Для доказательства заметим, что

$$\sum_{m=0}^{M-1} \cos \frac{l\pi m}{M} = \frac{1}{2} \sum_{m=0}^{M-1} \left( e^{i \frac{l\pi m}{M}} + e^{-i \frac{l\pi m}{M}} \right) = \\ = \frac{1}{2} \frac{1 - e^{il\pi}}{1 - e^{i \frac{l\pi}{M}}} + \frac{1}{2} \frac{1 - e^{-il\pi}}{1 - e^{-i \frac{l\pi}{M}}} = \begin{cases} 0, & \text{если } l \text{ четно и } 0 < l < 2M, \\ 1, & \text{если } l \text{ нечетно.} \end{cases}$$

Отсюда при  $k \neq r$  получаем

$$(\psi^{(k)}, \psi^{(r)}) = 2h \sum_{m=0}^M \sin \frac{k\pi m}{M} \sin \frac{r\pi m}{M} = 2h \sum_{m=0}^{M-1} \sin \frac{k\pi m}{M} \sin \frac{r\pi m}{M} = \\ = h \sum_{m=0}^{M-1} \cos \frac{(k-r)\pi m}{M} - h \sum_{m=0}^{M-1} \cos \frac{(k+r)\pi m}{M} = 0,$$

а при  $k = r$

$$(\psi^{(k)}, \psi^{(k)}) = h \sum_{m=0}^{M-1} \cos 0 - h \sum_{m=0}^{M-1} \cos \frac{2k\pi m}{M} = hM - h \cdot 0 = 1.$$

Любая сеточная функция  $v = (v_0, v_1, \dots, v_M)$  разлагается по ортонормальному базису (2) в сумму

$$v = c_1 \psi^{(1)} + \dots + c_{M-1} \psi^{(M-1)},$$

или

$$v_m = \sqrt{2} \sum_{k=1}^{M-1} c_k \sin \frac{k\pi m}{M}, \quad (4)$$

где

$$c_k = (v, \psi^{(k)}) = \sqrt{2} h \sum_{m=0}^M v_m \sin \frac{k\pi m}{M}.$$

Ясно, что благодаря ортонормальности базиса (2) имеем

$$(v, v) = c_1^2 + c_2^2 + \dots + c_{M-1}^2. \quad (5)$$

Сумма (4) и есть разложение сеточной функции  $v = \{v_m\}$  в конечный ряд Фурье, а равенство (5) — точный аналог равенства Парсеваля в обычной теории рядов Фурье.

Совершенно аналогично можно рассмотреть конечные ряды Фурье для функций на сеточном квадрате. Рассмотрим сетку

$$x_m = mh, \quad y_n = nh, \quad 0 \leq mh \leq 1, \quad 0 \leq nh \leq 1,$$

причем  $h = 1/M$ ,  $M$  — натуральное. Совокупность вещественных функций  $v = \{v_{mn}\}$ , определенных в точках сетки и обращающихся в нуль в точках, лежащих на границе квадрата, образует линейное пространство. Введем в нем скалярное умножение

$$(v, w) = h^2 \sum_{n=0}^M \sum_{m=0}^M v_{mn} w_{mn}.$$

В рассматриваемом линейном пространстве размерности  $(M-1)^2$  система функций

$$\psi^{(k, l)} = 2 \sin \frac{k\pi m}{M} \sin \frac{l\pi n}{M}, \quad k = 1, 2, \dots, M-1, \\ l = 1, 2, \dots, M-1,$$

образует ортонормальный базис

$$(\psi^{(k, l)}, \psi^{(r, s)}) = \begin{cases} 0, & \text{если } k \neq r \text{ либо } l \neq s, \\ 1, & \text{если } k = r \text{ и } l = s. \end{cases}$$

Это следует из (3), если заметить, что

$$(\psi^{(k, l)}, \psi^{(r, s)}) = \left( 2 \sum_{m=0}^M \sin \frac{k\pi m}{M} \sin \frac{r\pi m}{M} \right) \left( 2 \sum_{n=0}^M \sin \frac{l\pi n}{M} \sin \frac{s\pi n}{M} \right) = \\ = (\psi^{(k)}, \psi^{(r)}) (\psi^{(l)}, \psi^{(s)}).$$

Любая функция  $v = \{v_{mn}\}$ , обращающаяся в нуль на границе квадрата, разлагается в конечный двумерный ряд Фурье

$$v_{mn} = 2 \sum_{k, l=1}^{M-1} c_{kl} \sin \frac{k\pi m}{M} \sin \frac{l\pi n}{M}, \quad (6)$$

где

$$c_{kl} = (v, \psi^{(k, l)}).$$

Справедливо равенство Парсеваля

$$(v, v) = \sum_{k, l=1}^{M-1} c_{kl}^2. \quad (7)$$

Во всех примерах разностных краевых задач, решения которых мы запишем с помощью конечных рядов Фурье, используется выражение

$$\Delta_{xx} v_m \equiv \frac{1}{h^2} (v_{m+1} - 2v_m + v_{m-1}), \quad m = 1, 2, \dots, M-1. \quad (8)$$

Заметим, что

$$\Lambda_{xx} \sin \frac{k\pi m}{M} = \frac{1}{h^2} \left[ \sin \frac{k\pi(m+1)}{M} - 2 \sin \frac{k\pi m}{M} + \sin \frac{k\pi(m-1)}{M} \right] = \\ = \frac{2}{h^2} \left( \cos \frac{k\pi}{M} - 1 \right) \sin \frac{k\pi m}{M} = \mu_k \sin \frac{k\pi m}{M}, \quad m = 1, 2, \dots, M-1,$$

$$\text{где } \mu_k = -\frac{4}{h^2} \sin^2 \frac{k\pi}{2M}.$$

Другими словами, базис (2) состоит из собственных функций оператора  $\Lambda_{xx}$ , переводящего функции  $v = \{v_m\}$  из пространства функций, обращающихся в нуль при  $m = 0$  и  $m = M$ , в функции того же пространства по формулам

$$w_m = \frac{1}{h^2} (v_{m+1} - 2v_m + v_{m-1}), \quad m = 1, 2, \dots, M-1.$$

Собственной функции  $\psi^{(k)} = \sqrt{2} \sin \frac{k\pi m}{M}$  соответствует собственное значение

$$\mu_k = -\frac{4}{h^2} \sin^2 \frac{k\pi}{2M}, \quad k = 1, 2, \dots, M-1. \quad (9)$$

**2. Представление решений разностных схем для уравнения теплопроводности на отрезке.** В качестве первого примера, где удастся представить решение в виде конечного ряда Фурье, рассмотрим простейшую разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= 0, \\ m = 1, 2, \dots, M-1, \quad p = 0, 1, \dots, [T/\tau] - 1, & \\ u_0^p = u_M^p = 0, & \\ u_m^0 = \psi(mh) & \end{aligned} \right\} \quad (10)$$

для задачи теплопроводности на отрезке

$$\left. \begin{aligned} u_t - u_{xx} &= 0, & 0 \leq t \leq T, & \quad 0 \leq x \leq 1, \\ u(0, t) = u(1, t) &= 0, & 0 \leq t \leq T, & \\ u(x, 0) &= \psi(x), & 0 \leq x \leq 1. & \end{aligned} \right\} \quad (11)$$

Задачу (10) перепишем так:

$$\left. \begin{aligned} u_m^{p+1} &= u_m^p + \tau \Lambda_{xx} u_m^p = (E + \tau \Lambda_{xx}) u_m^p, \\ u_m^0 &= \psi(mh). \end{aligned} \right\} \quad (12)$$

Здесь  $E$  — тождественное отображение:  $E u_m^p = u_m^p$ , а  $E + \tau \Lambda_{xx}$  — оператор перехода от  $u^p$  к  $u^{p+1}$ , или оператор перехода со слоя на слой. Относительно сеточных функций  $u^p \equiv \{u_m^p\}$  аргумента  $m$

предполагается, что при каждом фиксированном  $p$  они принадлежат рассмотренному пространству, т. е.  $u_0^p = u_M^p = 0$ .

Будем искать решения уравнения (12) в виде

$$\lambda_k^p \psi^{(k)} \equiv \lambda_k^p \left( \sqrt{2} \sin \frac{k\pi m}{M} \right).$$

Подставляя это выражение в уравнение и сокращая обе части на  $\lambda_k^p \sqrt{2} \sin \frac{k\pi m}{M}$ , получим в силу (9) следующее выражение для  $\lambda_k$ :

$$\lambda_k = 1 + \tau \mu_k = 1 - \frac{4\tau}{h^2} \sin^2 \frac{k\pi}{2M}, \quad k = 1, 2, \dots, M-1.$$

Ввиду линейности уравнения (12) выражение

$$u^p = \sum_{k=1}^{M-1} c_k \lambda_k^p \psi^{(k)} \quad (13)$$

является его решением при любых произвольных постоянных  $c_k$ . При  $p = 0$  получаем

$$u_m^0 = \sum_{k=1}^{M-1} c_k \left( \sqrt{2} \sin \frac{k\pi m}{M} \right).$$

Выберем в качестве  $c_k$  коэффициенты разложения заданной функции  $u_m^0 = \psi(mh)$  в конечный ряд Фурье, т. е. положим

$$c_k = (\psi, \psi^{(k)}) = h \sum_{m=0}^M \psi(mh) \left( \sqrt{2} \sin \frac{k\pi m}{M} \right).$$

Тогда решение (13)

$$u_m^p = \sum_{k=1}^{M-1} c_k \left( 1 - \frac{4\tau}{h^2} \sin^2 \frac{\pi k}{2M} \right)^p \left( \sqrt{2} \sin \frac{k\pi m}{M} \right) \quad (14)$$

будет удовлетворять заданному начальному условию  $u_m^0 = \psi(mh)$ . Формула (14) и есть искомое представление решения задачи в виде конечного ряда Фурье.

Коэффициенты  $c_k^{(p)}$  разложения функции  $u_m^{(p)}$  аргумента  $m$  при фиксированном  $p$  по ортонормальному базису  $\psi^{(k)} = \sqrt{2} \sin \frac{k\pi m}{M}$

$$u_m^p = \sum c_k^{(p)} \psi^{(k)}$$

имеют вид

$$c_k^{(p)} = c_k \lambda_k^p.$$

Поэтому, принимая во внимание равенство Парсеваля, получим

$$\begin{aligned} (u^{p+1}, u^{p+1}) &= \sum_{k=1}^{M-1} |c_k^{(p+1)}|^2 = \sum_{k=1}^{M-1} |c_k \lambda_k^{p+1}|^2 \leq \\ &\leq \max_k |\lambda_k|^2 \sum_{k=1}^{M-1} |c_k \lambda_k^p|^2 = \max_k |\lambda_k|^2 (u^p, u^p), \end{aligned}$$

причем строгое равенство  $(u^{p+1}, u^{p+1}) = \max_k |\lambda_k|^2 (u^p, u^p)$  достигается, если в качестве  $u^0$  используется та  $\psi^{(h)}$ , для которой  $|\lambda_k|$  наибольшее.

Если  $\max_k |\lambda_k|^2 \leq 1$ , то имеет место неравенство

$$(u^{p+1}, u^{p+1}) \leq (u^p, u^p). \quad (15)$$

Положительно определенные квадратичные формы вида

$$(Au^p, u^p),$$

где  $A$  — матрица квадратичной формы, напоминают выражения для энергии в уравнениях математической физики. Поэтому неравенства вида

$$(Au^{p+1}, u^{p+1}) \leq (Au^p, u^p)$$

для решений разностных краевых задач называют обычно *энергетическими*.

Таким образом, оценка (15) есть простейшее энергетическое неравенство. В случае выполнения энергетического неравенства выбор норм  $\|\cdot\|_{U_h}$  и  $\|\cdot\|_{Fh}$  естественно связать с формой  $(Au^p, u^p)$ , положив, в частности,  $\|u^{(h)}\|_{U_h} = \max_p (Au^p, u^p)^{1/2}$ . Подобные нормы называются *энергетическими*.

Неравенство  $\max_k |\lambda_k|^2 \leq 1$  выполнено, как легко видеть, в случае, если

$$r = \frac{\tau}{h^2} \leq \frac{1}{2}.$$

При

$$r = \text{const} > \frac{1}{2}$$

и при достаточно малых значениях  $h$  найдутся  $\lambda_h$ ,  $|\lambda_h| > 1$ . Тогда устойчивости нет ни при каком разумном\*) выборе норм.

Рассмотрим разностную схему более общего вида

$$\begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - [(1 - \sigma) \Lambda_{xx} u^p + \sigma \Lambda_{xx} u^{p+1}] &= 0, \\ u_m^0 &= \psi(mh) \end{aligned}$$

\*) См. § 13.



для той же дифференциальной задачи о теплопроводности (11).  
Здесь  $\sigma$  — параметр.

Найдем решения вида

$$w_m^p = \lambda_k^p \sqrt{2} \sin \frac{k\pi m}{M}, \quad k = 1, 2, \dots, M-1,$$

где  $\lambda_k$  подлежат определению.

Подставляя это выражение в разностное уравнение, получим соотношение, которому должно удовлетворять  $\lambda_k$ :

$$\lambda_k = 1 + \tau(1 - \sigma)\mu_k + \tau\sigma\lambda_k\mu_k.$$

Отсюда

$$\lambda_k = \frac{1 - \frac{(1 - \sigma)\tau}{4h^2} \sin^2 \frac{k\pi}{2M}}{1 + \frac{\sigma\tau}{4h^2} \sin^2 \frac{k\pi}{2M}}, \quad k = 1, 2, \dots, M-1.$$

По-прежнему

$$(u^{p+1}, u^{p+1}) \leq \max_k |\lambda_k|^2 (u^p, u^p).$$

Энергетическое неравенство (15) имеет место, если

$$\max |\lambda_k| \leq 1$$

или

$$\left| 1 - \frac{(1 - \sigma)r}{4} \sin^2 \frac{k\pi}{2M} \right| \leq \left| 1 + \frac{\sigma r}{4} \sin^2 \frac{k\pi}{2M} \right|, \quad r = \frac{\tau}{h^2}.$$

Очевидно, что при  $1 \geq \sigma \geq 1/2$  это неравенство — и энергетическое неравенство (15) также — выполняется, каково бы ни было  $r$ . Если  $\sigma = 0$ , то разностная схема превращается в уже рассмотренную явную схему и, как мы видели, для выполнения энергетического неравенства (15) при всех  $h$  нужно, чтобы было  $r \leq 1/2$ .

**3. Представление решений разностных схем для двумерной задачи теплопроводности.** Рассмотрим теперь двумерную задачу теплопроводности

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \\ u(x, y, 0) &= \psi(x, y), \\ u(x, y, t)|_{\Gamma} &= 0, \end{aligned} \right\} \quad \begin{aligned} 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \\ 0 \leq t \leq T. \end{aligned} \quad (16)$$

Здесь через  $\Gamma$  обозначена боковая поверхность параллелепипеда  $0 \leq x, y \leq 1, 0 \leq t \leq T$ .

Построим сетку  $(x_m, y_n, t_p) = (mh, nh, p\tau)$ , причем будем считать  $h = 1/M$ , где  $M$  — натуральное. За  $D_h$  примем точки сетки,

лежащие внутри и на границе параллелепипеда  $0 \leq x, y \leq 1$ ,  $0 \leq t \leq T$ .

Обозначим

$$\Lambda_{xx} u_{mn}^p = \frac{u_{m+1, n}^p - 2u_{mn}^p + u_{m-1, n}^p}{h^2},$$

$$\Lambda_{yy} u_{mn}^p = \frac{u_{m, n+1}^p - 2u_{mn}^p + u_{m, n-1}^p}{h^2}.$$

Операторы  $\Lambda_{xx}$  и  $\Lambda_{yy}$  совершенно аналогичны, только первый действует по переменному  $m$ , в то время как  $n$  и  $p$  — параметры, а второй — по переменному  $n$ , а  $m$  и  $p$  — для него параметры.

Простейшая разностная схема для задачи (16) есть

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^p + \Lambda_{yy} u_{mn}^p, & 0 \leq mh, \quad nh \leq 1, \\ u_{mn}^0 &= \psi(mh, nh), & 0 \leq p\tau < T - \tau, \\ u_{mn}^p|_{\Gamma} &= 0. \end{aligned} \right\} \quad (17)$$

Ищем решения разностного уравнения при условии  $u_{mn}^p|_{\Gamma} = 0$  вида

$$u_{mn}^p = \lambda_{kl}^p \psi_{mn}^{(k, l)}.$$

Заметим, что

$$\begin{aligned} \Lambda_{xx} \psi^{(k, l)} &= \Lambda_{xx} (\psi_m^{(k)} \psi_n^{(l)}) = \psi_n^{(l)} \Lambda_{xx} \psi_m^{(k)} = \mu^{(k)} \psi_m^{(k)} \psi_n^{(l)} = \mu^{(k)} \psi^{(k, l)}, \\ \Lambda_{yy} \psi^{(k, l)} &= \Lambda_{yy} (\psi_m^{(k)} \psi_n^{(l)}) = \mu^{(l)} \psi^{(k, l)}. \end{aligned}$$

Поэтому для  $\lambda_{kl}$  получаем выражение

$$\frac{\lambda_{kl} - 1}{\tau} = [\mu^{(k)} + \mu^{(l)}]$$

или

$$\lambda_{kl} = 1 - \frac{4\tau}{h^2} \left( \sin^2 \frac{k\pi}{2M} + \sin^2 \frac{l\pi}{2M} \right).$$

Решение

$$u^p = \sum_{k, l=1}^{M-1} c_{kl} \lambda_{kl}^p \psi^{(k, l)} \quad (18)$$

удовлетворяет условиям на боковой границе при любом выборе постоянных  $c_{kl}$ . При  $p = 0$  это решение принимает вид

$$u^0 = \sum c_{kl} \psi^{(k, l)}.$$

Для того чтобы выполнялось заданное начальное условие

$$u_{mn}^0 = \psi(mh, nh) = \sum c_{kl} \psi_{mn}^{(k, l)},$$

в качестве  $c_{kl}$  надо взять коэффициенты Фурье функции  $\psi(mh, nh)$ , т. е.

$$c_{kl} = h^2 \sum_{m, n=0}^M \psi(mh, nh) \left( 2 \sin \frac{k\pi m}{M} \sin \frac{l\pi n}{M} \right). \quad (19)$$

В силу формулы (18) коэффициентом при  $\psi^{(k, l)}$  в разложении  $u^p$  в ряд Фурье служит число  $(c_{kl} \lambda_{kl}^p)$ . Поэтому

$$(u^p, u^p) = \sum_{k, l} |c_{kl} \lambda_{kl}^p|^2.$$

При любом фиксированном  $p$  в силу этого можно написать

$$\begin{aligned} (u^{p+1}, u^{p+1}) &= \sum_{k, l=1}^{M-1} |c_{kl} \lambda_{kl}^{p+1}|^2 \leq \\ &\leq \max_{k, l} |\lambda_{kl}|^2 \sum_{k, l} |c_{kl} \lambda_{kl}^p|^2 = \max_{k, l} |\lambda_{kl}|^2 (u^p, u^p). \end{aligned}$$

Равенство достигается, если в качестве  $\psi(mh, nh)$  задана та собственная функция  $\psi_{mn}^{(k, l)}$  оператора перехода  $E + \tau(\Lambda_{xx} + \Lambda_{yy})$  со слоя  $t = p\tau$  на слой  $t = (p+1)\tau$ , собственное число которой  $\lambda_{kl}$  принимает среди всех собственных чисел  $\lambda_{kl}$  наибольшее по модулю значение.

Если  $\max_{k, l} |\lambda_{kl}| \leq 1$ , то имеет место энергетическое неравенство

$$(u^{p+1}, u^{p+1}) \leq (u^p, u^p). \quad (20)$$

Когда  $k$  и  $l$  пробегает значения  $k, l = 1, 2, \dots, M-1$ , собственные числа пробегает некоторое конечное множество точек на вещественной прямой, лежащее левее точки  $\lambda = 1$ . Самая левая точка получается при  $k = l = M-1$ :

$$\begin{aligned} \lambda_{M-1, M-1} &= 1 - 8r \sin^2 \frac{(M-1)\pi}{2M} = \\ &= 1 - 8r \cos^2 \frac{\pi}{2M} = 1 - 8r + O\left(\frac{1}{M^2}\right). \end{aligned}$$

Поэтому неравенство  $\max |\lambda_{kl}| \leq 1$  выполняется при  $-1 \leq 1 - 8r$ ,  $r \leq 1/4$ .

Для неявной разностной схемы

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^{p+1} + \Lambda_{yy} u_{mn}^{p+1}, \\ u_{mn}^0 &= \psi(mh, nh), \\ u_{mn}^p|_{\Gamma} &= 0 \end{aligned} \right\}$$

решение имеет вид

$$u_{mn}^p = \sum_{k, l} c_{kl} \lambda_{kl}^p \Psi_{mn}^{(k, l)},$$

где

$$\lambda_{kl} = \frac{1}{1 + 4r \left( \sin^2 \frac{k\pi}{2M} + \sin^2 \frac{l\pi}{2M} \right)},$$

а коэффициенты  $c_{kl}$  определяются по-прежнему формулой (19). Здесь  $0 < \lambda_{kl} < 1$ , и энергетическое неравенство (20) имеет место при произвольном значении  $r = \tau/h^2$ .

**4. Представление решения разностной схемы для задачи о колебаниях струны.** Рассмотрим пример трехслойной схемы  $L_h u^{(h)} = f^{(h)}$ , аппроксимирующей задачу о колебаниях струны с закрепленными концами:

$$\left. \begin{aligned} u_{tt} - u_{xx} &= 0, & 0 \leq x \leq 1, & \quad 0 \leq t \leq T, \\ u(0, t) = u(1, t) &= 0, & 0 \leq t \leq T, \\ u(x, 0) &= \psi_0(x), & 0 \leq x \leq 1, \\ u_t(x, 0) &= \psi_1(x), & 0 \leq x \leq 1. \end{aligned} \right\}$$

Положим

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{p+1} - 2u_m^p + u_m^{p-1}}{\tau^2} - \Lambda_{xx} u_m^p = 0, \\ u_0^p = u_M^p = 0, \\ u_m^0 = \psi_0(mh), \\ u_m^1 = \bar{\psi}_m. \end{cases}$$

где

$$\bar{\psi}_m =$$

$$= u(mh, 0) + \tau u_t(mh, 0) + \frac{\tau^2}{2} u_{tt}(mh, 0) = \psi_0(mh) + \tau \psi_1(mh) + \frac{\tau^2}{2} \psi_0''(x).$$

Ищем решения разностного уравнения, удовлетворяющие условиям  $u_0^p = u_M^p = 0$ , имеющие вид

$$u_m^p = \lambda^p \sqrt{2} \sin \frac{k\pi m}{M} \quad (\equiv \lambda^p \psi^{(k)}), \quad (21)$$

пока не заботясь о выполнении начальных условий  $u_m^0 = \psi_0(mh)$  и  $u_m^1 = \bar{\psi}_m$ .

Получаем следующее уравнение для  $\lambda$ :

$$\frac{\lambda - 2 + \frac{1}{\lambda}}{\tau^2} - \mu_k = 0, \quad \mu_k = -\frac{4}{h^2} \sin^2 \frac{k\pi}{2M},$$

$$\lambda^2 - 2 \left( 1 - 2r^2 \sin^2 \frac{k\pi}{2M} \right) + 1 = 0, \quad r = \frac{\tau}{h},$$

$$\lambda_1(k) = 1 - 2r^2 \sin^2 \frac{k\pi}{2M} + \sqrt{\left( 1 - 2r^2 \sin^2 \frac{k\pi}{2M} \right)^2 - 1},$$

$$\lambda_2(k) = 1 - 2r^2 \sin^2 \frac{k\pi}{2M} - \sqrt{\left( 1 - 2r^2 \sin^2 \frac{k\pi}{2M} \right)^2 - 1}.$$

Таким образом, существуют два решения искомого вида (21):

$$\lambda_1^p(k) \psi^{(k)} \quad \text{и} \quad \lambda_2^p(k) \psi^{(k)}.$$

Ввиду линейности задачи выражение

$$u_m^p = \sum_{k=1}^{M-1} [\alpha_k \lambda_1^p(k) + \beta_k \lambda_2^p(k)] \psi_m^{(k)}$$

является решением при произвольном выборе чисел  $\alpha_k$  и  $\beta_k$ ,  $k = 1, 2, \dots, M-1$ . При  $p = 0$  и  $p = 1$  получаем соответственно

$$u_m^0 = \psi_0(mh) = \sum_{k=1}^{M-1} (\alpha_k + \beta_k) \psi^{(k)},$$

$$u_m^1 = \bar{\psi}_m = \sum_{k=1}^{M-1} [\alpha_k \lambda_1(k) + \beta_k \lambda_2(k)] \psi^{(k)}.$$

Эти соотношения определяют значения чисел  $\alpha_k$ ,  $\beta_k$ . Сумма  $\alpha_k + \beta_k$  должна быть коэффициентом Фурье разложения  $\psi_0(mh)$  по функциям  $\{\psi^{(k)}\}$ , т. е.

$$\alpha_k + \beta_k = h \sum_{m=0}^M \psi_0(mh) \left( \sqrt{2} \sin \frac{k\pi m}{M} \right).$$

Точно так же

$$\alpha_k \lambda_1(k) + \beta_k \lambda_2(k) = h \sum_{m=0}^M \bar{\psi}_m \left( \sqrt{2} \sin \frac{k\pi m}{M} \right).$$

Запись решения разностного уравнения в виде конечного ряда Фурье используется не только для выяснения условий, при которых имеют место энергетические неравенства. В дальнейшем мы многократно будем пользоваться такими представлениями с различными целями при качественном изучении модельных задач.

Надо отметить только, что представления решений в виде конечных рядов Фурье редко используются непосредственно для вычисления решения. Дело в том, что удобные вычислительные методы должны быть пригодны для широкого класса задач. Выписанные нами разностные схемы легко обобщаются на случай переменных коэффициентов и неправильных областей, причем мы можем ожидать сохранения таких свойств, как выполнение энергетического неравенства. Но каждое такое изменение задачи нарушает возможность записать ее решение в виде ряда Фурье: мы не можем обычно выписать собственные функции оператора перехода со слоя на слой и не можем вычислить отвечающие им собственные значения.

## ЗАДАЧИ

1. Для двумерной задачи теплопроводности в квадратной области с нулевыми значениями на границе рассмотреть разностную схему

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \sigma [\Lambda_{xx} u_{mn}^{p+1} + \Lambda_{yy} u_{mn}^{p+1}] + \\ &+ (1 - \sigma) [\Lambda_{xx} u_{mn}^p + \Lambda_{yy} u_{mn}^p], \quad 0 < mh, \quad nh < 1, \\ u_{mn}^p \Big|_{\Gamma} &= 0, \quad u_{mn}^0 = \psi(mh, nh) \end{aligned} \right\}$$

(обозначения введены в тексте параграфа). Выписать решение этой задачи в виде ряда Фурье. Выяснить, при каких значениях параметра  $\sigma$ ,  $0 \leq \sigma \leq 1$ , имеет место энергетическое неравенство  $(u^{p+1}, u^{p+1}) \leq (u^p, u^p)$  независимо от выбора соотношения шагов  $\tau = \tau/h^2$ .

При каких  $\sigma$  для любого  $u^p \neq 0$  выполнено строгое неравенство  $(u^{p+1}, u^{p+1}) < (u^p, u^p)$  независимо от выбора  $\tau$  и шага  $h$ ?

2. Записать решения дифференциальной задачи

$$\begin{aligned} u_t - u_{xx} &= 0, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq T, \\ u \Big|_{\Gamma} &= 0, \quad u(x, 0) = \psi(x) \end{aligned}$$

и разностной задачи

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} - \Lambda_{xx} u_{mn}^p &= 0, \\ u_{mn}^p \Big|_{\Gamma} &= 0, \quad u_{mn}^0 = \psi(mh, nh) \end{aligned} \right\}$$

соответственно в виде ряда Фурье и конечного ряда Фурье. Доказать путем сравнения этих рядов при  $\tau \leq 1/2$  в предположении ограниченности  $\psi''(x)$ , что решение разностной задачи сходится к решению дифференциальной задачи. Доказать, что при  $\tau > 1/2$  сходимость, вообще говоря, не имеет места.

3. Выписать в виде конечного ряда Фурье решение разностной задачи Дирихле для уравнения Пуассона в квадратной области  $0 \leq x, y \leq 1$ :

$$\Lambda_{xx} u_{mn} + \Lambda_{yy} u_{mn} = \varphi(mh, nh), \quad 0 < mh, \quad nh < 1,$$

при граничном условии:

$$\begin{aligned} \text{а)} \quad & u_{mn} \Big|_{\Gamma} = 0; \\ \text{б)} \quad & u_{mn} = \begin{cases} nh, & \text{если } m = 0, \\ 1 + mh, & \text{если } n = M, \\ mh, & \text{если } n = 0, \\ 1 + nh, & \text{если } m = M. \end{cases} \end{aligned}$$

Указание к 3, б).  $u_{mn} = mh + nh + Z_{mn}$ , где  $Z_{mn}$  удовлетворяет однородным условиям на границе.

## § 28. Принцип максимума

Мы уже видели в §§ 21 и 24 на примерах, как доказывается устойчивость с помощью принципа максимума. Здесь мы разберем еще два интересных примера, в которых этим приемом

удается доказать устойчивость: явную и неявную разностные схемы, аппроксимирующие краевую задачу для уравнения теплопроводности:

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - a^2(x, t) \frac{\partial^2 u}{\partial x^2} &= \varphi(x, t), & 0 < x < 1, & 0 \leq t \leq T, \\ u(x, 0) &= \psi_0(x), & 0 < x < 1, \\ u(0, t) &= \psi_1(t), & 0 \leq t \leq T, \\ u(1, t) &= \psi_2(t), & 0 \leq t \leq T. \end{aligned} \right\} \quad (1)$$

**1. Явная разностная схема.** Рассмотрим явную разностную схему

$$L_h u^{(h)} \equiv \left\{ \begin{aligned} \frac{u_m^{n+1} - u_m^n}{\tau} - a^2(mh, n\tau) \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} &= \varphi(mh, n\tau), \\ m = 1, 2, \dots, M-1; \quad n = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi_0(mh), \quad m = 0, 1, \dots, Mh, \\ u_0^n &= \psi_1(n\tau), \quad n = 1, 2, \dots, [T/\tau], \\ u_M^n &= \psi_2(n\tau), \quad n = 1, 2, \dots, [T/\tau]. \end{aligned} \right. \quad (2)$$

Здесь  $M = 1/h$  — целое число.

Спектральный признак Неймана в соединении с принципом замороженных коэффициентов приводит, как мы видели в § 26, к необходимому условию устойчивости

$$\frac{\tau}{h^2} \leq \frac{1}{2 \max_{x, t} a^2(x, t)}. \quad (3)$$

Докажем, что при этом условии устойчивость действительно имеет место, если определить нормы равенствами

$$\left. \begin{aligned} \|u^{(h)}\|_{U_h} &= \max_n \max_m |u_m^n|, \\ \|f^{(h)}\|_{F_h} &= \max \left( \max_m |\psi_0(mh)|, \max_n |\psi_1(n\tau)|, \right. \\ &\quad \left. \max_n |\psi_2(n\tau)|, \max_{m, n} |\varphi(x_m, t_n)| \right). \end{aligned} \right\} \quad (4)$$

Установим справедливость неравенства (*принцип максимума*)

$$\max_m |u_m^{n+1}| \leq \max \left[ \max_n |\psi_1(t_n)|, \max_n |\psi_2(t_n)|, \right. \\ \left. \max_m |u_m^n| + \tau \max_{m, n} |\varphi(x_m, t_n)| \right]. \quad (5)$$

Действительно, перепишем разностное уравнение, лежащее в основе схемы (2), придав ему вид

$$u_m^{n+1} = (1 - 2ra^2(x_m, t_n))u_m^n + ra^2(x_m, t_n)(u_{m-1}^n + u_{m+1}^n) + \tau\varphi(x_m, t_n), \quad m = 1, \dots, M-1. \quad (6)$$

При выполнении условия (3) выражение  $1 - 2ra^2(x_m, t_n)$  неотрицательно. Поэтому можно написать

$$\begin{aligned} |u_m^{n+1}| &\leq [1 - 2ra^2(x_m, t_n)] \max_k |u_k^n| + \\ &+ ra^2(x_m, t_n) (\max_k |u_k^n| + \max_k |u_k^n|) + \tau \max_{m, n} |\varphi(x_m, t_n)| = \\ &= \max_k |u_k^n| + \tau \max_{m, n} |\varphi(x_m, t_n)|, \quad m = 1, 2, \dots, M-1. \quad (7) \end{aligned}$$

Учитывая, что

$$u_0^{n+1} = \psi_1[(n+1)\tau], \quad u_M^{n+1} = \psi_2[(n+1)\tau], \quad (8)$$

отсюда выводим принцип максимума (5).

Разобьем решение  $u^{(h)}$  задачи  $L_h u^{(h)} = f^{(h)}$  на два слагаемых:  $u^{(h)} = v^{(h)} + w^{(h)}$ , определив  $v^{(h)}$  и  $w^{(h)}$  соответственно как решения задач

$$L_h v^{(h)} = \begin{cases} 0 \\ \psi_0(x_m), \\ \psi_1(t_n), \\ \psi_2(\tau_n) \end{cases}, \quad L_h w^{(h)} = \begin{cases} \varphi(x_m, t_n), \\ 0, \\ 0, \\ 0. \end{cases} \quad (9)$$

В силу оценки (5)

$$\begin{aligned} \max_m |v_m^{n+1}| &\leq \max [\max_k |\psi_1(t_k)|, \max_k |\psi_2(t_k)|, \max_m |v_m^n|], \\ \max_m |v_m^n| &\leq \max [\max_k |\psi_1(t_k)|, \max_k |\psi_2(t_k)|, \max_m |v_m^{n-1}|], \\ \max_m |v_m^{n-1}| &\leq \max [\max_k |\psi_1(t_k)|, \max_k |\psi_2(t_k)|, \max_m |v_m^{n-2}|], \\ &\dots \\ \max_m |v_m^1| &= \max [\max_k |\psi_1(t_k)|, \max_k |\psi_2(t_k)|, \max_m |\psi_0(x_m)|]. \end{aligned}$$

В силу той же оценки (5)

$$\begin{aligned} \max_m |\omega_m^{n+1}| &\leq \max_m |\omega_m^n| + \tau \max_{m, k} |\varphi(x_m, t_k)| \leq \\ &\leq \max_m |\omega_m^{n-1}| + 2\tau \max_{m, k} |\varphi(x_m, t_k)| \leq \\ &\dots \\ &\leq \max_m |\omega_m^0| + (n+1)\tau \max_{m, k} |\varphi(x_m, t_k)| \leq T \max_{m, k} |\varphi(x_m, t_k)|. \end{aligned}$$



Из оценок, установленных для  $v_m^{n+1}$  и  $\omega_m^{n+1}$ , следует

$$\begin{aligned} \max_m |u_m^{n+1}| &= \max_m |v_m^{n+1} + \omega_m^{n+1}| \leq \max_m |v_m^{n+1}| + \max_m |\omega_m^{n+1}| \leq \\ &\leq \max [\max_k |\psi_1(t_k)|, \max_k |\psi_2(t_k)|, \max_m |\psi_0(x_m)|] + \\ &\quad + T \max_{m, n} |\varphi(x_m, t_n)| \leq c \|f^{(h)}\|_{F_h}, \end{aligned} \quad (10)$$

где

$$c = 2 \max(1, T). \quad (11)$$

Неравенство справедливо при всех  $n$ . Поэтому

$$\|u^{(h)}\|_{U_h} \leq c \|f^{(h)}\|_{F_h}, \quad (12)$$

и устойчивость имеет место.

**2. Неявная разностная схема.** Теперь рассмотрим неявную разностную схему

$$L_h u^h = \begin{cases} \frac{u_m^{n+1} - u_m^n}{\tau} - a^2(x_m, t_n) \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} = \varphi(x_m, t_n), \\ u_m^0 = \psi_0(x_m), \\ u_0^n = \psi_1(t_n), \\ u_M^n = \psi_2(t_n). \end{cases} \quad (13)$$

Для того, чтобы, зная значения  $u_m^n$ ,  $m = 0, 1, \dots, M$ , вычислить значения  $u_m^{n+1}$ ,  $m = 0, 1, \dots, M$ , надо решить задачу

$$\left. \begin{aligned} \frac{u_m^{n+1}}{\tau} - a^2(x_m, t_n) \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} &= \frac{u_m^n}{\tau} + \varphi(x_m, t_n), \\ u_0^{n+1} &= \psi_1(t_{n+1}), \quad u_M^{n+1} = \psi_2(t_{n+1}). \end{aligned} \right\} \quad (14)$$

Эта задача после умножения обеих частей разностного уравнения на  $-\tau$  примет вид

$$\left. \begin{aligned} a_m v_{m-1} + b_m v_m + c_m v_{m+1} &= g_m, \quad m = 1, 2, \dots, M-1, \\ v_0 &= \alpha, \quad v_M = \beta, \end{aligned} \right\} \quad (15)$$

где

$$\begin{aligned} v_m &= u_m^{n+1}, \quad a_m = a^2(x_m, t_n) r, \\ b_m &= -2a^2(x_m, t_n) r - 1, \quad c_m = a^2(x_m, t_n) r, \\ g_m &= -u_m^n - \tau \varphi(x_m, t_n), \quad \alpha = \psi_1(t_{n+1}), \quad \beta = \psi_2(t_{n+1}). \end{aligned}$$

Коэффициенты  $a_m$ ,  $b_m$ ,  $c_m$  удовлетворяют условиям

$$a_m > 0, \quad c_m > 0, \quad |b_m| > a_m + c_m + \delta \quad (\delta > 0).$$

Поэтому, как показано в §§ 4, 5, задача имеет единственное решение

$$(v_0, v_1, \dots, v_M) = (u_0^{n+1}, u_1^{n+1}, \dots, u_M^{n+1}),$$

которое можно найти методом прогонки.

Для доказательства устойчивости осталось показать выполнение неравенства (12). Для этого докажем, что имеет место неравенство (5) (принцип максимума), из которого оценки (12), (10) выводятся дословно так же, как для явной схемы (2).

Из всех значений  $u_m^{n+1}$ , по модулю равных  $\max_m |u_m^{n+1}|$ , возьмем то, у которого индекс  $m$  принимает наименьшее значение  $m = m^*$ . Если  $m^* = 0$  или  $m^* = M$ , то в силу (8) выполнение неравенства (5) очевидно. Пусть  $m^* \neq 0$  и  $m^* \neq M$ . Выпишем отвечающее этому значению  $m = m^*$  уравнение (14):

$$\begin{aligned} ra^2(x_{m^*}, t_n) u_{m^*-1}^{n+1} - (1 + 2ra^2(x_{m^*}, t_n)) u_{m^*}^{n+1} + ra^2(x_{m^*}, t_n) u_{m^*+1}^{n+1} = \\ = -u_{m^*}^n - \tau\varphi(x_{m^*}, t_n). \end{aligned}$$

Пусть для определенности  $u_{m^*}^{n+1} > 0$ . Тогда левую часть этого равенства оценим так:

$$\begin{aligned} ra^2(x_{m^*}, t_n) u_{m^*-1}^{n+1} - (1 + 2ra^2(x_{m^*}, t_n)) u_{m^*}^{n+1} + ra^2(x_{m^*}, t_n) u_{m^*+1}^{n+1} = \\ = a^2(x_{m^*}, t_n) [(u_{m^*-1}^{n+1} - u_{m^*}^{n+1}) + (u_{m^*+1}^{n+1} - u_{m^*}^{n+1})] - u_{m^*}^{n+1} \leq -u_{m^*}^{n+1}. \end{aligned}$$

Поэтому

$$-u_{m^*}^{n+1} \geq -u_{m^*}^n - \tau\varphi(x_{m^*}, t_n),$$

$$\begin{aligned} \max_m |u_m^{n+1}| = u_{m^*}^{n+1} \leq |u_{m^*}^n + \tau\varphi(x_{m^*}, t_n)| \leq \\ \leq \max_m u_m^n + \tau \max_{m, n} |\varphi(x_m, t_n)|. \end{aligned}$$

**3. Сопоставление явной и неявной разностных схем.** Таким образом, неравенство (5), означающее справедливость принципа максимума, доказано. Вместе с тем доказана и устойчивость неявной разностной схемы (14) в нормах (4).

*Подчеркнем существенную разницу между явной и неявной разностными схемами (2) и (14). Первая из них требует для*

*устойчивости ограничения на шаг*

$$\tau \leq \frac{1}{2 \max a^2(x, t)} h^2,$$

*которое становится очень жестким, если коэффициент  $a^2(x, t)$  принимает большие значения хотя бы в малой окрестности какой-либо одной точки. Вторая, неявная разностная схема остается устойчивой при произвольном соотношении шагов  $h$  и  $\tau$ .*

Разностные схемы, которые подобно неявной схеме (14) остаются устойчивыми при произвольном соотношении шагов сетки, называют *абсолютно устойчивыми* или *безусловно устойчивыми*. Явная схема (2) не является абсолютно устойчивой.

## ПОНЯТИЕ О РАЗНОСТНЫХ СХЕМАХ ДЛЯ РАСЧЕТА ОБОБЩЕННЫХ РЕШЕНИЙ

### § 29. Обобщенное решение

Во всех рассмотренных до сих пор примерах мы предполагали, что существуют достаточно гладкие решения дифференциальных краевых задач, а в основу построения разностных схем клали приближенную замену производных в дифференциальном уравнении разностными отношениями. Однако дифференцируемых функций недостаточно для описания многих важных процессов физики. Так, например, физические эксперименты показывают, что распределения давления, плотности и температуры в сверхзвуковом течении вязкого газа описываются функциями, имеющими скачки — ударные волны. Скачки могут возникать с течением времени при гладких начальных данных.

Соответствующие дифференциальные краевые задачи не имеют гладких решений. Приходится расширить понятие решения и некоторым естественным способом ввести обобщенные решения, которые могут быть и разрывными. Для этого существуют два основных способа.

Первый способ состоит в том, чтобы записывать физические законы сохранения (массы, импульса, энергии и т. д.) не в дифференциальной, а в интегральной форме. Тогда они имеют смысл и для разрывных функций, которые нельзя дифференцировать, но интегрировать можно.

Второй способ состоит в искусственном введении в дифференциальные уравнения таких членов, при которых эти уравнения имеют гладкие решения. Эти искусственно введенные члены в случае газодинамических задач имеют смысл малой вязкости, выглаживающей разрывы течения. Затем коэффициенты при «вязких» членах устремляют к нулю, а предел, к которому стремится решение, принимают за обобщенное решение исходной задачи.

Мы поясним определение обобщенного решения и способов расчета на примере следующей задачи Коши:

$$\left. \begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} &= 0, & 0 < t < T, & \quad -\infty < x < \infty, \\ u(x, 0) &= \psi(x), & & \quad -\infty < x < \infty, \end{aligned} \right\} \quad (1)$$

которая является простейшей моделью задач газовой динамики среди обладающих свойством возникновения разрывных решений из гладких начальных данных.

**1. Механизм возникновения разрывов.** Предположим сначала, что задача (1) имеет гладкое решение  $u(x, t)$ . Введем линии  $x = x(t)$ , определяемые уравнением

$$\frac{dx}{dt} = u(x, t). \quad (2)$$

Эти линии называются *характеристиками* уравнения  $u_t + uu_x = 0$ .

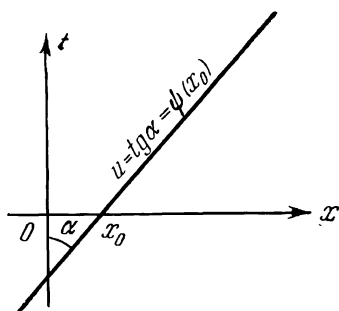


Рис. 28.

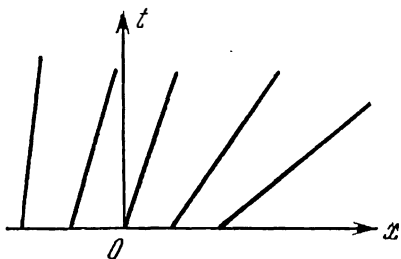


Рис. 29.

Вдоль каждой характеристики  $x = x(t)$  решение  $u(x, t)$  можно считать функцией, зависящей только от  $t$ :

$$u(x, t) = u[x(t), t] = u(t).$$

Тогда

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} \frac{dx}{dt} = \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0.$$

Поэтому вдоль характеристики решение постоянно,  $u(x, t) = \text{const}$ . Но в силу уравнения (2) из  $u = \text{const}$  следует, что характеристика есть прямая линия  $x = ut + x_0$ . Здесь  $x_0$  — абсцисса точки  $(x_0, 0)$ , из которой выходит характеристика, а  $u = \psi(x_0)$  — угловой коэффициент ее наклона к оси  $Ot$ . Заданием начальной функции  $u(x, 0) = \psi(x)$ , таким образом, наглядно определяется и картина характеристик, и значения решения  $u(x, t)$  в каждой точке полуплоскости  $t > 0$  (рис. 28).

Заметим сразу же, что в предположении существования гладкого решения  $u(x, t)$  характеристики не могут пересекаться, так как каждая характеристика приносила бы в точку пересечения свое значение решения и решение не было бы однозначной функцией. При монотонно возрастающей функции  $\psi(x)$  с ростом  $x_0$  угол  $\alpha$  увеличивается, характеристики не пересекаются (рис. 29). Но в случае убывания функции  $\psi(x)$  характеристики сходятся и пересечения неизбежны — независимо от гладкости функции  $\psi(x)$ . Гладкое решение задачи (1) перестает существовать с момента  $t = \bar{t}$ , когда хотя бы две характеристики пересекутся (рис. 30).

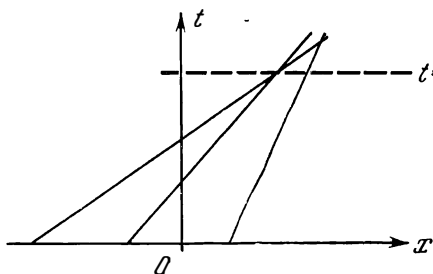


Рис. 30.

Графики функции  $u = u(x, t)$  при  $t = 0$ ,  $\frac{1}{2} \bar{t}$  и  $\bar{t}$ , изображены на рис. 31.

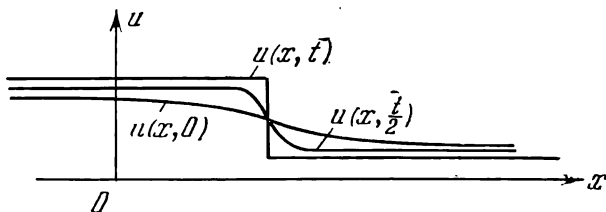


Рис. 31

**2. Определение обобщенного решения.** Напомним формулу Грина, которой будем пользоваться при определении обобщенного решения задачи (1). Пусть  $D$  — произвольная область с границей  $\Gamma$  на плоскости  $Oxt$ , и пусть  $\Phi_1(x, t)$  и  $\Phi_2(x, t)$  имеют в области  $D$  непрерывные вплоть до границы  $\Gamma$  частные производные. Тогда справедлива следующая формула Грина:

$$\iint_D \left( \frac{\partial \Phi_1}{\partial t} + \frac{\partial \Phi_2}{\partial x} \right) dx dt = \oint_{\Gamma} \Phi_1 dx - \Phi_2 dt. \quad (3)$$

Выражение  $\frac{\partial \Phi_1}{\partial t} + \frac{\partial \Phi_2}{\partial x}$  есть *дивергенция* вектора  $\Phi = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}$ . Формула Грина (3) означает, что интеграл от дивергенции векторного поля  $\Phi$  по области  $D$  равен потоку вектора  $\Phi$  через границу  $\Gamma$  этой области.

Переходим к определению понятия обобщенного решения. Запишем дифференциальное уравнение из задачи (1) в дивергентной форме:

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) = 0. \quad (4)$$

Проинтегрируем обе части уравнения (4) по произвольной области  $D$ , лежащей в полуплоскости  $t \geq 0$ . Получим

$$0 = \iint_D \left[ \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) \right] dx dt = \oint_{\Gamma} u dx - \frac{u^2}{2} dt.$$

Таким образом, каждое дифференцируемое решение уравнения (4) удовлетворяет интегральному соотношению

$$\oint_{\Gamma} u dx - \frac{u^2}{2} dt = 0, \quad (5)$$

где  $\Gamma$  — произвольный контур, лежащий в полуплоскости  $t > 0$ . Равенство (5) выражает некоторый закон сохранения: поток вектора  $\left( \frac{u}{u^2/2} \right)$  через любой замкнутый контур равен нулю.

Покажем, что, и наоборот, если гладкая функция удовлетворяет при любом контуре  $\Gamma$  интегральному закону сохранения (5), то в каждой точке  $(x_0, t_0)$ ,  $t_0 > 0$ , выполнено равенство (4). Предположим противное, и пусть для определенности в некоторой точке  $(x_0, t_0)$  будет

$$\left. \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) \right|_{\substack{x=x_0 \\ t=t_0}} > 0.$$

Тогда в силу непрерывности можно найти столь малый круг  $D$  с границей  $\Gamma$  и с центром в точке  $(x_0, t_0)$ , всюду в котором

$u_t + \left( \frac{u^2}{2} \right)_x > 0$ . Получим

$$0 = \oint_{\Gamma} u dx - \frac{u^2}{2} dt = \iint_D \left[ \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) \right] dx dt > 0.$$

Возникшее противоречие  $0 > 0$  доказывает, что из (5) в случае гладкой функции  $u(x, t)$  следует (4), так что (4) и (5) равносильны. Но в случае разрывной функции  $u(x, t)$  дифференциальное уравнение (1) или (4) на линии разрыва теряет смысл, а интегральное условие (5) смысла не теряет. Поэтому будем называть *обобщенным решением* уравнения (4) всякую кусочно-дифференцируемую функцию, удовлетворяющую при произвольном выборе контура  $\Gamma$  в полуплоскости  $t \geq 0$  условию (5).

**3. Условие на линии разрыва решения.** Пусть внутри области, где отыскивается решение, имеется линия  $x \equiv x(t)$ , на которой обобщенное решение  $u(x, t)$  терпит разрыв первого рода. Пусть при приближении к этой линии слева или справа получаем на ней соответственно

$$u(x, t) = u_{\text{лев}}(x, t),$$

$$u(x, t) = u_{\text{прав}}(x, t).$$

Оказывается, что значения  $u_{\text{лев}}(x, t)$ ,  $u_{\text{прав}}(x, t)$  и скорость движения точки разрыва  $\dot{x} = dx/dt$  не могут быть произвольны: они связаны между собой некоторым соотношением.

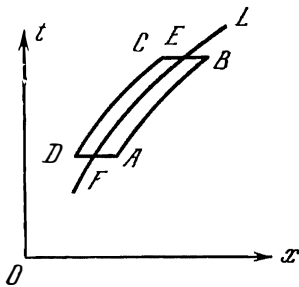


Рис. 32.

Пусть  $L$  является линией разрыва (рис. 32). Интеграл  $\int_{ABCD A} u dx - \frac{u^2}{2} dt$

по контуру  $ABCD A$ , как и по любому другому контуру, обращается в нуль. Когда отрезки  $BC$  и  $DA$  стягиваются к точкам  $E$  и  $F$  соответственно, интегралы по ним обращаются в нуль и получается равенство

$$\int_{L'} [u] dx - \left[ \frac{u^2}{2} \right] dt = 0,$$

или

$$\int_{L'} \left( [u] \frac{dx}{dt} - \left[ \frac{u^2}{2} \right] \right) dt = 0,$$

где  $[z] = z_{\text{прав}} - z_{\text{лев}}$  — скачок величины  $z$  на линии разрыва, а  $L'$  — произвольный участок этой линии  $L' = EF$ .

Ввиду произвольности участка  $L'$  в каждой точке линии  $L$  должна обращаться в нуль подынтегральная функция:

$$[u] \frac{dx}{dt} - \left[ \frac{u^2}{2} \right] = 0.$$

Отсюда

$$\frac{dx}{dt} = \left[ \frac{u^2}{2} \right] : [u] = \frac{u_{\text{прав}}^2 - u_{\text{лев}}^2}{2(u_{\text{прав}} - u_{\text{лев}})} = \frac{u_{\text{лев}} + u_{\text{прав}}}{2},$$

так что на линии разрыва выполнено условие

$$\frac{dx}{dt} = \frac{u_{\text{лев}} + u_{\text{прав}}}{2}. \quad (6)$$

Если бы мы записали уравнение  $u_t + uu_x = 0$  в другой дивергентной форме:

$$\frac{\partial}{\partial t} \left( \frac{u^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u^3}{3} \right) = 0, \quad (7)$$



то пришли бы аналогичным путем к другому интегральному соотношению:

$$\oint_{\Gamma} \frac{u^2}{2} dx - \frac{u^3}{3} dt = 0, \quad (8)$$

и к другому условию на линии разрыва:

$$\frac{dx}{dt} = \frac{2}{3} \frac{u_{лев}^2 + u_{лев} u_{прав} + u_{прав}^2}{u_{лев} + u_{прав}}. \quad (9)$$

Наклон (9) линии разрыва (или скорость  $\dot{x}$  ударной волны) не совпадает с наклоном (6), отвечающим первой дивергентной записи (4). Отсюда видно, что понятие обобщенного решения зависит от того, какой именно интегральный закон сохранения отражается заданным дифференциальным уравнением (1). В задачах математической физики интегральные законы сохранения имеют вполне определенный физический смысл.

На гладких функциях  $u$  все пять форм записи

$$\begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} &= 0, \\ \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) &= 0, \\ \frac{\partial}{\partial t} \left( \frac{u^2}{2} \right) + \frac{\partial}{\partial x} \left( \frac{u^3}{3} \right) &= 0, \\ \oint_{\Gamma} u dx - \frac{u^2}{2} dt &= 0, \\ \oint_{\Gamma} \frac{u^2}{2} dx - \frac{u^3}{3} dt &= 0 \end{aligned}$$

равносильны между собой.

В дальнейшем, рассматривая задачу Коши (1), мы будем иметь в виду выполнение интегрального закона сохранения (5) и вытекающего из него условия (6) на разрыве.

**4. Распад произвольного разрыва.** Пусть заданы разрывные начальные данные

$$u = \begin{cases} 2 & \text{при } x < 0, \\ 1 & \text{при } x > 0. \end{cases}$$

Построенное по этим начальным данным решение изображено на рис. 33.

Тангенс угла наклона линии разрыва  $\frac{dx}{dt} = \frac{2+1}{2} = \frac{3}{2}$  является средним арифметическим из тангенсов углов наклона характеристик по обе стороны от нее.

Зададим теперь в начальных условиях другой разрыв:

$$u = \begin{cases} 1 & \text{при } x < 0, \\ 2 & \text{при } x > 0. \end{cases}$$

Из рис. 34 видно, что возможны два способа построения решения. В первом способе мы получаем непрерывное решение, а во

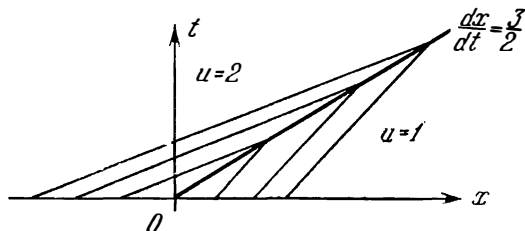


Рис. 33.

втором — разрывное при  $t > 0$ . Следует предпочесть непрерывное решение. В пользу этого говорит следующее рассуждение. Если несколько изменить начальные данные, задав их формулой

$$u = \begin{cases} 1 & \text{при } x \leq 0, \\ 2 & \text{при } x \geq \varepsilon, \\ 1 + x/\varepsilon & \text{при } 0 \leq x \leq \varepsilon, \end{cases}$$

то решение  $u$  определится однозначно. Оно изображено на рис. 35. При стремлении  $\varepsilon$  к нулю это решение переходит в непрерывное решение, изображенное на рис. 34, а. Запрет решения, изображенного на рис. 34, б, по причине его неустойчивости

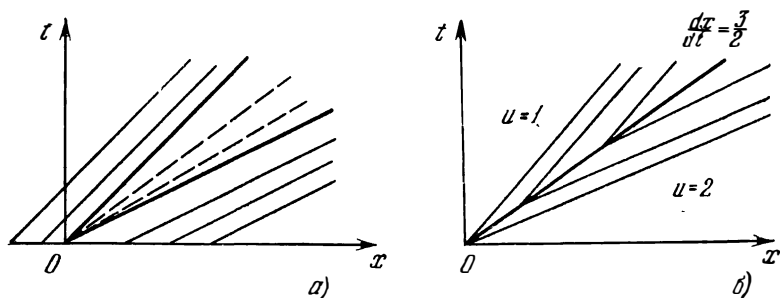


Рис. 34

относительно возмущения начальных данных аналогичен запрету ударных волн разрежения при математическом описании течения идеального газа.

**5. Другое определение обобщенного решения.** Для определения понятия обобщенного решения задачи (1) можно рассмотреть вспомогательную задачу

$$\left. \begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} &= \mu \frac{\partial^2 u}{\partial x^2}, \\ u(x, 0) &= \psi(x). \end{aligned} \right\} \quad (10)$$

Здесь дифференциальное уравнение уже не гиперболического, а параболического типа. Его решения сохраняют гладкость, если  $\psi(x)$  — гладкая функция. А если  $u(x, 0) = \psi(x)$  разрывно, то разрыв сглаживается. Параметр  $\mu > 0$  играет роль вязкости

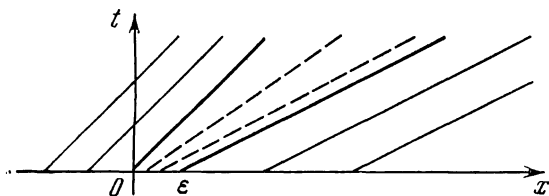


Рис. 35.

в газовой динамике. При  $\mu \rightarrow 0$  решение задачи (10) стремится к пределу, который можно принять за обобщенное решение задачи (1). Можно показать, что для задачи (1) последнее определение обобщенного решения равносильно определению с помощью закона сохранения (5).

### § 30. Построение разностных схем

Перейдем теперь к вопросу о построении разностных схем для задачи

$$\left. \begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} &= 0, \\ u(x, 0) &= \psi(x). \end{aligned} \right\} \quad (1)$$

Будем предполагать для определенности, что  $\psi(x) > 0$ . Тогда  $u(x, t) > 0$ . Первое, что кажется естественным, — это рассмотреть разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} + u_m^p \frac{u_m^p - u_{m-1}^p}{h} &= 0, \quad p = 0, 1, \dots, \\ m &= 0, \pm 1, \dots, \\ u_m^1 &= \psi(x_m). \end{aligned} \right\} \quad (2)$$

Замораживая коэффициент  $u_m^p$  в точке  $m = m_0$ , мы видим, что для возникающего уравнения с постоянными коэффициентами при переходе на слой  $t = (p + 1)\tau$  выполняется принцип максимума, если шаг  $\tau = \tau_p$  выбран из условия

$$r_p = \frac{\tau_p}{h} \leq \frac{1}{\max_m |u_m^p|}.$$

Поэтому можно ожидать устойчивости. Если решение задачи (1) гладкое, то аппроксимация задачи (1) задачей (2) не вызывает сомнения. Действительно, в этом случае экспериментальные расчеты заранее известных гладких решений подтверждают сходимость.

Однако если задача (1) имеет разрывное решение, то сходимости к обобщенному решению, заданному, скажем, интегральным законом сохранения

$$\oint_{\Gamma} u \, dx - \frac{u^2}{2} \, dt = 0, \quad (3)$$

ни в каком разумном смысле ожидать нет оснований. Ведь в используемую разностную схему (2) не заложена информация о том, какой именно закон сохранения — (8) из § 29, (3), или какой-нибудь другой — положен нами в основу определения обобщенного решения.

Поэтому при построении разностной схемы надо использовать либо интегральный закон сохранения, соответствующий искомому обобщенному решению, скажем закон (3), либо уравнение с искусственной вязкостью (10) § 29:

$$u_t + u_{ux} = \mu u_{xx}, \quad (4)$$

осуществляющее при  $\mu \rightarrow 0$  отбор интересующего нас обобщенного решения.

**1. Схема с искусственной вязкостью.** Отметим сразу же, что разностная схема с искусственно введенной малой вязкостью

$$\frac{u_m^{p+1} - u_m^p}{\tau} + u_m^p \frac{u_m^p - u_{m-1}^p}{h} = \mu \frac{u_{m-1}^p - 2u_m^p + u_{m+1}^p}{h^2},$$

$$u_m^0 = \psi(x_m)$$

имеет решение  $u^{(h)} = \{u_m^p\}$ , равномерно сходящееся при  $h \rightarrow 0$  и достаточно малом  $\tau = \tau(h, \mu)$  к искомому обобщенному решению задачи (1) вне сколь угодно малых окрестностей линий разрыва обобщенного решения. При этом  $\mu = \mu(h)$  должно при  $h \rightarrow 0$  достаточно медленно стремиться к нулю. Различные разностные схемы, использующие искусственную вязкость, успеш-

но применяются при газодинамических расчетах. Их недостатком является размазывание разрывов.

Остановимся подробно на построении разностных схем на основе интегрального закона сохранения (3).

Можно наметить два подхода. При первом используется не только сам интегральный закон сохранения (3), но также и вытекающее из него условие на разрыве

$$\frac{dx}{dt} = \frac{u_{\text{лев}} + u_{\text{прав}}}{2}. \quad (5)$$

При втором разрывы не выделяются и расчет ведется по единым формулам во всех расчетных точках.

**2. Метод характеристик.** Наиболее четко идея выделения разрыва при расчете обобщенного решения реализуется в методе характеристик, который можно считать одним из вариантов метода конечных разностей. Развитие возникающих в процессе расчета, т. е. при увеличении времени  $t$ , разрывов считается по особым формулам, использующим соотношение (5) на разрыве. Вне разрывов задания дифференциального уравнения во всех встречающихся нам формах равносильны между собой. Поэтому при построении расчетных формул в точках областей гладкости можно исходить из записи закона сохранения в дифференциальной форме, т. е. из дифференциального уравнения

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0.$$

Принципиальная схема одного из вариантов метода характеристик применительно к нашему примеру состоит в следующем. Отметим на оси  $Ox$  точки  $x_m = mh$ . Будем считать для определенности, что начальное условие  $u(x, 0) = \psi(x)$  задается гладкой функцией  $\psi(x)$ . Из каждой точки  $(x_m, 0)$  выпустим характеристику уравнения  $u_t + uu_x = 0$ .

Предположим, чтобы не усложнять изложение, что при заданной функции  $\psi(x)$  можно выбрать столь малое  $\tau$ , что на любом отрезке времени длины  $\tau$  каждая характеристика пересекается не более чем с одной из соседних характеристик.

Возьмем такие  $\tau$  и проведем прямые  $t = t_p = p\tau$ . Рассмотрим точки пересечения характеристик, выходящих из точек  $(x_m, 0)$ , с прямой  $t = \tau$  и перенесем в эти точки значения решения  $u(x_m, 0) = \psi(x_m)$  по характеристикам.

Если на участке  $0 \leq t \leq \tau$  никакие две характеристики не пересеклись, то делаем следующий шаг: продолжаем характеристики до пересечения с прямой  $t = 2\tau$  и переносим по характеристикам значения решения в точки пересечения. Если пересечения характеристик за время  $\tau < t \leq 2\tau$  опять не было, то

делаем следующий шаг и так до тех пор, пока на некотором участке  $t_p < t < t_{p+1}$  две характеристики, например выходящие из точек  $(x_m, 0)$  и  $(x_{m+1}, 0)$ , пересекутся (рис. 36). Тогда середину отрезка  $Q_m^{p+1}Q_{m+1}^{p+1}$  будем считать точкой, из которой выходит зарождающийся разрыв. Точки  $Q_m^{p+1}$  и  $Q_{m+1}^{p+1}$  заменяем одной точкой  $Q$ , приписывая ей два значения решения,  $u_{\text{лев}}$  и  $u_{\text{прав}}$ , принимая за эти величины значения

$$u_{\text{лев}} = u(Q_m^p) \quad \text{и} \quad u_{\text{прав}} = u(Q_{m+1}^p).$$

Из точки  $Q$  выпускаем линию разрыва до пересечения с прямой  $t = t_{p+2}$ . Угловой коэффициент разрыва определяем из условия на разрыве

$$\operatorname{tg} \alpha = \frac{u_{\text{лев}} + u_{\text{прав}}}{2}.$$

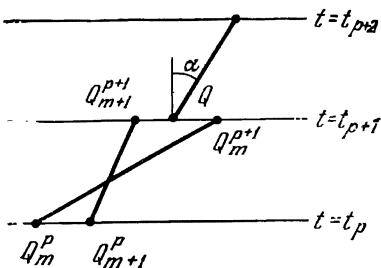


Рис. 36

Из точки пересечения разрыва с прямой  $t = t_{p+2}$  проводим характеристики назад до пересечения с прямой  $t = t_{p+1}$ , проведя их с угловыми коэффициентами  $u_{\text{лев}}$  и  $u_{\text{прав}}$  с предыдущего слоя. В точках пересечения этих характеристик с прямой  $t = t_{p+1}$  с помощью интерполяции по  $x$  находим значения  $u$  и принимаем их за левое и правое значения решения в точке разрыва, лежащей на прямой  $t_{p+2}$ . Это позволяет определить новый наклон разрыва как среднее арифметическое найденных значений слева и справа и продолжить разрыв еще на шаг  $\tau$  по времени.

Достоинство метода характеристик в том, что он позволяет следить за разрывами и адекватно их рассчитывать. Однако в процессе счета возникают все новые разрывы, в частности, малосущественные разрывы могут пересекаться, так что с течением времени картина усложняется. Логика расчета усложняется, требования к машинной памяти и расход машинного времени возрастают.

В этом недостаток метода характеристик, в котором разрывы выделены и считаются нестандартным образом.

**3. Дивергентные разностные схемы.** Разностные схемы, не использующие искусственно введенную вязкость и не использую-

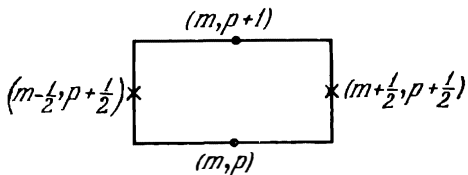


Рис. 37

шие условия на разрыве, как было выяснено, должны опираться на интегральные законы сохранения.

Проведем на плоскости  $Oxt$  сетку прямых  $t = p\tau$ ,  $x = (m + 1/2)h$ ,  $p = 0, 1, \dots$ ;  $m = 0, \pm 1, \dots$ . Отметим на сторонах возникающих прямоугольников их середины (рис. 37) и отнесем их к сетке  $D_h$  (оси координат на рисунке не показаны).

Искомой функцией  $[u]_h$  будем считать сеточную функцию, определенную в точках сетки  $D_h$  путем усреднения решения  $u(x, t)$  по той стороне сеточного прямоугольничка, которому принадлежит рассматриваемая точка сетки:

$$[u]_h \Big|_{\substack{x=x_m \\ t=t_p}} \equiv \bar{u}_m^p = \frac{1}{h} \int_{x_{m-1/2}}^{x_{m+1/2}} u(x, t_p) dx,$$

$$[u]_h \Big|_{\substack{x=x_{m+1/2} \\ t=t_{p+1/2}}} \equiv \bar{U}_{m+1/2}^{p+1/2} = \frac{1}{\tau} \int_{t_p}^{t_{p+1}} u(x_{m+1/2}, t) dt.$$

Приближенное решение  $u^{(h)}$  задачи определено на той же сетке  $D_h$ . Значения  $u^{(h)}$  в точках  $(x_m, t_p)$ , лежащих на горизонтальных сторонах прямоугольничков, будем обозначать  $u_m^p$ , а в точках  $(x_{m+1/2}, t_{p+1/2})$  вертикальных сторон — через  $U_{m+1/2}^{p+1/2}$ .

Величину  $u_m^p$  можно считать продолженной на всю сторону  $t = t_p$ ,  $x_m < x < x_{m+1}$  прямоугольничка, которому принадлежит точка  $(x_m, t_p)$ . Аналогично будем считать, что  $U_{m+1/2}^{p+1/2}$  определена на всем вертикальном промежутке

$$x = x_{m+1/2}, \quad t_p < t < t_{p+1}.$$

Таким образом,  $u^{(h)}$  будет функцией, определенной на прямых  $x = mh$ ,  $t = p\tau$ . Связь между величиной  $u_m^p$  и  $U_{m+1/2}^{p+1/2}$ ,  $p = 0, 1, \dots$ ;  $m = 0, \pm 1, \dots$ , установим, исходя из интегрального закона сохранения

$$\oint_{\Gamma} u \, dx - \frac{u^2}{2} \, dt = 0.$$

Рассмотрим в качестве контура  $\Gamma$  элементарный прямоугольник сетки и положим

$$-\oint_{\Gamma} u^{(h)} \, dx - \frac{(u^{(h)})^2}{2} \, dt = 0, \tag{6}$$

или в развернутом виде

$$h [u_m^{p+1} - u_m^p] + \frac{\tau}{2} [(U_{m+1/2}^{p+1/2})^2 - (U_{m-1/2}^{p+1/2})^2] = 0. \tag{7}$$

Если будет указано правило вычисления величин  $U_{m+1/2}^{p+1/2}$ ,  $m=0, \pm 1, \dots$ , по уже известным величинам  $u_m^p$ ,  $m=0, \pm 1, \dots$ , то формула (7) позволяет вычислить величины  $u_m^{p+1}$ ,  $m=0, \pm 1, \dots$ , т. е. продвинуться на один шаг по времени. Однако независимо от конкретного способа, который мы изберем для вычисления величины  $U_{m+1/2}^{p+1/2}$ , разностная схема вида (7) обладает

*свойством дивергентности*, которое состоит в следующем. Проведем в полуплоскости  $t \geq 0$  какой-либо замкнутый несамопересекающийся контур, целиком состоящий из сторон сеточных прямоугольников (рис. 38). Этот контур  $g_h$  ограничит некоторую область  $G_h$ , составленную из сеточных прямоугольников.

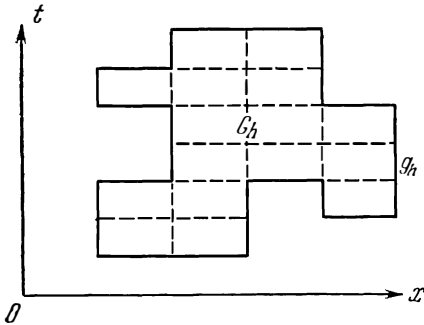


Рис. 38.

Уравнения (7) и (6) отличаются только формой записи. Поэтому можно считать, что мы суммируем уравнения (6). Получим

$$-\oint_{g_h} u^{(h)} dx - \frac{(u^{(h)})^2}{2} dt = 0. \quad (8)$$

Интегралы по тем сторонам прямоугольников, которые не лежат на границе  $g_h$  области  $G_h$ , но входят в выражение (6), после суммирования уравнений (6) взаимно уничтожатся.

Действительно, каждая из этих сторон принадлежит двум соседним прямоугольникам, так что интегрирование функции  $u^{(h)}$  по ней встречается дважды и ведется в противоположных направлениях (рис. 39).

Разностные схемы, при суммировании которых по точкам сеточной области  $G_h$  остаются только алгебраические суммы значений неизвестных или функций от них вдоль границы области, называют *дивергентными*, или *консервативными*. Такие схемы аналогичны дифференциальным уравнениям дивергентного вида

$$\operatorname{div} \Phi = \frac{\partial \Phi_1}{\partial t} + \frac{\partial \Phi_2}{\partial x} = 0,$$

при почленном интегрировании которых по двумерной области  $D$  в левой части возникает контурный интеграл (3) § 29. Разностная схема (2) недивергентна, схема (7) дивергентна.



Заметим следующее. Пусть сеточная функция  $u^{(h)}$ , удовлетворяющая уравнению (7), при  $h \rightarrow 0$  равномерно сходится к некоторой кусочно-непрерывной функции  $u(x, t)$  во всякой замкнутой области, не содержащей линий разрыва, и пусть  $u^{(h)}$  равномерно по  $h$  ограничена. Тогда  $u(x, t)$  удовлетворяет интегральному закону сохранения

$$\oint_g u \, dx - \frac{u^2}{2} \, dt = 0,$$

где  $g$  — произвольный кусочно-гладкий контур.

Это непосредственно следует из возможности приблизить контур  $g$  контуром  $g_h$ , из равенства (8) и предположенной сходимости \*)  $u^{(h)} \rightarrow u$ .

Чтобы схема (7) приобрела смысл, надо указать способ вычисления величин  $U_m^{p+1/2}$  по величинам  $u_{m+1/2}^p$ . В схеме С. К. Годунова, которую мы используем для иллюстрации понятия дивергентных схем, для этого используется решение следующей задачи о «распаде разрыва». Пусть в начальный момент решение  $u(x, 0)$  задано условиями

$$u(x, 0) = \begin{cases} u_{\text{лев}} & \text{при } x < 0, \\ u_{\text{прав}} & \text{при } x > 0, \end{cases}$$

где  $u_{\text{лев}} = \text{const}$  и  $u_{\text{прав}} = \text{const}$ . Тогда можно найти соответствующее обобщенное решение. Как это делается, мы видели в § 29 при разборе примера  $u_{\text{лев}} = 1$ ,  $u_{\text{прав}} = 2$  и примера  $u_{\text{лев}} = 2$ ,  $u_{\text{прав}} = 1$ . Нам важно знать значение  $U = u(0, t)$  решения  $u(x, t)$  при  $x = 0$ .

Читатель, построив картинки типа рис. 33, 34, изображающие решение  $u(x, t)$ , легко проверит, что на прямой  $x = 0$  решение принимает значения  $u_{\text{лев}}$ ,  $u_{\text{прав}}$  или 0 в зависимости от заданных начальных данных, и выяснит для каждой конкретной пары чисел  $u_{\text{лев}}$  и  $u_{\text{прав}}$  какое именно. Например, при  $u_{\text{лев}} > 0$ ,  $u_{\text{прав}} > 0$  будет  $u(0, t) \equiv u_{\text{лев}}$ , а при  $u_{\text{лев}} < 0$ ,  $u_{\text{прав}} < 0$  будет  $u(0, t) \equiv u_{\text{прав}}$ .

Величину  $U_m^{p+1/2} (= U)$  в схеме (7) будем определять из задачи о распаде разрыва, возникающего на границе  $x = x_{m+1/2}$  каждых двух участков, где заданы постоянные значения  $u_m^p (= u_{\text{лев}})$  и  $u_{m+1}^p (= u_{\text{прав}})$ .

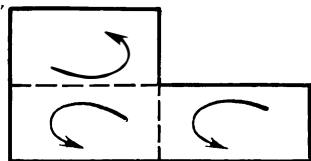


Рис. 39.

\*) Функция  $u = u(x, t)$  определена почти всюду, а функция  $u^{(h)} = u^{(h)}(x, t)$  — лишь на сетке прямых. Это формальное несоответствие можно преодолеть, считая, что при уменьшении  $h$  каждая новая сетка является подразделением старой, и говоря о сходимости в точках сетки, построенной для любого фиксированного  $h$  из числа допустимых.

Если, например,  $u_m^p > 0$ ,  $m = 0, \pm 1, \dots$ , то

$$U_{m+\frac{1}{2}}^{p+\frac{1}{2}} = u_{\text{лев}} = u_m^p, \quad m = 0, \pm 1, \dots,$$

и схема (7) примет вид

$$\frac{u_m^{p+1} - u_m^p}{\tau} + \frac{1}{h} \left[ \frac{(u_m^p)^2}{2} - \frac{(u_{m-1}^p)^2}{2} \right] = 0,$$

$$u_m^0 = \frac{1}{h} \int_{x_{m-\frac{1}{2}}}^{x_{m+\frac{1}{2}}} \psi(x) dx$$

или

$$\frac{u_m^{p+1} - u_m^p}{\tau} + \left( \frac{u_{m-1}^p + u_m^p}{2} \right) \frac{u_m^p - u_{m-1}^p}{h} = 0.$$

Легко видеть, что при

$$r = \frac{\tau}{h} \leq \frac{1}{\max_m |u_m^p|}$$

имеет место принцип максимума

$$\max_m |u_m^{p+1}| \leq \max_m |u_m^p| \leq \dots \leq \max_m |u_m^0| \leq \max_x |\psi(x)|.$$

Отсюда видно, что при  $\tau = \frac{1}{\max_x |\psi(x)|} h$  можно надеяться, что полученная разностная схема устойчива при некотором разумном выборе норм. Мы не будем фактически указывать эти нормы: экспериментальные расчеты подтверждают, что при измельчении сетки решение  $u^{(h)}$  задачи (7) с кусочно-монотонными и кусочно-гладкими начальными данными  $\psi(x)$  сходится к некоторой функции  $u(x, t)$ , имеющей конечное число разрывов, причем вне любой окрестности разрывов сходимость равномерная.

Схема (7) с вычислением  $U_{m+\frac{1}{2}}^{p+\frac{1}{2}}$  путем использования распада разрыва не является, конечно, единственной дивергентной схемой для задачи (1). Укажем, например, еще простейшую схему, основанную на идее пересчета. Эту идею мы изложили в п. 3 § 22. Для простоты ограничимся случаем  $\psi(x) > 0$ .

Сначала ищем вспомогательные величины  $\bar{u}$  по недивергентной неявной разностной схеме

$$\frac{\bar{u}_m^{p+\frac{1}{2}} - u_m^p}{\tau/2} + u_m^p \frac{\bar{u}_m^{p+\frac{1}{2}} - \bar{u}_{m-1}^{p+\frac{1}{2}}}{h} = 0.$$

Значение коэффициента при  $u_x$  в уравнении  $u_t + uu_x = 0$  заменяем через  $u_m^p$ , а не через  $\bar{u}_m^{p+\frac{1}{2}}$ , чтобы возникающая схема была линейна относительно подлежащих вычислению величин.

Далее полагаем

$$U_{m+1/2}^{p+1/2} = \frac{1}{2} (\bar{u}_m^{p+1/2} + \bar{u}_{m+1}^{p+1/2}) \quad (9)$$

и пользуемся схемой (7), (9).

Получаемая так дивергентная схема на гладком решении имеет второй порядок аппроксимации.

Эвристическое исследование с помощью спектрального признака Неймана при линейризации и замораживании коэффициента указывает на устойчивость при произвольном  $r = \tau/h$ . Проведем это исследование.

В результате линейризации и замораживания коэффициента придем к схеме вида

$$\left. \begin{aligned} \frac{\bar{u}_m^{p+1/2} - u_m^p}{\tau/2} + a \frac{\bar{u}_m^{p+1/2} - \bar{u}_{m-1}^{p+1/2}}{h} &= 0, \\ \frac{u_m^{p+1} - u_m^p}{\tau} + \frac{a}{h} \left[ \frac{\bar{u}_{m+1}^{p+1/2} + \bar{u}_m^{p+1/2}}{2} - \frac{\bar{u}_m^{p+1/2} + \bar{u}_{m-1}^{p+1/2}}{2} \right] &= 0. \end{aligned} \right\}$$

Для решения с начальными данными

$$u_m^p = e^{iam}$$

получим

$$\bar{u}_m^{p+1/2} = \mu e^{iam},$$

где

$$\mu = \frac{1}{1 + a \frac{r}{2} - a \frac{r}{2} e^{-ia}}.$$

Далее,

$$u_m^{p+1} = \lambda e^{iam},$$

где

$$\frac{\lambda - 1}{\tau} + \frac{a\mu}{h} \left( \frac{e^{ia} - e^{-ia}}{2} \right) = 0,$$

$$\lambda(\alpha) = \frac{2 + ar - are^{+ia}}{2 + ar - are^{-ia}}, \quad |\lambda(\alpha)| = 1.$$

## ЗАДАЧИ С ДВУМЯ ПРОСТРАНСТВЕННЫМИ ПЕРЕМЕННЫМИ

### ГЛАВА 10

#### ПОНЯТИЕ О РАЗНОСТНЫХ СХЕМАХ РАСЩЕПЛЕНИЯ

Разностные схемы расщепления — одно из важных средств при расчете решений многомерных нестационарных задач математической физики.

#### § 31. Конструкция схем расщепления

На описательном уровне идею конструкции схем расщепления можно изложить так.

Рассмотрим дифференциальную задачу вида

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= Au, & 0 < t < T, \\ u|_{t=0} & \text{ задано,} \end{aligned} \right\} \quad (1)$$

где  $A$  — некоторый оператор по пространственным переменным, например:

$$Au \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

Значения  $u(x, y, t_{p+1})$  по уже известным значениям  $u(x, y, t_p)$ ,  $t_p = p\tau$ , выразим формулой (где  $E$  — единичный оператор,  $E v = v$ )

$$\begin{aligned} u(x, y, t_p + \tau) &= u(x, y, t_p) + \tau \frac{\partial u}{\partial t} + O(\tau^2) = \\ &= u(x, y, t_p) + \tau Au(x, y, t_p) + O(\tau^2) = (E + \tau A) u(x, y, t_p) + O(\tau^2). \end{aligned}$$

Допустим, что правая часть уравнения (1) имеет вид

$$Au \equiv A_1 u + A_2 u.$$

Тогда расщепим уравнение (1)

$$\frac{\partial u}{\partial t} = A_1 u + A_2 u$$

на следующие два:

$$\left. \begin{aligned} \frac{\partial v}{\partial t} &= A_1 v, & t_p \leq t \leq t_{p+1}, \\ v(x, y, t_p) &= u(x, y, t_p), \end{aligned} \right\} \quad (2)$$

$$\left. \begin{aligned} \frac{\partial w}{\partial t} &= A_2 w, & t_p \leq t \leq t_{p+1}, \\ w(x, y, t_p) &= v(x, y, t_{p+1}). \end{aligned} \right\} \quad (3)$$

Заметим, что

$$w(x, y, t_{p+1}) = u(x, y, t_{p+1}) + O(\tau^2). \quad (4)$$

В самом деле,

$$\begin{aligned} v(x, y, t_{p+1}) &= (E + \tau A_1) v(x, y, t_p) + O(\tau^2) = \\ &= (E + \tau A_1) u(x, y, t_p) + O(\tau^2). \end{aligned}$$

Далее, с учетом последнего равенства имеем

$$\begin{aligned} w(x, y, t_{p+1}) &= (E + \tau A_2) w(x, y, t_p) + O(\tau^2) = \\ &= (E + \tau A_2) v(x, y, t_{p+1}) + O(\tau^2) = \\ &= (E + \tau A_2)(E + \tau A_1) u(x, y, t_p) + O(\tau^2) = \\ &= [E + \tau(A_1 + A_2)] u(x, y, t_p) + \tau^2 A_1 A_2 u(x, y, t_p) + O(\tau^2) = \\ &= (E + \tau A) u(x, y, t_p) + O(\tau^2) = u(x, y, t_{p+1}) + O(\tau^2). \end{aligned}$$

Равенство (4) и дает основание на каждом интервале времени  $t_p \leq t \leq t_{p+1}$  вместо задачи (1) последовательно решать задачи (2) и (3).

Для фактического решения уравнений (2) и (3) формально аппроксимируем эти уравнения какими-либо разностными. Тогда возникает некоторая разностная схема расщепления  $L_h u^{(h)} = f^{(h)}$ , позволяющая в два этапа вычислить  $u^{p+1}$  по уже известному  $u^p$  (первый этап — вычисление  $v^{p+1}$  по заданному  $v^p = u^p$ , а второй — вычисление  $w^{p+1} = u^{p+1}$  по уже вычисленному на первом этапе  $w^p = v^{p+1}$ ).

Высказанные соображения носят эвристический характер. После того как разностная схема расщепления

$$L_h u^{(h)} = f^{(h)} \quad (5)$$

для численного решения задачи (1) построена, надо как-либо проверить ее аппроксимацию и устойчивость.

В случае задачи Коши для двумерного уравнения теплопроводности

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, & 0 < t < T, & \quad -\infty < x, y < \infty, \\ u(x, y, 0) &= \psi(x, y) \end{aligned} \right\} \quad (6)$$

в качестве системы (2), (3) можно взять, например,

$$\left. \begin{aligned} \frac{\partial v}{\partial t} &= \frac{\partial^2 v}{\partial x^2}, & v(x, y, t_p) &= v(x, y, t_p), \\ \frac{\partial w}{\partial t} &= \frac{\partial^2 w}{\partial y^2}, & w(x, y, t_p) &= v(x, y, t_{p+1}). \end{aligned} \right\} \quad (7)$$

Указанное расщепление двумерного уравнения из задачи (6) на два одномерных уравнения (7) можно истолковать как приближенную замену процесса распространения тепла по плоскости  $Oxy$  за время  $t_p \leq t \leq t_{p+1}$  на два процесса. В первом из них, который описывается первым уравнением (7), вводятся (мысленно) теплонепроницаемые перегородки, препятствующие

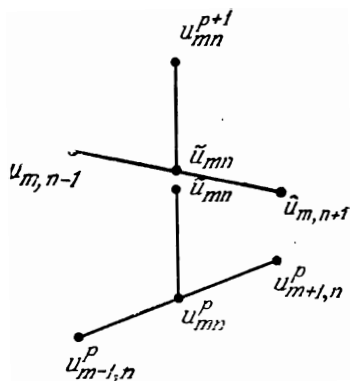


Рис. 40.

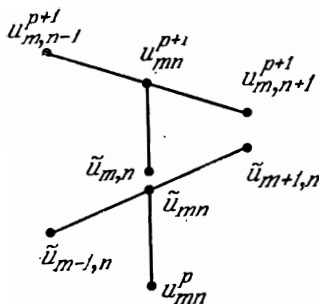


Рис. 41

распространению тепла в направлении оси  $Oy$ . Затем, по прошествии времени  $\tau$ , взамен этих перегородок вводятся перегородки, препятствующие распространению тепла в направлении оси  $Ox$ . Тогда распространение тепла, снова в течение времени  $\tau$ , описывается вторым уравнением.

Выберем сетку  $(x_m, y_n, t_p) = (mh, nh, p\tau)$ .

Разностную схему расщепления, отправляясь от (7), можно построить многими способами. Укажем два из них:

$$\left. \begin{aligned} \frac{\bar{u}_{mn} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^p, \\ \frac{u_{mn}^{p+1} - \bar{u}_{mn}}{\tau} &= \Lambda_{yy} \bar{u}_{mn}, \\ u_{mn}^0 &= \psi(x_m, y_n) \end{aligned} \right\} \quad (8)$$

И

$$\left. \begin{aligned} \frac{\tilde{u}_{mn} - u_{mn}^p}{\tau} &= \Lambda_{xx} \tilde{u}_{mn}, \\ \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} &= \Lambda_{yy} u_{mn}^{p+1}, \\ u_{mn}^0 &= \Psi(x_m, y_n). \end{aligned} \right\} \quad (9)$$

В обеих этих схемах расщепления положим

$$\tilde{u}_{mn} \equiv v_{mn}^{p+1} \equiv \omega_{mn}^p, \quad u_{mn}^{p+1} \equiv \omega_{mn}^{p+1}.$$

Напомним обозначения  $\Lambda_{xx}$  и  $\Lambda_{yy}$ , которые нам уже встречались:

$$\Lambda_{xx} u_{mn} = \frac{u_{m+1, n} - 2u_{mn} + u_{m-1, n}}{h^2},$$

$$\Lambda_{yy} u_{mn} = \frac{u_{m, n+1} - 2u_{mn} + u_{m, n-1}}{h^2}.$$

Схему (8) поясняет рис. 40, схему (9) — рис. 41.

Самое расщепление задачи (6) тоже неединственно.

Задачу (6) можно записать, например, так:

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2} \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + \frac{1}{2} \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), \\ u(x, y, 0) &= \Psi(x, y), \end{aligned} \right\} \quad (6')$$

и поставить ей в соответствие на отрезке  $t_p \leq t \leq t_{p+1}$  следующие две системы:

$$\left. \begin{aligned} \frac{\partial v}{\partial t} &= \frac{1}{2} \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right), \quad t_p \leq t \leq t_{p+1}, \\ v(x, y, t_p) &= u(x, y, t_p) \end{aligned} \right\} \quad (10)$$

И

$$\left. \begin{aligned} \frac{\partial w}{\partial t} &= \frac{1}{2} \left( \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right), \quad t_p \leq t \leq t_{p+1}, \\ w(x, y, t_p) &= v(x, y, t_{p+1}). \end{aligned} \right\} \quad (11)$$

Такое расщепление не есть расщепление по физическим соображениям, как в схеме (7). Разностную схему выберем так (рис. 42):

$$\left. \begin{aligned} \frac{\tilde{u}_{mn} - u_{mn}^p}{\tau} &= \frac{1}{2} (\Lambda_{xx} u_{mn}^p + \Lambda_{yy} \tilde{u}_{mn}), \\ \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} &= \frac{1}{2} (\Lambda_{xx} u_{mn}^{p+1} + \Lambda_{yy} \tilde{u}_{mn}), \\ u_{mn}^0 &= \Psi(x_m, y_n). \end{aligned} \right\} \quad (12)$$

Для вычисления  $u^{p+1}$  по схеме (12) переменных направлений надо сначала при каждом фиксированном  $m$  решить неявное уравнение для  $\tilde{u}_{mn}$ , в которое  $m$  входит как параметр. Потом для вычисления  $u_{mn}^{p+1}$  надо решить второе уравнение (12), неявное относительно  $u_{mn}^{p+1}$ , в которое  $n$  входит как параметр.

Схему (8) можно записать в виде (5), если положить

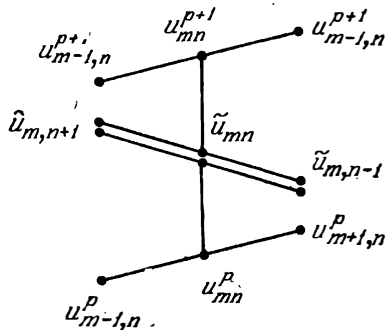


Рис. 42.

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} - \Lambda_{yy} \tilde{u}_{mn}, \\ u_{mn}^0, \end{cases}$$

где  $\tilde{u}_{mn} = u_{mn}^p + \tau \Lambda_{xx} u_{mn}^p$  определяется из первого уравнения (8). Тогда

$$f^{(h)} \equiv \begin{cases} 0 & m, n = 0, \pm 1, \dots; \quad p = 0, 1, \dots, [T/\tau] - 1 \\ \psi(x_m, y_n), & m, n = 0, \pm 1, \dots \end{cases}$$

Предоставим читателю записать схемы (9) и (12) в виде (5).

Читатель может проверить, что спектральный признак устойчивости Неймана, состоящий в ограниченности решений вида  $u_{mn}^p = \lambda^p e^{i(am + \beta n)}$ , выполнен для схемы (8) при  $r = \tau/h^2 \leq 1/2$ , а для схем (9) и (12) при любом  $r$ . Мы не будем останавливаться на исследовании условий устойчивости и доказательстве аппроксимации схем (8), (9) и (12).

### ЗАДАЧИ

1. Исследовать, при каких  $r = \tau/h^2$  выполнен спектральный признак Неймана для разностных схем расщепления (8), (9) и (12), приведенных в этом параграфе.
2. Проверить, что схема (8) аппроксимирует задачу (6) на достаточно гладком ограниченном решении  $u(x, y, t)$ .
3. То же, что и в задаче 2, но для разностных схем расщепления (9) и (12).

### § 32. Экономичные разностные схемы

Рассмотрим и исследуем примеры разностных схем расщепления для задачи о распространении тепла

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, & 0 \leq x, y \leq 1, & 0 \leq t \leq T, \\ u(x, y, 0) &= \psi(x, y), & 0 \leq x, y \leq 1, \\ u(x, y, t)|_{\Gamma} &= 0 \end{aligned} \right\} \quad (1)$$



в прямоугольной области  $0 \leq x, y \leq 1$  с границей  $\Gamma$ . Будем пользоваться обычной сеткой  $(x_m, y_n, t_p) = (mh, nh, p\tau)$ ,  $m, n = 1, 2, \dots, N$  и  $h = 1/N$ .

Разностная схема расщепления, которую мы приведем, в некоторых отношениях обладает принципиальными преимуществами перед простейшей явной

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^p + \Lambda_{yy} u_{mn}^p, \\ u_{mn}^0 &= \psi(x_m, y_n), \\ u_{mn}^p|_{\Gamma} &= 0 \end{aligned} \right\} \quad (2)$$

и простейшей неявной

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^{p+1} + \Lambda_{yy} u_{mn}^{p+1}, \\ u_{mn}^0 &= \psi(x_m, y_n), \\ u_{mn}^p|_{\Gamma} &= 0 \end{aligned} \right\} \quad (3)$$

разностными схемами.

Вычисления по явной схеме (2) очень просты. Для перехода от уже известного  $u^p$  к неизвестному  $u^{p+1} = \{u_{mn}^{p+1}\}$  требуется проделать арифметические действия в количестве, пропорциональном числу  $(N-1)^2$  неизвестных значений  $\{u_{mn}^{p+1}\}$ . В этом смысле явная схема неупрощаема. Разностные схемы, в которых число арифметических действий для перехода от  $u^p$  к  $u^{p+1} = \{u_{mn}^{p+1}\}$  пропорционально числу неизвестных значений, называются *экономичными*. Однако, будучи экономичной, явная схема устойчива лишь при жестком ограничении  $\tau \leq h^2/4$  на шаг  $\tau$  сетки. Приведенная выше простейшая неявная разностная схема (3), как мы уже знаем (§ 27, п. 3), абсолютно устойчива. Но она далеко не является экономичной. Для неизвестных  $\{u_{mn}^{p+1}\}$  приходится решать сложную (нерасщепляющуюся) систему линейных уравнений. Как известно из алгебры, для этого требуется произвести арифметические действия в количестве, пропорциональном не первой степени числа неизвестных, как в экономичных схемах, но кубу числа неизвестных, если пользоваться каким-либо методом исключения неизвестных.

Заметим, что сейчас ведутся поиски более экономичных способов точного решения общих линейных систем. Штрассеном указан алгоритм, требующий числа действий, пропорционального не третьей, а  $\lg_2 7$  степени числа неизвестных.

Разностная схема расщепления, которую мы построим, является экономичной и безусловно устойчивой, т. е. соединяет достоинства явной (2) и неявной (3) схем.

Относительно решения  $u(x, y, t)$  задачи (1) мы будем предполагать, что оно имеет непрерывные вплоть до границы  $\Gamma$  производные всех порядков, которые по ходу дела потребуются. Отметим, что на границе  $\Gamma$  все производные по пространственным переменным четного порядка (до того порядка, до которого они существуют и непрерывны) обращаются в нуль:

$$u_{xx}|_{\Gamma} = u_{xxxx}|_{\Gamma} = u_{xxyy}|_{\Gamma} = 0. \quad (4)$$

Так, на стороне  $x = 0$  границы  $\Gamma$  квадрата  $0 \leq x, y \leq 1$  обращаются в нуль  $\partial u / \partial t$  и  $\partial^2 u / \partial y^2$ . Поэтому в силу уравнения  $u_t = u_{xx} + u_{yy}$  также  $u_{xx} = 0$ . Дифференцируя уравнение по  $y$  дважды, получим

$$\frac{\partial u_{yy}}{\partial t} = u_{xxyy} + u_{yyyy}.$$

Но на стороне  $x = 0$  границы  $\Gamma$  будет

$$u_{yy} = 0, \quad u_{yyyy} = 0, \quad \frac{\partial u_{yy}}{\partial t} = 0,$$

а значит, в силу дифференциального уравнения также  $u_{xxyy} = 0$ .

Переходим к построению разностной схемы расщепления для задачи (1).

Задаче (1) на отрезке  $t_p \leq t \leq t_{p+1}$  поставим в соответствие две задачи:

$$\left. \begin{aligned} \frac{\partial v}{\partial t} &= \frac{\partial^2 v}{\partial x^2}, \\ v|_{\Gamma} &= 0, \quad v(x, y, t_p) = u(x, y, t_p), \end{aligned} \right\} \quad (5)$$

$$\left. \begin{aligned} \frac{\partial w}{\partial t} &= \frac{\partial^2 w}{\partial y^2}, \\ w|_{\Gamma} &= 0, \quad w(x, y, t_p) = v(x, y, t_{p+1}). \end{aligned} \right\} \quad (6)$$

Сеточную функцию

$$u^{(h)} = \{u_{mn}^p\}, \quad u_{mn}^0 = \psi(x_m, y_p), \quad u_{mn}^p|_{\Gamma} = 0$$

будем определять последовательно при  $p = 1, 2, \dots, [T/\tau]$  из уравнений

$$\left. \begin{aligned} \frac{\tilde{u}_{mn} - u_{mn}^p}{\tau} &= \Lambda_{xx} \tilde{u}_{mn}, \quad m, n = 1, 2, \dots, N-1 \\ \tilde{u}_{mn}|_{\Gamma} &= 0; \end{aligned} \right\} \quad (7)$$

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} &= \Lambda_{yy} u_{mn}^{p+1}, \quad m, n = 1, 2, \dots, N-1, \\ u_{mn}^{p+1}|_{\Gamma} &= 0. \end{aligned} \right\} \quad (8)$$

Задача (7) аналогична задаче (5), а задача (8) — задаче (6). При этом

$$v_{mn}^p = u_{mn}^p, \quad \omega_{mn}^p = v_{mn}^{p+1} = \tilde{u}_{mn}, \quad \omega_{mn}^{p+1} = u_{mn}^{p+1}.$$

В соответствии с разностной схемой расщепления (7), (8) сначала по известным значениям  $u^p = \{u_{mn}^p\}$  вычисляется вспомогательная функция  $\tilde{u}_{mn}$ , а потом из (8) вычисляется  $u^{p+1} = \{u_{mn}^{p+1}\}$ .

Заметим, что разностная задача (7) для отыскания  $\tilde{u}_{mn}$  при каждом фиксированном  $n$ ,  $n = 1, \dots, N-1$ , в точности совпадает с неявной разностной схемой

$$\frac{v_{mn}^{p+1} - v_{mn}^p}{\tau} = \Lambda_{xx} v_{mn}^{p+1}$$

для одномерного уравнения теплопроводности (5) на отрезке,  $0 \leq x \leq 1$ , в которое  $y$  входит только как параметр.

Разностная задача (7) при каждом фиксированном  $n$  решается прогонкой в направлении оси  $Ox$ .

Точно так же разностная задача (8) при каждом фиксированном  $m$  решается прогонкой по направлению оси  $Oy$ . Подчеркнем, что в силу свойств алгоритма прогонки общее число арифметических действий для вычисления  $u^{p+1} = \{u_{mn}^{p+1}\}$  оказывается пропорционально числу неизвестных значений, т. е. разностная схема (7), (8) является экономичной.

Для точного формулирования понятия аппроксимации и устойчивости запишем разностную схему (7), (8) в принятом нами на протяжении всей книги виде

$$L_h u^{(h)} = f^{(h)}. \quad (9)$$

Для этого положим

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} - \Lambda_{yy} u_{mn}^{p+1}, & m, n = 1, 2, \dots, N-1, \\ u_{mn}^{p+1} |_{\Gamma}, \\ u_{mn}^0, \end{cases} \quad (10)$$

где  $\tilde{u}_{mn}$  — решение вспомогательной задачи,

$$\left. \begin{aligned} \frac{\tilde{u}_{mn} - u_{mn}^p}{\tau} &= \Lambda_{xx} \tilde{u}_{mn}, & m, n = 1, 2, \dots, N-1, \\ \tilde{u}_{mn} |_{\Gamma} &= 0. \end{aligned} \right\} \quad (11)$$

За  $f^{(h)}$  надо принять в таком случае

$$f^{(h)} \equiv \begin{cases} 0, & m, n = 1, 2, \dots, N-1, \\ 0, & (x_m, y_n) \in \Gamma, \\ \psi(x_m, y_n). \end{cases} \quad (12)$$

В качестве нормы в  $U_h$  примем

$$\|u^{(h)}\|_{U_h} = \max_{m, n, \rho} |u_{mn}^\rho|.$$

К пространству  $F_h$  отнесем элементы  $g^{(h)}$  вида

$$g^{(h)} = \begin{cases} \varphi_{mn}^\rho, \\ 0, \\ \psi_{mn} \end{cases}$$

и норму в  $F_h$  определим равенством

$$\|g^{(h)}\|_{F_h} = \max_{m, n, \rho} |\varphi_{mn}^\rho| + \max_{m, n} |\psi_{mn}|.$$

Сначала докажем безусловную устойчивость разностной схемы (9), задаваемой уравнениями (10), (12), а аппроксимацию докажем позже. Ввиду линейности разностной схемы (9) для доказательства устойчивости надо показать, что задача  $L_h z^{(h)} =$

$= g^{(h)}$  имеет решение при произвольном  $g^{(h)} = \begin{Bmatrix} \varphi_{mn}^\rho \\ 0 \\ \psi_{mn} \end{Bmatrix} \in F_h$ , причем

$$\|z^{(h)}\|_{U_h} \leq c \|g^{(h)}\|_{F_h},$$

где  $c$  не зависит от  $h$ .

Запишем задачу  $L_h z^{(h)} = g^{(h)}$  в развернутом виде:

$$\left. \begin{aligned} \frac{z_{mn}^{\rho+1} - \tilde{z}_{mn}}{\tau} - \Lambda_{yy} z_{mn}^{\rho+1} &= \varphi_{mn}^{\rho+1}, \quad m, n = 1, 2, \dots, N-1, \\ z_{0n}^{\rho+1} &= z_{Nn}^{\rho+1} = 0, \end{aligned} \right\} \quad (13)$$

где  $\tilde{z}_{mn}$  — решение вспомогательной задачи

$$\left. \begin{aligned} \frac{\tilde{z}_{mn} - z_{mn}^\rho}{\tau} &= \Lambda_{xx} \tilde{z}_{mn}, \quad m, n = 1, 2, \dots, N-1, \\ \tilde{z}_{m,0} &= \tilde{z}_{m,N} = 0, \end{aligned} \right\} \quad (14)$$

причем

$$z_{mn}^0 = \psi_{mn}.$$

В силу принципа максимума, доказанного в § 28 для одномерной неявной разностной схемы, аппроксимирующей одно-

мерное уравнение теплопроводности, из уравнений (13) следует, что

$$\max_{m, n} |z_{mn}^{\rho+1}| \leq \max_{m, n} |\tilde{z}_{mn}| + \tau \max_{m, n, \rho} |\varphi_{mn}^{\rho}|.$$

Но из (14) следует, что

$$\max_{m, n} |\tilde{z}_{mn}| \leq \max_{m, n} |z_{mn}^{\rho}|.$$

Поэтому

$$\max_{m, n} |z_{mn}^{\rho+1}| \leq \max_{m, n} |z_{mn}^{\rho}| + \tau \max_{m, n, \rho} |\varphi_{mn}^{\rho}|.$$

Отсюда

$$\begin{aligned} \max_{m, n} |z_{mn}^{\rho+1}| &\leq \max_{m, n} |z_{mn}^{\rho-1}| + 2\tau \max_{m, n} |\varphi_{mn}^{\rho}| \leq \\ &\dots \dots \dots \leq \max_{m, n} |\psi_{mn}| + T \max_{m, n, \rho} |\varphi_{mn}^{\rho}| \leq \\ &\leq (1 + T) (\max_{m, n} |\psi_{mn}| + \max_{m, n, \rho} |\varphi_{mn}^{\rho}|) = (1 + T) \|g^{(h)}\|_{F_h}. \end{aligned}$$

Выписанное неравенство

$$\max_{m, n} |z_{mn}^{\rho+1}| \leq (1 + T) \|g^{(h)}\|_{F_h}$$

справедливо при любом  $\rho$ , поэтому

$$\|z^{(h)}\|_{U_h} \leq (1 + T) \|g^{(h)}\|_{F_h}$$

при произвольном соотношении шагов  $\tau$  и  $h$ . Это и означает, что рассматриваемая схема расщепления безусловно устойчива.

Перейдем к проверке аппроксимации. Будем предполагать, как обычно, что задача (1) имеет достаточно гладкое решение  $u(x, y, t)$ . Вычислим невязку  $\Delta f^{(h)}$ ,  $L_h[u]_h = f^{(h)} + \Delta f^{(h)}$ , возникающую при подстановке  $[u]_h$  в левую часть уравнения (9), и покажем, что  $\|\Delta f^{(h)}\|_{F_h} = O(\tau + h^2)$ .

В соответствии с тем, как мы определили оператор  $L_h$ , будет

$$L_h[u]_h = \begin{cases} \frac{u(x_m, y_n, t_{\rho+1}) - \tilde{u}_{mn}}{\tau} - \Lambda_{yy}u(x_m, y_n, t_{\rho+1}), \\ \quad m, n = 1, \dots, N-1, \\ 0 \text{ в точках } \Gamma, \\ u(x_m, y_n, 0), \quad m, n = 1, \dots, N-1, \end{cases} \quad (15)$$

где  $\tilde{u}_{mn}$  — решение вспомогательной задачи

$$\left. \begin{aligned} \frac{\tilde{u}_{mn} - u(x_m, y_n, t_p)}{\tau} - \Lambda_{xx}\tilde{u}_{mn} &= 0, \quad m, n = 1, 2, \dots, N-1, \\ \tilde{u}_{mn}|_{\Gamma} &= 0. \end{aligned} \right\} (16)$$

Решение  $\tilde{u}_{mn}$  вспомогательной задачи (16), как мы покажем ниже, имеет вид

$$\left. \begin{aligned} \tilde{u}_{mn} &= u(x_m, y_n, t_p) + \tau\Lambda_{xx}u(x_m, y_n, t_p) + O(\tau^2), \\ m, n &= 1, 2, \dots, N-1, \\ \tilde{u}_{mn}|_{\Gamma} &= u(x_m, y_n, t_p)|_{\Gamma} = 0. \end{aligned} \right\} (17)$$

Подставляя это выражение  $\tilde{u}_{mn}$  в (15), получим

$$\begin{aligned} \frac{u(x_m, y_n, t_{p+1}) - \tilde{u}_{mn}}{\tau} - \Lambda_{yy}u(x_m, y_n, t_{p+1}) &= \\ = \frac{u(x_m, y_n, t_{p+1}) - [u(x_m, y_n, t_p) + \tau\Lambda_{xx}u(x_m, y_n, t_p) + O(\tau^2)]}{\tau} &= \\ - \Lambda_{yy}u(x_m, y_n, t_{p+1}) = \frac{u(x_m, y_n, t_{p+1}) - u(x_m, y_n, t_p)}{\tau} &= \\ - \Lambda_{xx}u(x_m, y_n, t_p) - \Lambda_{yy}u(x_m, y_n, t_{p+1}) + O(\tau) &= \\ = \left( \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} \right)_{x_m, y_n, t_p} + O(\tau + h^2) &= O(\tau + h^2). \end{aligned}$$

Таким образом,

$$L_h[u]_h \equiv \left\{ \begin{array}{l} 0 \quad + O(\tau + h^2) \\ 0 \quad + 0 \\ \psi(x_m, y_n) + 0 \end{array} \right\} = f^{(h)} + \Delta f^{(h)},$$

где

$$\Delta f^{(h)} = \left\{ \begin{array}{ll} O(\tau + h^2) & \text{в точках } (x_m, y_n, t_p), \\ 0, & p = 0, 1, \dots, [T/\tau] - 1, \\ 0, & m, n = 1, 2, \dots, N-1. \end{array} \right.$$

Отсюда

$$\|\Delta f^{(h)}\|_{F_h} = O(\tau + h^2).$$

Осталось доказать приближенное представление (17) для решения  $\tilde{u}_{mn}$  задачи (16).

Сначала выскажем эвристические соображения, подсказавшие представление (17). Ясно, что при малых  $\tau$  выполняется приближенное равенство

$$\tilde{u}_{mn} \approx u(x_m, y_n, t_p).$$

При замене на этом основании в уравнении (16) выражения  $\Lambda_{xx}\tilde{u}_{mn}$  выражением  $\Lambda_{xx}u(x_m, y_n, t_p)$  возникло бы уравнение

$$\frac{\tilde{u} - u}{\tau} - \Lambda_{xx}u = 0,$$

из которого следует равенство  $\tilde{u} = u + \tau\Lambda_{xx}u$ , лишь остаточным членом  $O(\tau^2)$  отличающееся от (17). Переходим к доказательству справедливости (17).

Доопределим  $\Lambda_{xx}u(x_m, y_n, t_p)$  в точках  $\Gamma$ , положив  $\Lambda_{xx}u|_{\Gamma} = 0$ . Подставим

$$w_{mn} \equiv u(x_m, y_n, t_p) + \tau\Lambda_{xx}u(x_m, y_n, t_p)$$

вместо  $\tilde{u}_{mn}$  в уравнение (16). Получим

$$\begin{aligned} \frac{w_{mn} - u(x_m, y_n, t_p)}{\tau} - \Lambda_{xx}w_{mn} &\equiv \\ &\equiv \frac{\{u(x_m, y_n, t_p) + \tau\Lambda_{xx}u(x_m, y_n, t_p)\} - u(x_m, y_n, t_p)}{\tau} - \\ &- \Lambda_{xx}u(x_m, y_n, t_p) - \tau\Lambda_{xx}\Lambda_{xx}u(x_m, y_n, t_p) = \\ &= -\tau\Lambda_{xx}\Lambda_{xx}u(x_m, y_n, t_p). \end{aligned}$$

В предположении, что  $\partial^4 u / \partial x^4$  непрерывна и ограничена, и учитывая, что  $\frac{\partial^2 u}{\partial x^2} \Big|_{\Gamma} = 0$ , легко видеть, что  $\Lambda_{xx}\Lambda_{xx}u(x_m, y_n, t_p)$  ограничено. Поэтому

$$\begin{aligned} \frac{w_{mn} - u(x_m, y_n, t_p)}{\tau} - \Lambda_{xx}w_{mn} &= O(\tau), \\ w_{mn} |_{\Gamma} &= 0. \end{aligned}$$

Вычитая из этих уравнений уравнения (16) почленно, для разности  $z_{mn} = w_{mn} - \tilde{u}_{mn}$  получим

$$\left. \begin{aligned} z_{mn} - \tau\Lambda_{xx}z_{mn} &= O(\tau^2), \\ z_{mn} |_{\Gamma} &= 0, \end{aligned} \right\} \quad (18)$$

или в развернутом виде

$$\left. \begin{aligned} rz_{m-1, n} - 2\left(r + \frac{1}{2}\right)z_{mn} + rz_{m+1, n} &= O(\tau^2), \\ m, n &= 1, \dots, N-1, \\ z_{0n} = z_{Mn} &= 0, \quad r = \tau/h^2. \end{aligned} \right\} \quad (18')$$

Но эта задача для  $(z_{mn})$  имеет вид

$$\begin{aligned} a_m u_{m-1} + b_m u_m + c_m u_{m+1} &= g_m, \quad m = 1, \dots, N-1, \\ u_0 &= u_N = 0, \\ a_m > 0, \quad c_m > 0, \quad |b_m| &\geq a_m + c_m + \delta, \quad \delta = 1. \end{aligned}$$

В § 4 было доказано, что в таком случае

$$\max |u_m| \leq c \max |g_m|,$$

где  $c$  зависит только от  $\delta$ . Отсюда  $z_{mn} = O(\tau^2)$ , т. е.

$$\begin{aligned} \tilde{u}_{,mn} &= w_{mn} - z_{mn} = w_{mn} + O(\tau^2) = \\ &= u(x_m, y_n, t_p) + \tau \Lambda_{xx} u(x_m, y_n, t_p) + O(\tau^2), \end{aligned}$$

что совпадает с представлением (17), которое мы доказываем.

### ЗАДАЧИ

1. Для дифференциальной краевой задачи (1) о распространении тепла в квадратной области  $0 \leq x, y \leq 1$  предложить и исследовать разностную схему расщепления, аналогичную явной схеме расщепления (8) из § 31 для задачи Коши.

2. Для дифференциальной краевой задачи (1) предложить разностную схему, аналогичную схеме переменных направлений (12) из § 31. Доказать, что имеет место аппроксимация порядка  $O(\tau + h^2)$ .

3. Предложить разностную схему для решения задачи

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \quad 0 \leq t \leq T, \quad (x, y) \in D, \\ u(x, y, t)|_{\Gamma} &= \psi(x, y, t) \quad u(x, y, 0) = \psi(x, y) \end{aligned} \right\}$$

в случае криволинейной области  $D$  по аналогии с разностной схемой расщепления, рассмотренной для задачи (1) в тексте параграфа.

### § 33. Расщепление по физическим факторам

Идея расщепления используется не только для получения экономичных абсолютно устойчивых схем. Иногда производится расщепление сложной задачи на более простые, чтобы на каждом малом интервале времени  $t_p < t < t_{p+1}$  разделить во времени действие различных факторов, влияющих на процесс. Для возникающих при этом сравнительно простых задач легче построить адекватные им схемы, составляющие в совокупности разумную разностную схему расщепления для всей задачи.

В качестве примера укажем метод крупных частиц О. М. Белоцерковского и Ю. М. Давыдова (ЖВМ и МФ 11, № 1, 1971), предназначенный для расчета течений газа при сильной деформации вещества и больших колебаниях плотности. Этот метод, как и метод Харлоу частиц в ячейках, по замечанию Н. Н. Яненко, можно трактовать как некоторую разностную



схему расщепления для уравнений газовой динамики. Все вещество разбивается сеткой неподвижных прямых (речь идет о двумерной задаче) на ячейки. Вещество, попавшее в ячейку в момент  $t_p$ , и есть крупная частица. Ей приписываются импульс и полная энергия. Затем строится разностная схема, моделирующая изменение скоростей, импульсов и полной энергии крупных частиц под влиянием одного только давления, без учета тех членов системы уравнений газовой динамики, которые описывают перенос вещества, импульса и энергии. Это — первый шаг разностной схемы расщепления. На втором шаге производится пересчет полученных на первом шаге промежуточных величин по разностной схеме, учитывающей остальные члены уравнений газовой динамики, т. е. только перетекание вещества из ячейки в соседние ячейки и соответствующий перенос импульса и энергии. Так получают крупные частицы и соответствующие им импульс и энергия на момент времени  $t_{p+1} = t_p + \tau$ .

ЭЛЛИПТИЧЕСКИЕ ЗАДАЧИ

§ 34. Простейшая разностная схема для задачи Дирихле

Здесь мы проверим, что простейшая разностная схема (13) § 24 аппроксимирует задачу Дирихле (12) § 24 со вторым относительно  $h$  порядком и устойчива, а следовательно, применима для приближенного вычисления решения задачи Дирихле.

Задачу Дирихле для уравнения Пуассона в квадратной области  $D = (0 \leq x, y \leq 1)$  с границей  $\Gamma$  запишем в виде

$$\left. \begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= \varphi(x, y), & 0 \leq x, y \leq 1, \\ u|_{\Gamma} &= \psi(s), \end{aligned} \right\} \quad (1)$$

где  $s$  — длина дуги, отсчитываемая вдоль границы  $\Gamma$ , функции  $\varphi(x, y)$  и  $\psi(s)$  заданы.

Совокупность точек  $(x_m, y_n) = (mh, nh)$  сетки  $h = 1/M, M$  — целое, попавших внутрь квадрата или на его границу, обозначим через  $D_h$ . Точки  $D_h$ , лежащие строго внутри квадрата  $D$ , будем считать внутренними точками сеточного квадрата  $D_h$ ; совокупность внутренних точек обозначим  $D_h^0$ . Точки  $D_h$ , лежащие на границе  $\Gamma$  квадрата  $D$ , будем считать граничными точками сеточной области  $D_h$ ; совокупность граничных точек обозначим  $\Gamma_h$ . Разностную схему (13) § 24

$$L_h u^{(h)} = f^{(h)} \quad (2)$$

запишем в виде

$$L_h u^{(h)} \equiv \begin{cases} \Lambda_{xx} u_{mn} + \Lambda_{yy} u_{mn} = \varphi(x_m, y_n), & (x_m, y_n) \in D_h^0, \\ u_{mn} = \psi(s_{mn}), & (x_m, y_n) \in \Gamma_h, \end{cases} \quad (3)$$

где  $\psi(s_{mn})$  — значение функции  $\psi(s)$  в точке  $(x_m, y_n)$ , принадлежащей  $\Gamma_h$ .

**1. Аппроксимация.** Правая часть  $f^{(h)}$  разностной схемы (2) имеет вид

$$f^{(h)} = \begin{cases} \varphi(x_m, y_n), & (x_m, y_n) \in D_h^0, \\ \psi(s_{mn}), & (x_m, y_n) \in \Gamma_h. \end{cases} \quad (4)$$

В предположении, что решение  $u(x, y)$  задачи (1) имеет ограниченные четвертые производные, с помощью формулы Тейлора устанавливается равенство

$$\begin{aligned} \Lambda_h u &\equiv \Lambda_{xx} u + \Lambda_{yy} u = \\ &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{h^2}{24} \left( \frac{\partial^4 u(x + \xi h, y)}{\partial x^4} + \frac{\partial^4 u(x, y + \eta h)}{\partial y^4} \right). \end{aligned} \quad (5)$$

Поэтому для решения  $u(x, y)$  задачи (1) имеем

$$L_h[u]_h = \begin{cases} \varphi(x_m, y_n) + O(h^2), & (x_m, y_n) \in D_h^0, \\ \psi(s_{mn}) + 0, & (x_m, y_n) \in \Gamma_h. \end{cases} \quad (6)$$

Таким образом, невязка  $\delta f^{(h)}$ , возникающая при подстановке  $[u]_h$  в левую часть разностной схемы (2), имеет вид

$$\delta f^{(h)} = \begin{cases} O(h^2), & (x_m, y_n) \in D_h^0, \\ 0, & (x_m, y_n) \in \Gamma_h. \end{cases} \quad (7)$$

В пространстве  $F_h$ , состоящем из элементов вида

$$f^{(h)} = \begin{cases} \varphi_{mn}, & (x_m, y_n) \in D_h^0, \\ \psi_{mn}, & (x_m, y_n) \in \Gamma_h, \end{cases}$$

введем норму

$$\|f^{(h)}\|_{F_h} = \max_{(mh, nh) \in D_h^0} |\varphi_{mn}| + \max_{(mh, nh) \in \Gamma_h} |\psi_{mn}|. \quad (8)$$

Тогда

$$\|\delta f^{(h)}\|_{F_h} = O(h^2).$$

Таким образом, разностная краевая задача (3) аппроксимирует задачу Дирихле (1) со вторым порядком относительно  $h$ .

**2. Устойчивость.** Определим норму в пространстве  $U_h$  функций, заданных на сетке  $D_h$ , положив

$$\|u^{(h)}\|_{U_h} = \max_{(mh, nh) \in D_h} |u_{mn}|. \quad (9)$$

Для доказательства устойчивости разностной схемы (3), к которому мы переходим, в соответствии с определением устойчивости надо установить, что задача (2) однозначно разрешима при произвольной правой части  $f^{(h)}$  (это свойство не зависит от выбора норм) и что выполнена оценка вида

$$\|u^{(h)}\|_{U_h} \leq c \|f^{(h)}\|_{F_h}, \quad (10)$$

где  $c$  не зависит ни от  $h$ , ни от  $f^{(h)}$ .

Лемма 1. Пусть функция  $v^{(h)} = \{v_{mn}\}$  определена на сетке  $D_h$  и во внутренних точках  $(x_m, y_n) = (mh, nh) \in D_h^0$  удовлетворяет условию

$$\Lambda_h v^{(h)}|_{(mh, nh)} \geq 0, \quad (mh, nh) \in D_h^0. \quad (11)$$

Тогда наибольшее на сетке  $D_h$  значение  $v^{(h)}$  достигается хотя бы в одной точке  $\Gamma_h$ .

Доказательство. Допустим противное. Выберем среди точек сетки  $D_h$ , в которых  $v^{(h)}$  достигает своего наибольшего значения, какую-нибудь одну точку  $(x_m, y_n)$ , имеющую самую большую абсциссу. По нашему предположению  $(x_m, y_n)$  — внутренняя точка, причем  $v_{mn}$  строго больше, чем  $v_{m+1, n}$ . В точке  $(mh, nh)$  будет

$$\begin{aligned} \Lambda_h v^{(h)}|_{(mh, nh)} &\equiv \\ &\equiv \frac{(v_{m+1, n} - v_{mn}) + (v_{m, n+1} - v_{mn}) + (v_{m-1, n} - v_{mn}) + (v_{m, n-1} - v_{mn})}{h^2} < 0, \end{aligned}$$

поскольку первая скобка в числителе отрицательна, а остальные неположительны. Противоречие с (11).

Лемма 2. Пусть функция  $v^{(h)} = \{v_{mn}\}$  определена на сетке  $D_h$  и во внутренних точках  $(mh, nh) \in D_h^0$  удовлетворяет условию

$$\Lambda_h v^{(h)}|_{(mh, nh)} \leq 0, \quad (mh, nh) \in D_h^0. \quad (12)$$

Тогда наименьшее на сетке  $D_h$  значение  $v^{(h)}$  достигается хотя бы в одной точке границы.

Лемма 2 доказывается аналогично лемме 1.

Теорема (принцип максимума). Каждое решение разностного уравнения

$$\Lambda_h v^{(h)}|_{(mh, nh)} = 0, \quad (mh, nh) \in D_h^0, \quad (13)$$

достигает своего наибольшего и наименьшего значения в некоторых точках  $\Gamma_h$ .

Доказательство получается объединением утверждений лемм 1 и 2.

Свойство решений разностного уравнения (13) аналогично свойству решений  $v(x, y)$  уравнения Лапласа  $v_{xx} + v_{yy} = 0$  принимать наибольшее и наименьшее значения на границе области, где эти решения определены.

Из принципа максимума следует, что задача

$$L_h u^{(h)} = \begin{cases} \Lambda_h u^{(h)}|_{(mh, nh)} = 0, & (mh, nh) \in D_h^0, \\ u^{(h)}|_{(mh, nh)} = 0, & (mh, nh) \in \Gamma_h, \end{cases} \quad (14)$$

имеет только нулевое решение  $u^{(h)} = 0$ , поскольку наибольшее и наименьшее значения этого решения принимаются в точках  $\Gamma_h$ , где  $u_{mn} = 0$ . Следовательно, определитель системы линейных уравнений (3) отличен от нуля и разностная краевая задача (2) однозначно разрешима при произвольной правой части  $f^{(h)}$ .

Переходим к доказательству оценки (10).

В силу формулы (5) для произвольного многочлена  $P(x, y)$  второй (и даже третьей) степени

$$P(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$$

выполнено равенство

$$\Lambda_h P = \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2}, \quad (15)$$

так как четвертые производные от  $P(x, y)$ , входящие в выражение остаточного члена формулы (5), обращаются в нуль.

Используя функции  $\varphi_{mn}$  и  $\psi_{mn}$  из правой части системы (3) и фиксировав  $R > \sqrt{2}$ , построим вспомогательную функцию

$$P^{(h)}(x, y) = \frac{1}{4} [R^2 - (x^2 + y^2)] \max_{(mh, nh) \in D_h^0} |\varphi_{mn}| + \max_{(mh, nh) \in \Gamma_h} |\psi_{mn}|,$$

которую будем рассматривать только в точках сетки  $D_h$ . Это отражено значком  $h$  в обозначении  $P^{(h)}(x, y)$ . В силу (15) всюду в точках  $D_h^0$

$$\Lambda_h P^{(h)} \Big|_{\substack{x=mh \\ y=nh}} \equiv - \max_{(rh, sh) \in D_h^0} |\varphi_{rs}|, \quad (mh, nh) \in D_h^0.$$

Поэтому разность решения  $u^{(h)}$  задачи (3) и функции  $P^{(h)}$  удовлетворяет в точках  $D_h^0$  равенствам

$$\Lambda_h (u^{(h)} - P^{(h)}) = \Lambda_h u^{(h)} - \Lambda_h P^{(h)} = \varphi_{mn} + \max_{r, s} |\varphi_{rs}| \geq 0.$$

В силу леммы 1 разность  $u^{(h)} - P^{(h)}$  принимает свое наибольшее значение на границе  $\Gamma_h$ . Но на границе  $\Gamma_h$  эта разность

$$\begin{aligned} u^{(h)} \Big|_{\Gamma_h} - P^{(h)} \Big|_{\Gamma_h} &= [\psi_{mn} - P^{(h)}]_{mn} = \\ &= [\psi_{mn} - \max_{(rh, sh) \in \Gamma_h} |\varphi_{rs}|] + \frac{1}{4} [(x^2 + y^2) - R^2] \max_{(rh, sh) \in D_h^0} |\varphi_{rs}| \end{aligned}$$

неположительна, так как в квадрате  $D$  всюду  $x^2 + y^2 \leq R^2$  и обе квадратные скобки в правой части неположительны. Поскольку наибольшее значение  $u^{(h)} - P^{(h)}$  неположительно, то всюду на  $D_h$

$$u^{(h)} - P^{(h)} \Big|_{(mh, nh)} \leq 0, \quad \text{или} \quad u^{(h)} \leq P^{(h)}.$$

Аналогично, для функции  $u^{(h)} + P^{(h)}$  в точках  $D_h^0$

$$\Lambda_h(u^{(h)} + P^{(h)}) \leq 0,$$

а в точках  $\Gamma_h$  сумма  $u^{(h)} + P^{(h)}$  неотрицательна. В силу леммы 2 всюду на  $D_h$

$$u^{(h)} + P^{(h)} \geq 0, \quad \text{или} \quad -P^{(h)} \leq u^{(h)}.$$

Таким образом, всюду на  $D_h$

$$|u_{mn}| \leq |P_{mn}^{(h)}| \leq \frac{1}{4} R^2 \max_{(rh, sh) \in D_h^0} |\varphi_{rs}| + \max_{(rh, sh) \in \Gamma_h} |\psi_{rs}|.$$

Отсюда вытекает неравенство (10):

$$\begin{aligned} \max |u_{mn}| = \|u^{(h)}\|_{U_h} &\leq c \|f^{(h)}\|_{F_h} = \\ &= c \left( \max_{(rh, sh) \in D_h^0} |\varphi_{rs}| + \max_{(rh, sh) \in \Gamma_h} |\psi_{rs}| \right), \end{aligned}$$

где

$$c = \max \left[ 1, \frac{R^2}{4} \right],$$

завершающее доказательство устойчивости.

В случае задачи Дирихле для эллиптического уравнения с переменными коэффициентами

$$\begin{aligned} \frac{\partial}{\partial x} \left[ k_1(x, y) \frac{\partial u}{\partial x} \right] + \frac{\partial}{\partial y} \left[ k_2(x, y) \frac{\partial u}{\partial y} \right] &= \varphi(x, y), \quad (x, y) \in D, \\ u|_{\Gamma} &= \psi(s), \end{aligned}$$

где  $k_1(x, y)$  и  $k_2(x, y)$  — положительные в прямоугольнике  $D$  гладкие функции, разностную схему можно построить аналогично. Используя во внутренних точках сетки  $D_h^0$  замену выражений  $\frac{\partial}{\partial x} \left( k_1 \frac{\partial u}{\partial x} \right)$  и  $\frac{\partial}{\partial y} \left( k_2 \frac{\partial u}{\partial y} \right)$  разностными отношениями по приближенным формулам

$$\begin{aligned} \frac{\partial}{\partial x} \left[ k_1(x, y) \frac{\partial u(x, y)}{\partial x} \right] &\approx \Lambda_{xx} u(x, y) \equiv \\ &\equiv \frac{1}{h} \left[ k_1(x + h/2, y) \frac{u(x + h, y) - u(x, y)}{h} - \right. \\ &\quad \left. - k_1(x - h/2, y) \frac{u(x, y) - u(x - h, y)}{h} \right], \\ \frac{\partial}{\partial y} \left[ k_2(x, y) \frac{\partial u(x, y)}{\partial y} \right] &\approx \tilde{\Lambda}_{yy} u(x, y) \equiv \\ &\equiv \frac{1}{h} \left[ k_2(x, y + h/2) \frac{u(x, y + h) - u(x, y)}{h} - \right. \\ &\quad \left. - k_2(x, y - h/2) \frac{u(x, y) - u(x, y - h)}{h} \right], \end{aligned}$$

получим разностную схему (2) вида

$$L_h u^{(h)} = \begin{cases} \tilde{\Delta}_{xx} u^{(h)} + \tilde{\Delta}_{yy} u^{(h)} = \varphi(mh, nh), & (mh, nh) \in D_h^0, \\ u|_{\Gamma_h} = \psi(s_{mn}), & (mh, nh) \in \Gamma_h. \end{cases}$$

Пользуясь формулой Тейлора, можно убедиться в том, что имеется второй порядок аппроксимации. Можно было бы доказать устойчивость построенной схемы, преодолевая некоторые дополнительные трудности, по сравнению с рассмотренными нами при разборе примера.

На практике, при решении конкретных задач, обычно ограничиваются обоснованиями принципиального характера на модельных задачах, типа проведенного выше. Конкретные суждения о погрешности получаются, как правило, не из теоретических оценок, а из сравнения между собой результатов расчетов, выполненных на сетках с различными значениями шага  $h$ .

После того, как разностная краевая задача, аппроксимирующая дифференциальную, построена, нужно еще указать не слишком трудоемкий способ ее решения. Ведь при малом  $h$  задача (2) есть система скалярных уравнений очень высокого порядка. В разобранный нами примере решение разностных уравнений — сложная и интересная задача, но мы отложим ее рассмотрение до §§ 35, 36.

### ЗАДАЧИ

1. Доказать, что если во внутренних точках области  $D_h$  функция  $u^{(h)}$  удовлетворяет уравнению

$$\Delta_h u^{(h)}|_{(mh, nh)} = 0, \quad m, n = 1, 2, \dots, M-1, \\ Mh = 1,$$

то либо  $u^{(h)}$  принимает всюду на  $D_h$  одинаковые значения, либо наибольшее и наименьшее значения функции  $u^{(h)}$  не достигаются ни в одной внутренней точке сетки  $D_h$  (усиленный принцип максимума).

2. Если во всех внутренних точках области  $D_h$  выполнено условие  $\Delta_h u^{(h)} \geq 0$ , причем хотя бы в одной точке неравенство строгое, то  $u^{(h)}$  не достигает своего наибольшего значения ни в одной внутренней точке.

3. Рассмотрим разностную схему  $L_h u^{(h)} = f^{(h)}$  вида

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_{m+1, n} + u_{m, n+1} + u_{m-1, n} + u_{m, n-1} - 4u_{mn}}{h^2} = \\ = \varphi(mh, nh), & (mh, nh) \in D_h^0, \\ u_{mn} = \psi_1(s_{mn}) & \text{на } \Gamma_h^{(1)}, \\ \frac{u_{1, n} - u_{0, n}}{h} = \psi_2(nh), & n = 1, \dots, M-1. \end{cases}$$

Эта разностная схема аппроксимирует задачу (рис. 43)

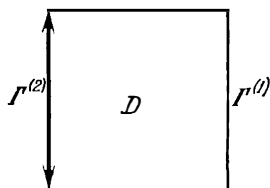


Рис. 43.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \varphi(x, y) \quad (x, y) \in D,$$

$$u(x, y)|_{\Gamma^{(1)}} = \psi_1(s), \quad (x, y) \in \Gamma^{(1)},$$

$$\frac{\partial u}{\partial n} \Big|_{\Gamma^{(2)}} = \psi_2(s), \quad (x, y) \in \Gamma^{(2)}.$$

а) Доказать, что при любых  $\varphi(mh, nh)$ ,  $\psi_1(s_{mn})$ ,  $\psi_2(nh)$  задача  $L_h u^{(h)} = f^{(h)}$  имеет единственное решение.

б) Доказать, что если  $\varphi(mh, nh)$  неотрицательно, а  $\psi_1(s_{mn})$  и  $\psi_2(nh)$  неположительны, то  $u^{(h)}$  неположительно.

в) Доказать, что при любых  $\varphi$ ,  $\psi_1$  и  $\psi_2$  имеет место оценка вида

$$\max_{m, n} |u_{mn}| \leq c \left( \max_{m, n} |\varphi_{mn}| + \max_{(mh, nh) \in \Gamma^{(1)}} |\psi_1(s_{mn})| + \max_n |\psi_2(nh)| \right),$$

где  $c$  — некоторая постоянная, не зависящая от  $h$ . Вычислить  $c$ .

### § 35. Метод установления

**1. Идея метода установления.** Для вычисления решений многих стационарных задач математической физики, описывающих равновесные состояния, рассматривают последние как результат установления развивающегося во времени процесса, расчет которого часто оказывается проще, чем прямой расчет равновесного состояния.

Мы проиллюстрируем применение метода установления примером алгоритма для вычисления решения разностной задачи Дирихле

$$\left. \begin{aligned} \Delta_{xx} u_{mn} + \Delta_{yy} u_{mn} &= \varphi(x_m, y_n), \quad m, n = 1, 2, \dots, M-1, \\ u_{mn} |_{\Gamma} &= \psi(s_{mn}), \end{aligned} \right\} \quad (1)$$

аппроксимирующей дифференциальную задачу Дирихле

$$\left. \begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= \varphi(x, y), \quad 0 \leq x, y \leq 1, \\ u |_{\Gamma} &= \psi(s). \end{aligned} \right\} \quad (2)$$

В случае задачи (1), которым мы будем заниматься, удастся провести теоретический анализ различных алгоритмов установления с помощью конечных рядов Фурье. Отметим сразу же, что для решения разностных эллиптических задач, подобных задаче (1), разработаны гораздо более эффективные итерационные методы; некоторые из них будут изложены в §§ 36, 37.

Способы точного решения задачи (1), выдерживающие обобщения на случай переменных коэффициентов и областей с кри-



волинейной границей, например метод исключения Гаусса, при сколько-нибудь больших  $M$  становятся неудобными и не применяются.

Изложим сначала наводящие соображения. Решение  $u(x, y)$  задачи (2) можно понимать как не зависящую от времени температуру в точке  $(x, y)$  пластинки, находящейся в тепловом равновесии. Функции  $\varphi(x, y)$  и  $\psi(s)$  означают в таком случае соответственно распределение источников тепла и температуру на границе.

Рассмотрим вспомогательную нестационарную задачу о распространении тепла

$$\left. \begin{aligned} \frac{\partial U}{\partial t} &= \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} - \varphi(x, y), \\ U|_{\Gamma} &= \psi(s), \\ U(x, y, 0) &= \psi_0(x, y), \end{aligned} \right\} \quad (3)$$

где  $\varphi$  и  $\psi$  те же, что и в задаче (2), а  $\psi_0(x, y)$  произвольно.

Поскольку источники тепла  $\varphi(x, y)$  и температура на границе  $\psi(s)$  не зависят от времени, то естественно ожидать, что и решение  $U(x, y, t)$  с течением времени будет меняться все медленнее, распределение температур  $U(x, y, t)$  в пределе при  $t \rightarrow \infty$  превращается в равновесное распределение температур  $u(x, y)$ , описываемое задачей (2). Поэтому вместо стационарной задачи (2) можно решать нестационарную задачу (3) до того времени  $t$ , пока ее решение перестанет меняться в пределах интересующей нас точности. В этом и состоит идея решения стационарных задач *методом установления*.

В соответствии с этим вместо задачи (2) будем решать задачу (3), а вместо разностной схемы (1) для задачи (2) рассмотрим и сопоставим три различные разностные схемы для задачи (3).

Именно, рассмотрим простейшую явную разностную схему

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^p + \Lambda_{yy} u_{mn}^p - \varphi(x_m, y_n), \\ u_{mn}^{p+1}|_{\Gamma} &= \psi(s_{mn}), \\ u_{mn}^0 &= \psi_0(x_m, y_n). \end{aligned} \right\} \quad (4)$$

Рассмотрим также простейшую неявную разностную схему

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^{p+1} + \Lambda_{yy} u_{mn}^{p+1} - \varphi(x_m, y_n), \\ u_{mn}^{p+1}|_{\Gamma} &= \psi(s_{mn}), \\ u_{mn}^0 &= \psi_0(x_m, y_n). \end{aligned} \right\} \quad (5)$$

Наконец, исследуем схему переменных направлений (12) § 31:

$$\left. \begin{aligned} \frac{\tilde{u}_{mn} - u_{mn}^p}{\tau} &= \frac{1}{2} [\Lambda_{xx} \tilde{u}_{mn} + \Lambda_{yy} u_{mn}^p - \varphi(x_m, y_n)], \\ \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} &= \frac{1}{2} [\Lambda_{xx} \tilde{u}_{mn} + \Lambda_{yy} u_{mn}^{p+1} - \varphi(x_m, y_n)], \\ u_{mn}^{p+1} |_{\Gamma} &= \tilde{u}_{mn} |_{\Gamma} = \psi(s_{mn}), \\ u_{mn}^0 &= \psi_0(x_m, y_n). \end{aligned} \right\} \quad (6)$$

Будем считать, что  $\psi_0(x_m, y_n)$  задано так, чтобы на границе выполнялось равенство

$$\psi_0 |_{\Gamma} = \psi(s_{mn}). \quad (7)$$

Вычисление  $u^{p+1} = \{u_{mn}^{p+1}\}$  по уже известному  $u^p = \{u_{mn}^p\}$  для схемы (4) осуществляется по явным формулам.

Вычисление  $u^{p+1} = \{u_{mn}^{p+1}\}$  при уже вычисленном  $u^p = \{u_{mn}^p\}$  по схеме (5) требует решения задачи

$$\left. \begin{aligned} \Lambda_{xx} u_{mn}^{p+1} + \Lambda_{yy} u_{mn}^{p+1} - \frac{u_{mn}^{p+1}}{\tau} &= \varphi(x_m, y_n) - \frac{u_{mn}^p}{\tau}, \\ u_{mn}^{p+1} |_{\Gamma} &= \psi(s_{mn}). \end{aligned} \right\}$$

Эта задача ничем не проще исходной задачи (1). Поэтому простейшую неявную схему не имеет смысла использовать для приближенного вычисления. Наконец, вычисление  $u^{p+1} = \{u_{mn}^{p+1}\}$  по уже известным  $u^p = \{u_{mn}^p\}$  по схеме (6) осуществляется прогонками в направлении оси  $Ox$  для вычисления решений  $\{\tilde{u}_{mn}\}$  одномерных задач при каждом фиксированном  $n$ , а затем прогонками в направлении оси  $Oy$  для вычисления решений  $\{u_{mn}^{p+1}\}$  одномерных задач при каждом фиксированном  $m$ .

Количество арифметических действий при этом пропорционально числу неизвестных. Для каждой из двух оставленных нами для дальнейшего изучения разностных схем (4) и (6) рассмотрим разность

$$\epsilon_{mn}^p \equiv u_{mn}^p - u_{mn} \quad (8)$$

между сеточной функцией  $u^p = \{u_{mn}^p\}$  и точным решением  $u = \{u_{mn}\}$  задачи (1), существование которого доказано в § 34.

Выясним условия, при которых погрешность  $\epsilon_{mn}^p$  решения  $u_{mn}^p$  нестационарной задачи стремится к нулю с ростом  $p$ , а также характер этого стремления к нулю; выберем оптимальным об-

разом шаг  $\tau$  и оценим объем вычислительной работы, необходимый для уменьшения нормы первоначальной погрешности

$$\varepsilon_{mn}^0 = \psi_0(x_m, y_n) - u_{mn}$$

в заданное число раз.

**2. Анализ явной схемы установления.** Решение  $\{u_{mn}\}$  задачи (1), очевидно, удовлетворяет уравнениям

$$\left. \begin{aligned} \frac{u_{mn} - u_{mn}}{\tau} &= \Lambda_{xx} u_{mn} + \Lambda_{yy} u_{mn} - \varphi(x_m, y_n), \\ u_{mn} |_{\Gamma} &= \psi(s_{mn}), \\ u_{mn} &= u_{mn}. \end{aligned} \right\}$$

Вычитая эти равенства из уравнений (4) почленно, получим для погрешности  $\varepsilon_{mn}^p$  следующую разностную задачу:

$$\left. \begin{aligned} \frac{\varepsilon_{mn}^{p+1} - \varepsilon_{mn}^p}{\tau} &= \Lambda_{xx} \varepsilon_{mn}^p + \Lambda_{yy} \varepsilon_{mn}^p, \\ \varepsilon_{mn}^{p+1} |_{\Gamma} &= 0, \\ \varepsilon_{mn}^0 &= \psi_0(x_m, y_n) - u_{mn}. \end{aligned} \right\} \quad (9)$$

Заметим, что сеточная функция  $\varepsilon_{mn}^p$  при каждом  $p$ ,  $p = 0, 1, \dots$ , обращается в нуль на границе  $\Gamma$ . Ее можно считать элементом линейного пространства функций, определенных на сетке  $(x_m, y_n) = (mh, nh)$ ,  $m, n = 0, 1, \dots, N$ , и обращающихся в нуль в точках  $\Gamma$ . Норму в этом пространстве определим, как в § 27, равенством

$$\|\varepsilon^p\| = \left( \sum_{m,n} |\varepsilon_{mn}^p|^2 \right)^{1/2}.$$

В § 27 мы получили представление для решения задачи (17) в виде конечного ряда Фурье. Эта задача только обозначением неизвестной функции отличается от разностной схемы (9) для погрешности  $\varepsilon^p = \{\varepsilon_{mn}^p\}$ . Поэтому

$$\varepsilon^p = \sum_{r,s} (c_{rs} \lambda_{rs}^p) \psi^{(r,s)}, \quad (10)$$

где  $c_{rs}$  — коэффициенты разложения начальной погрешности  $\varepsilon^0 = \{\varepsilon_{mn}^0\}$  в конечный ряд Фурье, а числа  $\lambda_{rs}$  задаются формулой

$$\lambda_{rs} = 1 - \frac{4\tau}{h^2} \left( \sin^2 \frac{r\pi}{2M} + \sin^2 \frac{s\pi}{2M} \right). \quad (11)$$

Числа  $c_{rs}^p \equiv c_{rs} \lambda_{rs}^p$  являются коэффициентами разложения погрешности  $\varepsilon^p = \{\varepsilon_{mn}^p\}$  в ряд Фурье по ортонормальному базису  $\psi^{(r, s)}$ . Поэтому

$$\|\varepsilon^p\| = \left( \sum |c_{rs} \lambda_{rs}^p|^2 \right)^{1/2}, \quad \|\varepsilon^0\| = \left( \sum |c_{rs}|^2 \right)^{1/2}. \quad (12)$$

Отсюда видно, что

$$\frac{\|\varepsilon^p\|}{\|\varepsilon^0\|} \leq \left\{ \max_{r, s} |\lambda_{rs}| \right\}^p. \quad (13)$$

При этом всегда можно задать  $\varepsilon^0$  так, чтобы равенство достигалось. Для этого нужно взять  $\varepsilon^0 = \psi^{(r', s')}$ , где  $(r', s')$  — та пара номеров, при которой

$$\max_{r, s} |\lambda_{rs}| = |\lambda_{r' s'}|.$$

Таким образом, для стремления  $\|\varepsilon^p\|/\|\varepsilon^0\|$  к нулю при  $p \rightarrow \infty$  нужно, чтобы выполнялось неравенство

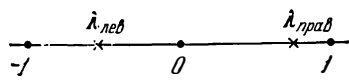
$$\max_{r, s} |\lambda_{rs}| < 1.$$

Наиболее быстрое убывание погрешности получится при таком выборе  $\tau$ , при котором  $\max_{r, s} |\lambda_{rs}|$  принимает наименьшее возможное значение. Из формулы (11) находим самую левую и самую правую точки  $\lambda_{rs}$ :

$$\lambda_{\text{лев}} = 1 - \frac{8\tau}{h^2} \cos^2 \frac{\pi}{2M},$$

$$\lambda_{\text{прав}} = 1 - \frac{8\tau}{h^2} \sin^2 \frac{\pi}{2M}$$

(рис. 44). Увеличивая  $\tau$ , начиная от  $\tau = 0$ , мы вызываем сдвиг обеих этих точек влево. При том значении  $\tau$ , при котором эти точки будут симметричны относительно точки  $\lambda = 0$ ,



$$-\lambda_{\text{лев}} = \lambda_{\text{прав}}, \quad (14)$$

Рис. 44.

дальнейшее увеличение  $\tau$  нецелесообразно. Действительно, при таком увеличении правая точка  $\lambda_{\text{прав}}$  будет продолжать приближаться к нулю, но зато левая, которая станет больше нее по модулю  $\max |\lambda_{rs}| = -\lambda_{\text{лев}}$ , будет удаляться от нуля.

При том  $\tau$ , при котором  $\lambda_{\text{лев}} = -1$ , и при больших  $\tau$  погрешность  $\varepsilon^p$  вообще не будет стремиться к нулю.

Итак, оптимальное  $\tau = h^2/4$  находим из условия (14). При этом

$$\max_{r, s} |\lambda_{rs}| = 1 - 2 \sin^2 \frac{\pi}{2M}.$$

Поэтому для уменьшения нормы первоначальной погрешности  $\epsilon^0 = \{\epsilon_{mn}^0\}$  в заданное число  $e$  раз требуется проделать такое число  $p$  шагов итерационного процесса (4), чтобы

$$\left(1 - 2 \sin^2 \frac{\pi}{2M}\right)^p < e^{-1}.$$

Отсюда

$$p \geq - \frac{1}{\ln \left(1 - 2 \sin^2 \frac{\pi}{2M}\right)} \approx \frac{2M^2}{\pi^2}.$$

Подсчитаем число арифметических действий, необходимых для уменьшения ошибки в  $e$  раз. На каждый переход от  $u^p$  к  $u^{p+1}$  требуется  $cM^2$  арифметических действий. Поэтому их общее число  $c p M^2 = O(M^4)$ .

**3. Схема переменных направлений.** Займемся теперь исследованием поведения погрешности  $\epsilon^p = \{\epsilon_{mn}^p\}$  для схемы (6).

Аналогично предыдущему убеждаемся, что погрешность  $\epsilon^p$  в этом случае удовлетворяет разностной краевой задаче

$$\left. \begin{aligned} \frac{\bar{\epsilon}_{mn} - \epsilon_{mn}^p}{\tau/2} &= \Lambda_{xx} \bar{\epsilon}_{mn} + \Lambda_{yy} \epsilon_{mn}^p, \\ \frac{\epsilon_{mn}^{p+1} - \bar{\epsilon}_{mn}}{\tau/2} &= \Lambda_{xx} \bar{\epsilon}_{mn} + \Lambda_{yy} \epsilon_{mn}^{p+1}, \\ \bar{\epsilon}_{mn} |_{\Gamma} &= \epsilon_{mn}^p |_{\Gamma} = 0, \\ \epsilon_{mn}^0 &= \psi_0(x_m, y_n) - u_{mn}. \end{aligned} \right\} \quad (15).$$

Решение задачи (15) было выписано в виде конечного ряда Фурье в § 27. Как и для задачи (9), оно имеет вид (10):

$$\epsilon^p = \sum (c_{rs} \lambda_{rs}^p) \psi^{(r,s)},$$

где  $c_{rs}$  — коэффициенты разложения начальной погрешности

$$\epsilon^0 = \sum c_{rs} \psi^{(r,s)}$$

в конечный ряд Фурье, но числа  $\lambda_{rs}$ , на которые умножается гармоника  $\psi^{(r,s)}$  при переходе от  $\epsilon^p$  к  $\epsilon^{p+1}$ , теперь другие:

$$\lambda_{rs} = \frac{\left(1 - 2\tau M^2 \sin^2 \frac{\pi r}{2M}\right) \left(1 - 2\tau M^2 \sin^2 \frac{\pi s}{2M}\right)}{\left(1 + 2\tau M^2 \sin^2 \frac{\pi r}{2M}\right) \left(1 + 2\tau M^2 \sin^2 \frac{\pi s}{2M}\right)}. \quad (16).$$

Как и при анализе сходимости схемы (4), выполнено неравенство (13):

$$\frac{\|\epsilon^p\|}{\|\epsilon^0\|} \leq \left\{ \max_{r,s} |\lambda_{rs}| \right\}^p,$$

причем равенство достигается при некотором специальном задании  $\varepsilon^0 = \{\varepsilon_{mn}^0\}$ .

Из выражения (16) для  $\lambda_{rs}$  видно, что при любом  $\tau$  выполнено неравенство  $|\lambda_{rs}| < 1$  и, следовательно, имеет место стремление  $\|\varepsilon^p\|$  к нулю. Далее,  $\lambda_{rs} = \lambda_r \cdot \lambda_s$ , где

$$\lambda_k = \frac{1 - 2\tau M^2 \sin^2 \frac{\pi k}{2M}}{1 + 2\tau M^2 \sin^2 \frac{\pi k}{2M}}, \quad k = 1, 2, \dots, M-1.$$

Поэтому  $\max_{r,s} |\lambda_{rs}|$  достигается при  $r = s = r'$ , где  $r'$  — тот номер, при котором величина  $|\lambda_{r'}|$  максимальна. Очевидно, что функция  $y = (1-x)/(1+x)$  монотонна. Поэтому

$$\lambda_s = \frac{1 - 2\tau M^2 \sin^2 \frac{\pi s}{2M}}{1 + 2\tau M^2 \sin^2 \frac{\pi s}{2M}}$$

лежит между точками

$$\lambda_{\text{лев}} = \frac{1 - 2\tau M^2 \cos^2 \frac{\pi}{2M}}{1 + 2\tau M^2 \cos^2 \frac{\pi}{2M}}$$

и

$$\lambda_{\text{прав}} = \frac{1 - 2\tau M^2 \sin^2 \frac{\pi}{2M}}{1 + 2\tau M^2 \sin^2 \frac{\pi}{2M}}$$

на вещественной прямой. Увеличение  $\tau$  вызывает сдвиг точек  $\lambda_{\text{лев}}$  и  $\lambda_{\text{прав}}$  влево. Поэтому значение  $\max_s |\lambda_s|$  будет наименьшим при том  $\tau$ , при котором  $-\lambda_{\text{лев}} = \lambda_{\text{прав}}$ , т. е. при  $\tau \approx \frac{1}{\sqrt{2} \pi M}$ .

В этом случае

$$\max |\lambda_{rs}| = 1 - \frac{\sqrt{2} \pi}{M} + O\left(\frac{1}{M^2}\right).$$

Для уменьшения нормы погрешности  $\|\varepsilon^p\|$  в заданное число  $e$  раз по сравнению с первоначальным значением нормы погрешности  $\|\varepsilon^0\|$  число шагов  $p$  должно быть найдено из условия

$$\left(1 - \frac{\pi \sqrt{2}}{M}\right)^p \leq e^{-1}, \text{ откуда}$$

$$p \approx \frac{M}{\pi \sqrt{2}} = O(M).$$

Каждый переход от  $u^p$  к  $u^{p+1}$  требует  $cM^2$  арифметических операций. Следовательно, общее число арифметических операций для уменьшения ошибки в  $e$  раз будет  $O(M^3)$ , а для уменьшения в заданное число  $k$  раз будет  $O(M^3 \ln k)$ .

Мы видим, что при больших  $M$  второй из рассмотренных нами процессов установления, использующий схему переменных направлений, приводит к уменьшению ошибки в заданное число раз ценой меньших затрат арифметических действий, чем метод установления, основанный на использовании простейшей явной разностной схемы (4): при достаточно больших значениях  $M$  (при мелкой сетке) схема переменных направлений оказывается выгоднее.

**4. Выбор точности.** Сделаем замечание о точности, которой следует добиваться, решая задачу (1) методом установления или другим итерационным методом, дающим последовательные приближения  $u^1, u^2, \dots, u^p$ . Разностная задача (1) аппроксимирует задачу (2) на гладком решении  $u(x, y)$  с порядком  $h^2 = 1/M^2$ . Поэтому точное решение  $u^{(h)}$  задачи (1) отличается от искомой таблицы  $[u]_h$  на величину порядка  $1/M^2$ . В связи с этим нет смысла вычислять решение  $u^{(h)}$  задачи (1) с большей точностью. Если считать, что нулевое приближение  $u^0 = \psi_0$  задано с погрешностью порядка 1, то число  $k$ , в которое мы хотим уменьшить норму погрешности, должно быть выбрано порядка  $M^2$ .

Добиваться уменьшения первоначальной погрешности более чем в  $O(M^2)$  раз было бы нецелесообразной затратой вычислительной работы.

При вычислениях на конкретной фиксированной сетке практически итерируют до тех пор, пока последовательные приближения  $u^p, u^{p+1}, \dots$  перестанут меняться в пределах удовлетворяющей нас точности.

**5. Границы применимости методов.** Разностная схема (4) и анализ убывания ошибки, проведенный нами, переносятся на разностные схемы, аппроксимирующие другие краевые задачи для эллиптического уравнения с переменными коэффициентами в области с криволинейными границами. Важно только, чтобы оператор  $\tilde{\Lambda}_h$ , аналогичный оператору  $-\Lambda_h \equiv -(\Lambda_{xx} + \Lambda_{yy})$  в схеме (1), рассматриваемый на сеточных функциях, удовлетворяющих однородным граничным условиям, был самосопряженным и чтобы его собственные значения  $\mu_j$  были одного знака:

$$0 < \mu_{\min} < \mu_j < \mu_{\max}.$$

В этом случае для анализа используются конечные ряды Фурье не по функциям

$$\psi^{(r, s)} = 2 \sin \frac{\pi r}{M} \sin \frac{\pi s}{M},$$

а по ортонормальной системе собственных функций этого самосопряженного оператора  $\tilde{\Lambda}_h$ . Факт существования и полноты такой системы собственных функций известен, а их конкретный вид нигде не используется.

Разностная схема (6) переменных направлений выдерживает обобщение на случай задачи Дирихле с переменными коэффициентами в области с криволинейной границей. (Однако анализ Фурье становится невозможным.) В случае краевых условий вида  $\alpha u + \beta \frac{\partial u}{\partial n} \Big|_{\Gamma} = \psi$  прямое обобщение схемы (6) не приводит к распадению алгоритма на одномерные прогонки.

### ЗАДАЧИ

1. Написать по аналогии с рассмотренными схемами (4) и (6) явную и неявную схемы решения установлением задачи Дирихле

а) для уравнения Лапласа с переменными коэффициентами:

$$\frac{\partial}{\partial x} \left[ k_1(x, y) \frac{\partial u}{\partial x} \right] + \frac{\partial}{\partial y} \left[ k_2(x, y) \frac{\partial u}{\partial y} \right] = 0, \quad 0 \leq x, y \leq 1$$

$$u|_{\Gamma} = \psi(x, y)|_{\Gamma}$$

б) для квазилинейного уравнения

$$\frac{\partial}{\partial x} \left[ k_1(u) \frac{\partial u}{\partial x} \right] + \frac{\partial}{\partial y} \left[ k_2(u) \frac{\partial u}{\partial y} \right] = 0, \quad 0 \leq x, y \leq 1,$$

$$u|_{\Gamma} = \psi(x, y)|_{\Gamma}.$$

2. Показать, что в методе переменных направлений для решения разностной задачи Дирихле

$$\left. \begin{aligned} \Lambda_{xx} u_{mn} + \Lambda_{yy} u_{mn} &= \varphi(x_m, y_n), \\ m, n &= 1, 2, \dots, M-1; Mh = 1, \\ u_{mn}|_{\Gamma} &= \psi(x, y)|_{\Gamma} \end{aligned} \right\}$$

итерациями можно выбрать итерационный параметр  $\tau$  так, чтобы после первой же итерации разложение погрешности  $\varepsilon^p$  в конечный ряд Фурье не содержало любой наперед заданной гармоники  $\psi^{(r, s)}$ .

## § 36. Итерации с переменным шагом

1. Идея Ричардсона. Механизм сходимости простейшей схемы установления (4) § 35

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^p + \Lambda_{yy} u_{mn}^p - \varphi(x_m, y_n), \\ u_{mn}^{p+1} \Big|_{\Gamma} &= \psi(s_{mn}), \\ u_{mn}^0 &= \psi_0(x_m, y_n) \end{aligned} \right\} \quad (1)$$



состоит, как мы видели, в погашении каждой из гармоник  $\psi^{(r, s)}$ , по которым разлагается погрешность  $\epsilon_{mn}^0 = u_{mn} - u_{mn}^0$  нулевого приближения в конечный ряд Фурье. Если

$$\epsilon^p = \sum_{r, s} c_{rs}^p \psi^{(r, s)}, \quad (2)$$

то коэффициенты Фурье погрешности следующего приближения

$$\epsilon^{p+1} = \sum_{r, s} c_{rs}^{p+1} \psi^{(r, s)}$$

выражаются через  $c_{rs}^p$  по формулам (см. (10), (11), § 35):

$$c_{rs}^{p+1} = (1 - \tau \mu_{rs}) c_{rs}^p, \quad \text{где } \mu_{rs} = 4M^2 \left( \sin^2 \frac{\pi r}{2M} + \sin^2 \frac{\pi s}{2M} \right). \quad (3)$$

При фиксированном выборе  $\tau$  не все гармоники погашаются одинаково быстро. Более сильно погашаются те гармоники  $\psi^{(r, s)}$ , для которых множитель погашения  $\lambda_{rs} \equiv 1 - \tau \mu_{rs}$  ближе к нулю, т. е. те, для которых  $\mu_{rs} \approx 1/\tau$ . Это наводит на мысль от шага к шагу менять параметр  $\tau$ , чтобы поочередно хорошо гасились все гармоники  $\psi^{(r, s)}$ , и в результате за несколько шагов все гармоники гасились бы более или менее равномерно.

В этом состоит идея Ричардсона для решения самосопряженных линейных систем уравнений с матрицей, все собственные значения которой имеют одинаковый знак.

**2. Чебышевский набор параметров.** Итерационный процесс Ричардсона задается формулами

$$\left. \begin{aligned} u_{mn}^{p+1} &= u_{mn}^p + \tau_{p+1} [\Lambda_n u_{mn}^p - \Phi(x_m, y_n)], \\ m, n &= 1, 2, \dots, M-1, \\ u_{mn}^{p+1} |_{\Gamma} &= \psi(s_{mn}); \quad \{u_{mn}^0\} \text{ задано} \end{aligned} \right\} \quad (4)$$

с итерационным параметром  $\tau_{p+1}$ , зависящим от номера итерации. Ричардсон указал приемлемый, но не оптимальный набор параметров  $\{\tau_p\}$ . Изложим результаты об оптимальном наборе итерационных параметров  $\{\tau_p\}$  и оценке скорости убывания нормы погрешности  $\|\epsilon^p\|$ . В силу формулы (3) очевидно, что коэффициенты Фурье  $c_{rs}^k$  погрешности  $\epsilon^k$  на  $k$ -м шаге выражаются через коэффициенты Фурье  $c_{rs}^0$  начальной погрешности  $\epsilon^0$  по формулам

$$c_{rs}^k = c_{rs}^0 \prod_{j=1}^k (1 - \tau_j \mu_{rs}), \quad r, s = 1, 2, \dots, M-1.$$

Введем обозначение  $Q_k(\mu)$ , положив

$$Q_k(\mu) \equiv \sum_{j=1}^k (1 - \tau_j \mu). \quad (5)$$

Тогда

$$\begin{aligned} \|\varepsilon^k\|^2 &= \sum_{r,s} |c_{rs}^k|^2 = \sum_{r,s} |Q_k(\mu_{rs}) c_{rs}^0|^2 \leq \\ &\leq \max_{r,s} |Q_k(\mu_{rs})| \sum_{r,s} |c_{rs}^0|^2 = \max_{r,s} |Q_k(\mu_{rs})| \cdot \|\varepsilon^0\|^2. \end{aligned}$$

Очевидно, что неравенство

$$\|\varepsilon^k\| \leq \max_{r,s} |Q_k(\mu_{rs})| \cdot \|\varepsilon^0\|$$

при некоторых  $\varepsilon^0$  становится точным равенством. Числа  $\mu_{rs}$ , задаваемые формулой (3), разбросаны по отрезку

$$a = \mu_{\min} \leq \mu \leq \mu_{\max} = b, \quad (6)$$

где

$$\left. \begin{aligned} a = \mu_{\min} &= 8M^2 \sin^2 \frac{\pi}{2M} \approx 2\pi^2, \\ b = \mu_{\max} &= 8M^2 \cos^2 \frac{\pi}{2M} \approx 8M^2. \end{aligned} \right\} \quad (6')$$

Мы не будем опираться на фактическое знание чисел  $\mu_{rs}$ , так как это — случайное обстоятельство, имеющее место только для нашего примера. Будем пользоваться лишь тем, что известны границы  $a$  и  $b$  отрезка (6), на котором они лежат. Поэтому, задав  $k$ , поставим задачу такого определения итерационных параметров  $\tau_1, \tau_2, \dots, \tau_k$ , чтобы среди всех многочленов  $Q_k(\mu)$  степени  $k$ , удовлетворяющих условию

$$Q(0) = 1, \quad (7)$$

многочлен (5) на отрезке  $a \leq \mu \leq b$  наименее уклонялся от нуля:

$$Q_k^* = \max_{a \leq \mu \leq b} |Q_k(\mu)| \text{ минимально.} \quad (8)$$

Эта задача теории аппроксимации функций решена в 1892 году А. А. Марковым. Искомый многочлен  $Q_k(\mu) \equiv \tilde{T}_k(\mu)$  выражается через многочлен Чебышева (см., например, В. Л. Гончаров, Теория интерполирования и приближения функций, М., 1954)

$$T_k(x) \equiv \cos k \arccos x \equiv \frac{1}{2} [(x + \sqrt{x^2 - 1})^k + (x - \sqrt{x^2 - 1})^k],$$

наименее уклоняющийся от нуля на отрезке  $-1 \leq x \leq 1$  среди всех многочленов степени  $k$ , с коэффициентом единица при  $x^k$ . Именно, если сделать линейное преобразование

$$x = \frac{b + a - 2\mu}{b - a}, \quad (9)$$

переводящее отрезок  $a \leq \mu \leq b$  в отрезок  $-1 \leq x \leq 1$ , а точку  $\mu = 0$  в точку  $x_0 = \frac{b+a}{b-a} > 1$ , то

$$\tilde{T}_k(\mu) = \frac{T_k(x)}{T_k(x_0)} = \frac{(x + \sqrt{x^2 - 1})^k + (x - \sqrt{x^2 - 1})^k}{(x_0 + \sqrt{x_0^2 - 1})^k + (x_0 - \sqrt{x_0^2 - 1})^k}. \quad (10)$$

Набор итерационных параметров  $\tau_1, \tau_2, \dots, \tau_k$ , при которых возникает многочлен (10), определяется из условия, чтобы нули  $\mu_j = 1/\tau_j$  многочлена  $\tilde{T}_k(\mu)$  при преобразовании (9) переходили в нули  $x_j$  многочлена Чебышева  $T_k(x)$ :

$$\left. \begin{aligned} \tau_j &= \frac{2}{b+a - (b-a)x_j}, \\ x_j &= \cos \frac{\pi(2j-1)}{2k}, \quad j = 1, 2, \dots, k. \end{aligned} \right\} \quad (11)$$

Оценим наибольшее отклонение  $Q_k^*$  многочлена  $Q_k(\mu) \equiv \tilde{T}_k(\mu)$  от нуля на отрезке  $a \leq \mu \leq b$ . Как известно из теории аппроксимации функций, многочлен Чебышева  $T_k(x)$  принимает наибольшее по модулю значение на отрезке  $-1 \leq x \leq 1$  в  $k+1$  точках, к числу которых принадлежат концы этого отрезка. Поэтому из (10) следует, что

$$Q_k^* = \frac{T_k(1)}{T_k(x_0)} = \frac{2}{(x_0 + \sqrt{x_0^2 - 1})^k + (x_0 - \sqrt{x_0^2 - 1})^k}. \quad (12)$$

Далее, из (9) получаем

$$\left. \begin{aligned} x_0 &= \frac{b+a}{b-a} = \frac{1+\eta}{1-\eta} = 1 + 2\eta + O(\eta^2), \\ \eta &= \frac{a}{b} = \frac{\mu_{\min}}{\mu_{\max}} \approx \frac{\pi^2}{4M^2}. \end{aligned} \right\} \quad (13)$$

Поэтому при больших  $M$

$$x_0 \pm \sqrt{x_0^2 - 1} = 1 \pm 2\sqrt{\eta} + O(\eta),$$

откуда, с учетом (13), следует

$$\begin{aligned} Q_k^* &= \frac{2}{[1 + 2\sqrt{\eta} + O(\eta)]^k + [1 - 2\sqrt{\eta} + O(\eta)]^k} = \\ &= 2 : \{e^{k \ln(1 + 2\sqrt{\eta} + O(\eta))} + e^{k \ln(1 - 2\sqrt{\eta} + O(\eta))}\} \approx \\ &\approx 2 : \{e^{k\pi/M} + e^{-k\pi/M}\}. \end{aligned}$$

Считая, что норма первоначальной ошибки  $\varepsilon^0$  порядка единицы,  $\|\varepsilon^0\| \approx 1$ , в силу замечания п. 4 § 35 о разумном порядке

точности, которого следует добиваться, решая задачу итерациями, надо выбрать  $k$  из условия  $Q_k^* \approx M^{-2}$ , т. е.

$$k \approx \frac{2 \ln M + \ln 2}{\pi} M \approx \frac{2 \ln M + \ln 2}{2 \sqrt{\eta}}. \quad (14)$$

Для погашения первоначальной погрешности в  $e$  раз надо выбрать  $k$  из условия  $Q_k^* \leq e^{-1}$ , т. е.

$$k \approx \frac{1 + \ln 2}{\pi} M \approx \frac{1 + \ln 2}{2 \sqrt{\eta}} = O(M). \quad (15)$$

Выбрав  $k$  из этого условия, можно затем первые  $k$  шагов итерации принять за первый цикл итераций и повторять весь цикл с тем же набором параметров  $\tau_1, \tau_2, \dots, \tau_k$ . Для уменьшения нормы погрешности в  $M^2$  раз потребуется такое число  $\nu$  циклов, чтобы  $e^{-\nu} \sim 1/M^2$ ,  $\nu \sim 2 \ln M$ . Общее число элементарных шагов итерационного процесса за  $\nu$  циклов будет

$$k\nu \approx \left( \frac{1 + \ln 2}{\pi} M \right) 2 \ln M = O(M \ln M).$$

Оно лишь в конечное число раз

$$2 \frac{1 + \ln 2}{2 + \ln 2 / (\ln M)} \leq 1 + \ln 2$$

превосходит число (14) элементарных шагов итерационного процесса, не использующего заикливание. Таким образом, использование цикла с  $k \approx (1 + \ln 2) / (2 \sqrt{\eta})$  дает упрощение без существенного увеличения числа итераций.

Использовать циклы длины  $k \ll 1 / (2 \sqrt{\eta})$  нецелесообразно. Например, при  $k = 1$  процесс Ричардсона (4) превратится в процесс простых итераций (1) с оптимальным выбором  $\tau$ . Число шагов процесса для уменьшения нормы первоначальной ошибки  $\|e^0\|$  в  $e$  раз будет  $\approx 2M^2/\pi^2$ , как показано в п. 2 § 35. Это число в  $O(M)$  раз превышает число шагов, необходимых для этой же цели при выборе длины  $k$  цикла в соответствии с (15).

**3. Нумерация итерационных параметров.** Формулы (11) задают оптимальный набор итерационных параметров  $\tau_1, \tau_2, \dots, \tau_k$  (при заданном фиксированном  $k$ ). Переставим как-либо члены последовательности  $\tau_1, \tau_2, \dots, \tau_k$ , расположив их в некоторой очередности  $\kappa^{(k)} = (\kappa_1, \kappa_2, \dots, \kappa_k)$ , и будем вести итерации по формулам

$$\left. \begin{aligned} u^{p+1} &= u^p + \tau_{\kappa_{p+1}} (\Lambda_{\kappa} u^p - \Phi), \\ u^{p+1}|_{\Gamma} &= \psi, \quad u^0 \text{ задано.} \end{aligned} \right\} \quad (16)$$

При идеальной реализации алгоритма (16) результат финальной,  $k$ -й итерации не зависит от выбранной очередности  $\kappa^{(k)} = (\kappa_1, \kappa_2, \dots, \kappa_k)$ . Но при реальном расчете, который ведется на машине с конечным числом знаков, этот порядок крайне важен. От него при больших  $k$  резко зависит чувствительность результата к ошибкам округления, допускаемым на промежуточных шагах процесса, т. е. вычислительная устойчивость алгоритма. Прежде чем приводить приемлемые порядки  $\kappa^{(k)} = (\kappa_1, \kappa_2, \dots, \kappa_k)$ , заметим, что исходный алгоритм (4), отвечающий  $\kappa^{(k)} = (1, 2, \dots, k)$ , практически непригоден.

Разберемся в механизме возникновения неустойчивости в этом случае. Пусть начальная погрешность  $\varepsilon^0$  имеет вид  $\varepsilon^0 =$

$= \sum c_{rs}^0 \psi^{(r, s)}$ ,  $c_{rs}^0 \approx 1$ , и расчет ведется точно, без ошибок округления. Тогда коэффициенты погрешности  $l$ -го приближения  $\varepsilon^l = \sum c_{rs}^l \psi^{(r, s)}$  выражаются формулой

$$c_{rs}^l = \prod_{j=1}^l (1 - \tau_j \mu_{rs}) c_{rs}^0.$$

Проследим за эволюцией  $c_{rs}^l$  с ростом  $l$  при  $r = M - 1$ ,  $s = M - 1$ . В этом случае

$$\left. \begin{aligned} \mu_{rs} &= \mu_{M-1, M-1} = \mu_{\max} = b \sim M^2, \\ \xi^l &\equiv \frac{c_{M-1, M-1}^l}{c_{M-1, M-1}^0} = \prod_{j=1}^l (1 - \tau_j b). \end{aligned} \right\} \quad (17)$$

Рассмотрим линейные функции  $1 - \tau_j \mu$ ,  $j = 1, 2, \dots, k$ , нули которых  $\mu_j = 1/\tau_j$  определяются формулами (11). Из этих формул видно, что при  $\frac{2j-1}{2k} < \frac{1}{6}$ , или  $j < \frac{k+3}{6}$ , справедливо неравенство  $\mu_j < \frac{b}{2}$ , и поэтому (рис. 45)

$$|1 - \tau_j b| > 1 \quad \text{при} \quad j < \frac{k+3}{6}. \quad (18)$$

Если  $k \sim \frac{1}{2\sqrt{\eta}} \sim M$ , а  $j \sim 1$ , то  $\mu_j \sim a + \frac{1}{M^2}$ ,  $\tau_j \sim \frac{1}{a} - \frac{1}{M^2}$ , и поэтому

$$|1 - \tau_j b| \sim \frac{b}{a} \sim M^2.$$

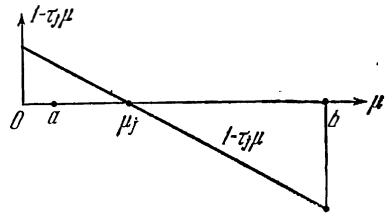


Рис. 45.

Таким образом, величина  $\xi^l$ , определенная формулой (17), увеличивается сначала примерно в  $M^2$  раз за один шаг, а потом медленнее. В силу (18) этот рост продолжается во всяком случае до тех пор, пока  $l \leq (k+3)/6$ , так что при  $l \sim k/6$  величина  $c_{M-1, M-1}^l$ , а вместе с ней и  $\|\varepsilon^l\|$ , окажется очень большой и тем большей, чем больше число  $k$ . При этом порядки величин значений приближения  $u^l = \{u_{m_i}^l\}$  могут выйти за пределы допустимых для данной машины уже при умеренных  $k, k \ll M$ .

Если гипотетически считать, что этого не произошло и что счет продолжается точно, то к шагу  $l = k$  величина  $c_{M-1, M-1}^l$  вновь уменьшится, так что  $\xi^k \leq Q_k^*$ .

Но дело в том, что даже если и не произошло переполнения ячеек машинной памяти при  $l \sim k/6$ , то неизбежные малые относительные ошибки округления при  $l \sim k/6$  велики по абсолютной величине. Они носят случайный характер, так что в их разложении в конечный ряд Фурье будут присутствовать все члены, в частности член вида

$$\bar{c}_{1,1} \psi^{(1,1)},$$

причем  $\bar{c}_{1,1}$  — величина не малая по абсолютной величине.

Покажем, что при дальнейших шагах итерации ошибка  $\bar{c}_{1,1} \psi^{(1,1)}$ , внесенная на гармонику  $\psi^{(1,1)}$  в результате округления на шаге  $l \sim k/6$ , не претерпевает существенного погашения и недопустимо искажает результат. Вызванный ею вклад  $\bar{c}_{1,1}^k \psi^{(1,1)}$  в приближение  $u^k$ , полученное на последнем,  $k$ -м шаге, выражается формулой

$$\bar{c}_{1,1}^k = \left[ \prod_{j=l+1}^{1k} (1 - \tau_j \mu_{11}) \right] \bar{c}_{1,1} = \left[ \prod_{j=l+1}^k (1 - \tau_j a) \right] \bar{c}_{1,1}.$$

Но при  $j > (k+3)/6$ , очевидно,

$$\mu_j > \frac{1}{2} \left[ b + a - \frac{b-a}{2} \right] > \frac{b}{4} \sim M^2.$$

Поэтому

$$1 - \tau_j a \sim 1 - \frac{1}{M^2}, \quad \left| \prod_{j=l+1}^k (1 - \tau_j a) \right| \sim \left( 1 - \frac{1}{M^2} \right)^{k-1} \sim 1,$$

так что  $\bar{c}_{1,1}^k \sim \bar{c}_{1,1}$ , и погашения погрешности округления не произошло. Практическая непригодность нумерации параметров  $\kappa^{(k)} = (1, 2, \dots, k)$  показана.

Если на  $l$ -м шаге процесса (13) внесена погрешность округления

$$c_{rs} \psi^{(r,s)},$$

то на  $k$ -м шаге она разовьется в

$$\left[ \prod_{j=i+1}^k (1 - \tau_j \mu_{rs}) \right] c_{rs} \psi^{(r, s)}.$$

Поэтому целесообразно стремиться к такому выбору  $\kappa^{(k)} = (\kappa_1, \kappa_2, \dots, \kappa_k)$ , при котором имеет место оценка

$$\max_{a \leq \mu \leq b} \left[ \prod_{j=i+1}^k (1 - \tau_j \mu) \right] < A \quad (19)$$

при умеренном значении  $A$ .

Пусть  $c_{rs}^0 \psi^{(r, s)}$  — компонента погрешности  $\varepsilon^0$  нулевого приближения. К  $l$ -му шагу она разовьется в

$$\left[ \prod_{j=1}^l (1 - \tau_j \mu_{rs}) \right] c_{rs}^0 \psi^{(r, s)}.$$

Если норма этой функции велика, то ее округление дает значительный по абсолютной величине вклад во все гармоники. Может оказаться, что вклад, полученный некоторыми гармониками, при дальнейших итерациях не погашается и сильно искажает результат. Поэтому естественно стремиться к такому выбору  $\kappa^{(k)} = (\kappa_1, \kappa_2, \dots, \kappa_k)$ , при котором выполнено неравенство

$$\max_{a \leq \mu \leq b} \left| \prod_{j=1}^l (1 - \tau_j \mu) \right| < B \quad (20)$$

при умеренном значении  $B$ .

В работе В. И. Лебедева и С. А. Финогенова, ЖВМ и МФ 11, № 2, (1971), и в работе А. А. Самарского [23] указаны различные целесообразные способы нумерации параметров и освещена история вопроса. Приведем результаты работы В. И. Лебедева и С. А. Финогенова. В ней предполагается, что  $k$  является степенью числа 2, т. е.  $k = 2^i$ , и указаны рекуррентные формулы для построения  $\kappa^{(k)}$ .

Именно, при  $i = 1$

$$\kappa^{(k)} = (1, 2).$$

Если при  $k = 2^{i-1}$  порядок  $\kappa^{(2^{i-1})}$  уже определен:

$$\kappa^{(2^{i-1})} = (\kappa_1, \kappa_2, \dots, \kappa_{2^{i-1}}),$$

то полагаем

$$\kappa^{(2^i)} = (\kappa_1, 2^i + 1 - \kappa_1, \kappa_2, 2^i + 1 - \kappa_2, \dots, 2^i + 1 - \kappa_{2^{i-1}}) \quad (21)$$

В частности, при  $i = 2$ ,  $i = 3$ ,  $i = 4$  последовательно получаем

$$(1, 4, 2, 3); \quad (1, 8, 4, 5, 2, 7, 3, 6); \\ (1, 16, 8, 9, 4, 13, 5, 12, 2, 15, 7, 10, 3, 14, 6, 11).$$

При указанном способе (21) упорядочения итерационных параметров числа  $A$  и  $B$  в неравенствах (19) и (20) можно брать независимыми от  $k$  и от  $l$ .

Алгоритм упорядочения параметров, указанный А. А. Самарским, формулируется несколько сложнее, но при этом  $k$  не обязательно степень двойки. Число  $k$  может иметь вид  $k = (2j + 1)2^i$ . Вместо (19), (20) установлены другие оценки, гарантирующие устойчивость в некотором смысле.

**4. Метод Дугласа — Рэкфорда.** В схеме переменных направлений (6) § 35 будем считать итерационный параметр  $\tau$  зависящим от номера шага, положив

$$\frac{\tilde{u}_{mn} - u_{mn}^p}{\tau_{p+1}} = \frac{1}{2} [\Lambda_{xx} \tilde{u}_{mn} + \Lambda_{yy} u_{mn}^p - \varphi(x_m, y_n)], \\ \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau_{p+1}} = \frac{1}{2} [\Lambda_{xx} \tilde{u}_{mn} + \Lambda_{yy} u_{mn}^{p+1} - \varphi(x_m, y_n)], \\ u_{mn}^{p+1} |_{\Gamma} = \tilde{u}_{mn} |_{\Gamma} = \psi(s_{mn}), \quad m, n = 1, 2, \dots, M-1.$$

Для погрешности  $\epsilon^k = u^k - u$  получим выражение

$$\epsilon^k = \sum_{r,s=1}^{M-1} c_{rs}^k \psi^{(r,s)}, \quad c_{rs}^k = \left[ \prod_{j=1}^k \lambda_r(\tau_j) \lambda_s(\tau_j) \right] c_{rs}^0,$$

где

$$\lambda_i(\tau) \equiv \frac{1 - 2\tau M^2 \sin^2 \frac{\pi i}{2M}}{1 + 2\tau M^2 \sin^2 \frac{\pi i}{2M}}, \quad i = 1, 2, \dots, k.$$

При заданном  $k$  оптимальным является такой выбор чисел  $\tau_1, \tau_2, \dots, \tau_k$ , при котором величина

$$\max_{r,s} \left| \prod_{j=1}^k \lambda_r(\tau_j) \lambda_s(\tau_j) \right|$$

принимает наименьшее значение. Если не пользоваться точным знанием чисел  $\lambda_r(\tau)$  и  $\lambda_s(\tau)$ , а лишь границами, где они лежат, возникает задача чебышевского типа для произведения дробно-линейных функций, подобная рассмотренной в п. 2 для многочленов.

Постановка этой задачи, как впрочем и предложение использовать для решения уравнения Пуассона процесс установления



со схемой переменных направлений, принадлежит Дугласу, Писману и Рэкфорду. В работе Дугласа и Рэкфорда 1956 года, которую мы здесь изложим, было дано приближенное решение этой задачи. При их выборе итерационных параметров число шагов итерационного процесса, нужное для уменьшения ошибки в  $e$  раз, есть  $O(\ln M)$ , а число арифметических действий есть  $O(M^2 \ln M)$ .

Покажем, что, задав произвольно положительное  $q, q < 1$ , можно так выбрать итерационные параметры  $\tau_1, \tau_2, \dots, \tau_k$  в количестве  $k = O(\ln M)$ , чтобы выполнялись неравенства

$$|[\lambda_r(\tau_1)\lambda_s(\tau_1)][\lambda_r(\tau_2)\lambda_s(\tau_2)] \dots [\lambda_r(\tau_k)\lambda_s(\tau_k)]| < q, \quad (22)$$

$$r, s = 1, 2, \dots, M-1.$$

Тогда  $\|\epsilon^k\| \leq q \|\epsilon^0\|$ . Если производить первые  $k$  итераций с итерационными параметрами  $\tau_1, \tau_2, \dots, \tau_k$ , затем следующие  $k$  итераций, снова используя  $\tau_1, \tau_2, \dots, \tau_k$ , то для уменьшения нормы погрешности в  $e$  раз потребуется, очевидно, некоторое не зависящее от  $M$  число таких циклов, содержащих по  $k = O(\ln M)$  итераций каждый.

Обоснуем равенство (22) и при этом объясним, как можно выбрать параметры  $\tau_1, \tau_2, \dots, \tau_k$ . Очевидно, что

$$|\lambda_i(\tau)| < 1, \quad i = 1, 2, \dots, M-1, \quad \tau > 0.$$

Поэтому для выполнения при любых  $r, s = 1, 2, \dots, M-1$  неравенства (22) достаточно, чтобы при любом  $i = 1, 2, \dots, M-1$  хотя бы один из  $k$  сомножителей  $\lambda_i(\tau_p), p = 1, 2, \dots, k$ , удовлетворял неравенству

$$|\lambda_i(\tau_p)| = \left| \frac{1 - 2\tau_p M^2 \sin^2 \frac{\pi i}{2M}}{1 + 2\tau_p M^2 \sin^2 \frac{\pi i}{2M}} \right| \leq \sqrt{q}. \quad (23)$$

Все числа  $2M^2 \sin^2 \frac{\pi i}{2M}, i = 1, 2, \dots, M-1$ , принадлежат отрезку

$$a \leq 0,5\pi^2 \leq \mu \leq 2M^2 = b. \quad (24)$$

Итак, для выполнения (22) достаточно ввиду (23), чтобы для каждого значения  $\mu$  из отрезка (24) хотя бы при одном  $\tau, \tau = \tau_1, \tau_2, \dots, \tau_k$ , выполнялись неравенства

$$-\sqrt{q} < \frac{1 - \tau_p \mu}{1 + \tau_p \mu} < \sqrt{q},$$

и тем более достаточно, чтобы выполнялись неравенства

$$-\sqrt{q} \leq 1 - \tau_p \mu \leq \sqrt{q}.$$

Для этого нужно, чтобы для каждого  $\mu$  из отрезка (24) при некотором  $\tau_p$ ,  $p = 1, 2, \dots, k$ , выполнялись неравенства

$$1 - \sqrt{q} \leq \tau_p \mu \leq 1 + \sqrt{q}. \quad (25)$$

Зададим  $\mu_p$  и  $\tau_p$  соответственно формулами

$$\begin{aligned} \mu_p &= \left( \frac{1 + \sqrt{q}}{1 - \sqrt{q}} \right)^{p-1} a, \quad p = 1, 2, \dots, k, \\ \tau_p &= \frac{1 - \sqrt{q}}{\mu_p}, \quad p = 1, 2, \dots, k. \end{aligned} \quad (26)$$

Тогда при возрастании  $\mu$  от  $\mu_p$  до  $\mu_{p+1}$  число  $\tau_p \mu$  пробегает отрезок (25).

Очевидно, что, выбрав  $k$  из условия  $\mu_k \geq b$ , т. е.

$$\left. \begin{aligned} k &\geq A \ln \frac{b}{a} + 1 = A \left( 2 \ln M + \ln \frac{4}{\pi^2} \right) + 1, \\ A &= \frac{1}{\ln \frac{1 + \sqrt{q}}{1 - \sqrt{q}}}, \end{aligned} \right\} \quad (27)$$

мы и получим по формуле (26) интересующую нас последовательность  $\tau_1, \tau_2, \dots, \tau_k$ .

### ЗАДАЧИ

1. Можно ли выбрать итерационные параметры  $\tau_1, \tau_2, \dots, \tau_k$  так, чтобы за конечное число шагов итерационного процесса (4) получилось точное решение разностной задачи Дирихле?

Сколько для этого надо сделать итераций? Выдерживает ли такой прием решения обобщение на случай, когда точные значения собственных чисел  $\mu_r$  неизвестны?

2. Объяснить механизм вычислительной неустойчивости, развивающейся при расчете по формулам (16) при

$$x^{(k)} = (k, k-1, k-2, \dots, 2, 1)$$

и больших значениях  $k$  и  $M$ .

Какие гармоники  $\psi^{(r, s)}$  будут преобладать в ряде Фурье для погрешности  $\varepsilon^k$ , полученной в результате расчета при  $x^{(k)} = (k, k-1, \dots, 1)$  с ошибками округления?

3. Пусть  $A$  — самосопряженный оператор, собственные значения которого лежат на отрезке  $0 < \mu_{\min} < \mu < \mu_{\max}$ . Какому условию должно удовлетворять число обусловленности  $\eta = \mu_{\min}/\mu_{\max}$ , чтобы для решения уравнения  $Ax = \varphi$  сходилась и был вычислительно устойчив процесс Ричардсона

$$x^{p+1} = x^p - \tau_{p+1} (Ax^p - \varphi)$$

при произвольном выборе  $\tau_j = 1/\mu_j$ ,  $\mu_{\min} < \mu_j < \mu_{\max}$ ,  $j = 1, 2, \dots, k$ ,  $k$  произвольно?

4. Почему при использовании схемы Дугласа—Рэкфорда очередность использования параметров  $\tau_1, \tau_2, \dots, \tau_k$  не влияет существенным образом на вычислительную устойчивость итерационного процесса?

5. Считая, что затраты машинного времени на один шаг итерационного процесса Дугласа — Рэкфорда в двадцать раз больше, чем на один шаг процесса Ричардсона, прикинуть по формулам (15) и (27), при каких  $M$  преимущества метода Дугласа — Рэкфорда начинают проявляться.

### § 37. Метод Федоренко

В работе Р. П. Федоренко, ЖВМ и МФ 1, № 5 (1961), предложен метод итерационного решения разностных эллиптических задач, названный им релаксационным. Для уменьшения нормы первоначальной погрешности в  $\epsilon$  раз этот метод требует всего  $cM^2$  арифметических действий, где  $M$  — число шагов сетки по одному направлению, а  $c$  — некоторая постоянная, не зависящая от  $M$ . Напомним, что наиболее быстросходящийся среди рассмотренных выше (и вообще всех других известных) методов — метод Дугласа — Рэкфорда — требует для той же цели  $O((\ln M)M^2)$  арифметических действий.

Границы применимости метода Федоренко почти такие же, как у простейшего метода установления. Дополнительным ограничением является требование «плавности», «гладкости» первых собственных функций, которое для эллиптических задач обычно выполнено.

В простых примерах экономия машинного времени по сравнению с лучшими в смысле скорости сходимости другими итерационными методами убедительно подтверждается уже при  $M \approx 50$ . Надо иметь в виду, что логика организации расчета релаксационным методом существенно сложнее, как мы увидим, чем логика расчета, скажем, по схеме Ричардсона. Поэтому машинное время существенно зависит от качества программы.

Простейшая оценка скорости сходимости (для разностной аппроксимации уравнения Пуассона в квадратной области на квадратной сетке при заданных значениях решения на границе) была получена Р. П. Федоренко, ЖВМ и МФ 4, № 3 (1964).

В работе Н. С. Бахвалова, ЖВМ и МФ 6, № 5 (1966), исследована сходимость метода Федоренко и был получен тот же результат для разностного аналога первой краевой задачи в прямоугольнике для общего эллиптического уравнения с гладкими коэффициентами

$$a_{11} \frac{\partial^2 u}{\partial x^2} + 2a_{12} \frac{\partial^2 u}{\partial x \partial y} + a_{22} \frac{\partial^2 u}{\partial y^2} + a_1 \frac{\partial u}{\partial x} + a_2 \frac{\partial u}{\partial y} + au = f_0.$$

Наконец, Г. П. Астраханцев, ЖВМ и МФ 11, № 2 (1971), получил аналогичный результат для разностной аппроксимации третьей краевой задачи для самосопряженного эллиптического уравнения в произвольной двумерной области с гладкой границей.

Обоснования довольно громоздки, поэтому мы ограничимся качественным описанием идеи метода и самого алгоритма Федоренко, отсылая за доказательствами к оригинальным работам и обзорной статье Р. П. Федоренко, УМН 28, в. 2 (1973).

1. Идея метода. При решении итерациями задачи

$$\left. \begin{aligned} \Delta_h u_{mn} - \varphi(x_m, y_n) &= 0, \quad m, n = 1, 2, \dots, M-1, \\ u_{mn}|_{\Gamma} &= \psi(s_{mn}) \end{aligned} \right\} \quad (1)$$

будем отправляться от простейшего процесса установления (4) § 35

$$\left. \begin{aligned} u_{mn}^{p+1} &= u_{mn}^p + \tau (\Delta_h u_{mn}^p - \varphi(x_m, y_n)), \quad m, n = 1, 2, \dots, M-1, \\ u_{mn}^{p+1}|_{\Gamma} &= \psi(s_{mn}), \quad \{u_{mn}^0\} \text{ задано,} \end{aligned} \right\} \quad (2)$$

который в целом сходится очень медленно, но неравномерно на различных гармониках. Погрешность  $\varepsilon^p = u^p - u$  в соответствии с (10) § 35 записывается в виде конечного ряда Фурье

$$\varepsilon^p = \sum_{r, s=1}^{M-1} [\lambda_{rs}(\tau)]^p c_{rs}^0 \psi^{(r, s)}, \quad (3)$$

где  $c_{rs}^0$  — коэффициенты разложения погрешности  $\varepsilon^0 = u^0 - u$  нулевого приближения,  $\lambda_{rs} = 1 - 4\tau M^2 \left( \sin^2 \frac{\pi r}{2M} + \sin^2 \frac{\pi s}{2M} \right)$ . Числа  $\lambda_{rs}$  лежат на отрезке  $\lambda_{\text{лев}} \leq \lambda \leq \lambda_{\text{прав}}$ , где

$$\begin{aligned} \lambda_{\text{лев}} &= \lambda_{M-1, M-1} \approx 1 - 8\tau M^2, \\ \lambda_{\text{прав}} &= \lambda_{1, 1} \approx 1 - 2\pi^2 \tau. \end{aligned}$$

Положим

$$\tau = \frac{3}{16M^2}. \quad (4)$$

Если при этом условии хотя бы одно из чисел  $r$  или  $s$  больше, чем  $M/2$ , то  $|\lambda_{rs}| < 0,6$ . (5)

Поэтому вклад высокочастотных гармоник  $\psi^{(r, s)}$ ,  $r > M/2$  или  $s > M/2$  в погрешность (3) за один шаг итерационного процесса уменьшается почти вдвое и вскоре становится малым. После нескольких итераций по формуле (2) погрешность  $\varepsilon^p$  будет существенно содержать только гладкую компоненту (гармоники  $\psi^{(r, s)}$ ,  $r < M/2$ ,  $s < M/2$ ), потому что низкочастотные гармоники  $\psi^{(r, s)}$  умножаются на числа  $\lambda_{rs}$ , которые ближе к единице. Очень медленно гасится вклад первой гармоники  $\psi^{(1, 1)}$ : при сделанном выборе  $\tau$

$$\lambda_{1, 1} \approx 1 - \frac{3\pi^2}{8M^2} (\approx 1). \quad (6)$$

Обозначим полученное в процессе итераций (2) приближение  $u^p$  через  $U$ , а погрешность  $\varepsilon^p = u^p - u = U - u$  через  $v$ . Если бы мы знали погрешность  $v$ , то нашли бы искомое решение  $u = U - v$ . Однако относительно  $v$  мы знаем только, что оно удовлетворяет уравнению

$$\Delta_h v = \xi, \quad v|_{\Gamma} = 0, \quad (7)$$

где  $\xi$  — известная сеточная функция — это невязка, возникающая при подстановке  $u^p = U$  в уравнение (1):

$$\xi = \Lambda_h u^p - \varphi = \Lambda_h u^p - \Lambda_h u = \Lambda_h (u^p - u) = \Lambda_h v.$$

Задача (7) для определения поправки  $v$  проще исходной задачи (1) лишь в том отношении, что относительно  $v$  заранее известно, что это гладкая сеточная функция. Поэтому для определения  $v$  вместо задачи (7) можно приближенно рассматривать такую же задачу на вдвое более крупной сетке, которая (при четном  $M$ ) является подсеткой исходной сетки:

$$\Lambda_{2h} v^* = \xi^*, \quad v^* |_{\Gamma^*} = 0. \quad (1^*)$$

Звездочкой мы обозначили величины на укрупненной сетке. Задачу (1\*) решаем итерациями по формулам

$$\left. \begin{aligned} (v_{mn}^*)^{p+1} &= (v_{mn}^*)^p + \tau^* [\Lambda_{2h} (v_{mn}^*)^p - \xi_{mn}^*], \\ m, n &= 1, 2, \dots, M^* - 1, \\ (v_{mn}^*)^{p+1} |_{\Gamma^*} &= 0 \end{aligned} \right\} \quad (2^*)$$

приняв за нулевое приближение  $(v_{mn}^*)^0 \equiv 0$ . Здесь  $M^* = M/2$ ,  $\tau^* = \frac{3}{16(M^*)^2} = 4\tau$ .

Каждый шаг итераций (2\*) вчетверо менее трудоемок, чем шаг итераций (2), потому что расчетных точек вчетверо меньше. Кроме того, благодаря  $\tau^* = 4\tau$  быстрее происходит погашение самой медленно убывающей компоненты ошибки. В соответствии с (6)

$$\lambda_{1,1}^* = 1 - \frac{3\pi^2}{8(M^*)^2} = 1 - 4 \frac{3\pi^2}{8M^2} < \lambda_{1,1},$$

и для погашения вклада  $\psi^{(1,1)}$  в  $e$  раз нужно вчетверо меньше итераций. Результат итераций по формуле (2\*) обозначим  $V^*$ . Проинтерполируем  $V^*$  на исходную сетку (линейно). Гладкие компоненты будут получены почти правильно. Возникающая при интерполяции погрешность будет относительно интерполируемой гладкой функции мала, но ее фурье-разложение будет содержать все гармоники (погрешность интерполяции негладкая из-за изломов в узлах интерполяции). Кроме того, негладкая компонента  $V^*$ , не имеющая отношения к искомой поправке, при интерполяции тоже дает случайный вклад в негладкую составляющую полученной при интерполяции функции  $V$ . Итак, гладкая компонента разности  $U - V$  близка к гладкой компоненте искомого решения  $u = U - v$ , но негладкая компонента не очень мала и носит случайный характер.

Поэтому следует проделать еще несколько шагов исходного итерационного процесса (2), приняв  $U - V$  за начальное приближение. Это быстро погасит привнесенную при интерполяции негладкую составляющую погрешности, которую итерации (2) гасят почти вдвое за один шаг.

**2. Описание алгоритма.** Ускорение сходимости, достигнутое за счет использования укрупненной сетки и итерационного процесса (2\*), может оказаться недостаточным. При большом  $M$  (мелкой сетке) задача (1\*) на укрупненной сетке все еще трудоемка. Поэтому при ее решении в свою очередь целесообразно проделать еще одно укрупнение сетки вдвое, при решении задачи на вчетверо укрупненной сетке снова использовать процесс удвоения сетки и увеличения  $\tau$  и т. д. В экспериментах Р. П. Федоренко сетка укрупняется не вдвое, а втрое. При  $M \approx 100$  оказывается достаточным двух укрупнений. Будем считать для простоты, что  $M = 2^k$ , т. е.  $M$  является некоторой степенью двойки.

На исходной сетке делаем несколько шагов итераций (2), чтобы «выгладить» погрешность. Погрешность нам неизвестна, поэтому можно следить за этим по невязке  $\Delta_h u^p - \varphi$ , которая тоже выглаживается. Результат вычислений  $U = u^p$  запоминаем. Затем для поправки  $v$  рассматриваем задачу на укрупненной сетке, делаем несколько итераций (2\*), чтобы выгладить «поправку к поправке» и результат  $V^*$  запоминаем (он занимает вчетверо меньше места в памяти, чем  $U$ ). Для вычисления поправки к  $V^*$  рассматриваем задачу на еще вдвое укрупненной сетке, делаем несколько итераций с шагом  $\tau^{**} = 4\tau^* = 16\tau$  и запоминаем результат  $V^{**}$ . Этот процесс вычисления поправок к поправкам на вдвое укрупненных сетках продолжаем  $k$  раз, пока не дойдем до самой крупной сетки и поправки  $V^{(k*)}$ .

Затем начинаем возвращение на мелкую сетку. Сначала с самой крупной сетки интерполируем полученную там последнюю поправку  $V^{(k*)}$  на сетку вдвое более мелкую, вносим эту проинтерполированную поправку в  $V^{((k-1)*)}$  и делаем несколько итераций, чтобы погасить привнесенную при интерполяции ошибку. Результат этих итераций интерполируем на еще вдвое более мелкую сетку, уточняем с его помощью хранящуюся для этой сетки поправку  $V^{((k-2)*)}$ , делаем несколько итераций и производим следующую интерполяцию. На предпоследнем шаге после внесения в  $V^*$  поправки и итераций получим поправку  $V^*$ , которую интерполируем на исходную сетку. Прделав несколько итераций (2) над  $U - V$ , получим результат.

## ПОНЯТИЕ О ВАРИАЦИОННО-РАЗНОСТНЫХ И ПРОЕКЦИОННО-РАЗНОСТНЫХ СХЕМАХ

В этой главе мы изложим способ построения разностных схем, основанный на использовании той или иной вариационной или проекционной постановки краевой задачи, решение которой требуется численно найти. Этот способ, называемый иногда методом конечных элементов, позволяет строить пригодные разностные схемы на нерегулярных сетках, а также при меньших предположениях о гладкости искомого решения и коэффициентов уравнения. Благодаря появляющейся свободе в выборе сеток узлы можно располагать гуще в тех частях области определения искомого решения, где решение ведет себя более сложно или где нас интересуют более мелкие детали его поведения.

Возможность целесообразно располагать узлы позволяет достигать требуемой точности при меньшем числе узлов сетки.

Метод конечных элементов можно интерпретировать как одну из возможных конкретизаций классических вариационных методов решения краевых задач. Поэтому мы сначала (§ 38) расскажем о классических вариационном и проекционном методах, а затем (§ 39) о вариационно-разностных схемах.

### § 38. Вариационные и проекционные методы

**1. Вариационная постановка краевых задач.** Многие дифференциальные краевые задачи математической физики допускают естественные вариационные постановки.

Мы ограничимся рассмотрением двух простых примеров таких задач и их вариационных постановок, иллюстрирующих, однако, суть дела. В этих примерах речь пойдет о различных краевых задачах для уравнения Пуассона в некоторой ограниченной области  $D$  плоскости  $x, y$  с кусочно-гладкой границей  $\Gamma$ .

Обозначим через  $\mathcal{W}$  линейное пространство всех непрерывных в области  $D$  и на ее границе  $\Gamma$  функций, обладающих также

ограниченными производными первого порядка, которые могут иметь разрывы лишь на конечном множестве прямых (для каждой функции  $w(x, y)$  — своих). Введем в линейном пространстве  $\mathcal{W}$  норму, положив для каждой функции  $w \in \mathcal{W}$

$$\|w\|_{\mathcal{W}} = \left[ \iint_D w^2 dx dy + \iint_D \left[ \left( \frac{\partial w}{\partial x} \right)^2 + \left( \frac{\partial w}{\partial y} \right)^2 \right] dx dy \right]^{1/2}. \quad (1)$$

Пополнение пространства  $\mathcal{W}$  приводит к полному пространству  $\mathcal{W}_2^1$  С. Л. Соболева.

Переходим к рассмотрению примеров.

**Пример 1.** Рассмотрим первую краевую задачу (задачу Дирихле)

$$\left. \begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= f(x, y), & (x, y) \in D, \\ u|_{\Gamma} &= \varphi(s), \end{aligned} \right\} \quad (A)$$

где  $s$  — длина дуги вдоль границы  $\Gamma$  области  $D$ , а  $f(x, y)$ ,  $\varphi(s)$  — заданные функции, удовлетворяющие всем условиям того, чтобы решение  $u(x, y)$  задачи (A) имело непрерывные вторые производные всюду в области  $D$  и на ее границе  $\Gamma$ .

**Теорема 1.** Среди всех функций  $w \in \mathcal{W}$ , удовлетворяющих граничному условию

$$w|_{\Gamma} = \varphi(s), \quad (2)$$

решение  $u(x, y)$  задачи (A) придает выражению (функционалу)

$$I(w) \equiv \iint_D \left[ \left( \frac{\partial w}{\partial x} \right)^2 + \left( \frac{\partial w}{\partial y} \right)^2 + 2fw \right] dx dy \quad (3)$$

наименьшее численное значение.

**Доказательство.** Пусть  $w(x, y) \in \mathcal{W}$ ,  $w|_{\Gamma} = \varphi(s)$ , — некоторая фиксированная функция. Введем обозначение  $\xi(x, y) \equiv w(x, y) - u(x, y)$ , так что

$$w(x, y) = u(x, y) + \xi(x, y).$$

Поскольку  $u(x, y)$  имеет непрерывные вторые производные, а  $w(x, y) \in \mathcal{W}$ , то также  $\xi(x, y) \in \mathcal{W}$ , причем  $\xi|_{\Gamma} = 0$ . Докажем равенство

$$I(w) \equiv I(u + \xi) = I(u) + \iint_D \left[ \left( \frac{\partial \xi}{\partial x} \right)^2 + \left( \frac{\partial \xi}{\partial y} \right)^2 \right] dx dy, \quad (4)$$

из которого следует справедливость теоремы, поскольку в случае  $w(x, y) \neq u(x, y)$  функция  $\xi(x, y)$  не обращается тождественно в нуль, так что второе слагаемое в правой части формулы



(4) строго положительно, и  $I(w) > I(u)$ . Очевидно,

$$I(u + \xi) \equiv \iint_D \left[ \left( \frac{\partial u}{\partial x} + \frac{\partial \xi}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} + \frac{\partial \xi}{\partial y} \right)^2 + 2f(u + \xi) \right] dx dy =$$

$$= I(u) + \iint_D \left[ \left( \frac{\partial \xi}{\partial x} \right)^2 + \left( \frac{\partial \xi}{\partial y} \right)^2 \right] dx dy +$$

$$+ 2 \iint_D \left( \frac{\partial u}{\partial x} \frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \xi}{\partial y} + f\xi \right) dx dy.$$

Остается проверить, что третье слагаемое в правой части обращается в нуль. Действительно, из очевидных тождеств

$$\frac{\partial u}{\partial x} \frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \xi}{\partial y} \equiv \frac{\partial}{\partial x} \left( \xi \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left( \xi \frac{\partial u}{\partial y} \right) -$$

$$- \xi \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \equiv \frac{\partial}{\partial x} \left( \xi \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left( \xi \frac{\partial u}{\partial y} \right) - \xi f \quad (5)$$

следует

$$\iint_D \left( \frac{\partial u}{\partial x} \frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \xi}{\partial y} + f\xi \right) dx dy =$$

$$= \iint_D \left[ \frac{\partial}{\partial x} \left( \xi \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left( \xi \frac{\partial u}{\partial y} \right) \right] dx dy = \int_{\Gamma} \xi \frac{\partial u}{\partial n} ds = 0, \quad (6)$$

где  $\frac{\partial u}{\partial n}$  — производная по внутренней нормали.

В предпоследнем переходе в цепочке равенств (6) мы воспользовались теоремой из векторного анализа, в силу которой интеграл от дивергенции векторного поля по области равен потоку этого векторного поля через границу области. В данном случае этот поток  $\int_{\Gamma} \xi \frac{\partial u}{\partial n} ds$  обращается в нуль, так как  $\xi|_{\Gamma} =$

$= 0$ . Теорема доказана.

Таким образом, задача (A) допускает следующую вариационную постановку: *среди всех функций класса  $W$ , удовлетворяющих условию (2), найти ту, которая придает наименьшее значение функционалу  $I(w)$ , определенному формулой (3).*

Пример 2. Рассмотрим третью краевую задачу

$$\left. \begin{aligned} \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} &= f(x, y), & (x, y) \in D, \\ \frac{\partial v}{\partial n} + \sigma(s) v|_{\Gamma} &= \varphi(s), \end{aligned} \right\} \quad (B)$$

где  $f(x, y)$ ,  $\varphi(s)$  и  $\sigma(s) \geq \sigma_0 > 0$  — заданные функции, а  $\frac{\partial v}{\partial n}$  — производная в направлении внутренней нормали.

**Теорема 2.** Среди всех функций  $w \in W$  решение  $v$  задачи (B) придает функционалу

$$J(w) \equiv \iint_D \left[ \left( \frac{\partial w}{\partial x} \right)^2 + \left( \frac{\partial w}{\partial y} \right)^2 + 2fw \right] dx dy + \int_{\Gamma} [\sigma(s) w^2 - 2\varphi(s) w] ds \quad (7)$$

наименьшее значение.

**Доказательство.** Пусть  $w \in W$  — какая-нибудь фиксированная функция. Обозначим

$$\eta(x, y) \equiv w(x, y) - v(x, y).$$

Докажем равенство

$$J(w) \equiv J(v + \eta) = J(v) + \left\{ \iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy + \int_{\Gamma} \sigma(s) \eta^2 ds \right\}, \quad (8)$$

из которого следует, что в случае  $w \neq v$ , т. е.  $\eta \neq 0$  выполнено неравенство  $J(w) > J(v)$ , справедливость которого утверждает-ся в теореме.

Очевидно,

$$\begin{aligned} J(w) &\equiv J(v + \eta) = \\ &= \iint_D \left[ \left( \frac{\partial v}{\partial x} + \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial y} + \frac{\partial \eta}{\partial y} \right)^2 + 2f(v + \eta) \right] dx dy + \\ &\quad + \int_{\Gamma} [\sigma(s) (v + \eta)^2 - 2\varphi(s) (v + \eta)] ds = \\ &= J(v) + \left\{ \iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy + \int_{\Gamma} \sigma(s) \eta^2 ds \right\} + \\ &+ 2 \left\{ \iint_D \left[ \frac{\partial v}{\partial x} \frac{\partial \eta}{\partial x} + \frac{\partial v}{\partial y} \frac{\partial \eta}{\partial y} + f\eta \right] dx dy + \int_{\Gamma} [\sigma(s) v - \varphi(s)] \eta ds \right\}. \quad (9) \end{aligned}$$

Остается показать, что выражение, стоящее в правой части (9) во вторых фигурных скобках, обращается в нуль. Действительно, преобразовывая двойной интеграл в этом выражении аналогично (6), получим

$$\begin{aligned} &\iint_D \left[ \frac{\partial v}{\partial x} \frac{\partial \eta}{\partial y} + \frac{\partial v}{\partial y} \frac{\partial \eta}{\partial x} + f\eta \right] dx dy + \int_{\Gamma} [\sigma(s) v - \varphi(s)] \eta ds = \\ &= \int_{\Gamma} \frac{\partial v}{\partial n} \cdot \eta ds + \int_{\Gamma} [\sigma(s) v - \varphi(s)] \eta ds = \int_{\Gamma} \eta \left[ \frac{\partial v}{\partial n} + \sigma v - \varphi \right] ds = 0, \end{aligned}$$

поскольку  $\frac{\partial v}{\partial n} + \sigma v|_{\Gamma} \equiv \varphi$ . Теорема доказана.

Таким образом, третья краевая задача для уравнения Пуассона (В) допускает следующую вариационную постановку: *среди всех функций  $\omega \in W$  найти ту, которая придает наименьшее значение функционалу  $J(\omega)$* , введенному равенством (7).

Обратим внимание на то, что различие в вариационных постановках краевых задач (А) и (В) не исчерпывается различием в функционалах  $I(\omega)$  и  $J(\omega)$ . При минимизации функционала  $J(\omega)$  допустимыми считаются все функции  $\omega \in W$ , а при минимизации функционала  $I(\omega)$  — лишь те функции  $\omega \in W$ , которые удовлетворяют краевому условию  $\omega|_{\Gamma} = \varphi(s)$  задачи (А).

Это различие дало повод называть краевое условие задачи (В) *естественным*, поскольку при вариационной постановке оно не накладывает никаких ограничений на допустимые функции.

**2. Сходимость минимизирующих последовательностей.** Точным решением задачи (А), как мы видели, является та функция  $\omega(x, y) \equiv u(x, y)$ , которая придает наименьшее значение функционалу  $I(\omega)$  среди всех *допустимых* функций, т. е. функций  $\omega \in W$ , удовлетворяющих условию  $\omega|_{\Gamma} = \varphi(s)$ .

Численное решение вариационной задачи об отыскании  $u(x, y)$  состоит в построении такой функции  $\omega \in W$ ,  $\omega|_{\Gamma} = \varphi(s)$ , которая придает функционалу  $I(\omega)$  хотя и не наименьшее значение  $I(u)$ , но «близкое» к наименьшему. Точнее, для вычисления должен быть указан способ построения членов такой последовательности допустимых функций  $\omega_N(x, y) \in W$ ,  $\omega_N|_{\Gamma} = \varphi(s)$ , для которой

$$\lim_{N \rightarrow \infty} I(\omega_N) = I(u).$$

Такая последовательность допустимых функций называется *минимизирующей*. Выбирая член  $\omega_N(x, y)$  минимизирующей последовательности с достаточно большим номером  $N$ , можно добиться того, чтобы  $I(\omega_N)$  отличалось от  $I(u)$  сколь угодно мало.

Совершенно аналогично для вариационной постановки задачи (В) минимизирующей последовательностью допустимых функций будет всякая последовательность  $\omega_N(x, y) \in W$ , для которой

$$\lim_{N \rightarrow \infty} J(\omega_N) = J(v).$$

Способы построения минимизирующих последовательностей для вариационных задач (метод Ритца и вариационно-разностные схемы) будут указаны в этой главе ниже.

Здесь мы докажем только, что минимизирующие последовательности сходятся квадратически в среднем вместе с первыми производными к решениям  $u$  и  $v$  соответствующих вариационных задач, так что их члены можно считать приближениями к

решениям. Именно, будут доказаны следующие два предложения.

**Теорема 3.** Пусть  $w \in W$ ,  $w|_{\Gamma} = \varphi(s)$ . Тогда

$$\|w - u\|_W^2 \leq \alpha [I(w) - I(u)], \quad (10)$$

где  $\alpha$  — некоторая постоянная, полностью определяемая формой области  $D$  и не зависящая от функции  $w$ .

**Теорема 4.** Пусть  $w$  — произвольная функция из  $W$ . Тогда

$$\|w - v\|_W^2 \leq \beta [J(w) - J(v)], \quad (11)$$

где постоянная  $\beta > 0$  зависит только от формы области  $D$  и от числа  $\min \sigma(s) = \sigma_0 > 0$ , но не от функции  $w$ .

Из неравенств (10) и (11), очевидно, следует сходимость

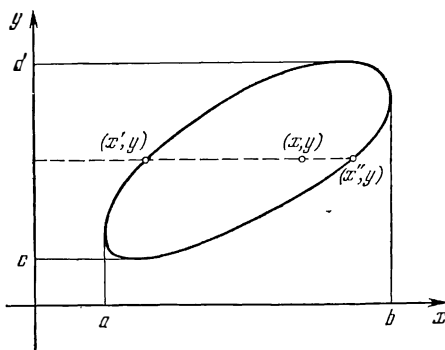


Рис. 46.

соответствующих вариационным постановкам краевых задач (А) и (В) минимизирующих последовательностей к  $u$  и  $v$  соответственно, поскольку при замене  $w$  членом  $w_N$  соответствующей минимизирующей последовательности правые, а значит, и левые части неравенств (10) и (11) при  $N \rightarrow \infty$  стремятся к нулю.

Доказательство теорем 3 и 4 основано на следующей лемме.

**Лемма.** Пусть функция  $\eta(x, y) \in W$  и  $\sigma(s) \geq \sigma_0 > 0$ . Тогда справедливо следующее неравенство:

$$\iint_D \eta^2 dx dy \leq \bar{\beta} \left\{ \iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy + \int_{\Gamma} \sigma(s) \eta^2 ds \right\}. \quad (12)$$

Здесь  $\bar{\beta}$  — некоторая постоянная, которая полностью определяется областью  $D$  и числом  $\sigma_0$  и не зависит от функции  $\eta(x, y) \in W$ .

Доказательство неравенства (12) проведем при дополнительном предположении, что каждая прямая  $y = \text{const}$  пересекает границу  $\Gamma$  области  $D$  не более чем в двух точках. Это предположение не вызвано существом дела, но благодаря ему доказательство становится коротким.

Пусть  $x, y$  — точка внутри области  $D$  (рис. 46). Тогда

$$\eta(x, y) = \eta(x', y) + \int_{x'}^x \frac{\partial \eta(t, y)}{\partial t} dt. \quad (13)$$

Возведем обе части неравенства (13) в квадрат и воспользуемся очевидным неравенством  $2AB \leq A^2 + B^2$ , справедливым для любых двух чисел  $A$  и  $B$ :

$$\eta^2(x, y) \leq 2 \left[ \eta^2(x', y) + \left( \int_{x'}^x \frac{\partial \eta(t, y)}{\partial t} dt \right)^2 \right]. \quad (14)$$

Воспользуемся неравенством Буняковского

$$\begin{aligned} \left( \int_{x'}^x \frac{\partial \eta(t, y)}{\partial t} dt \right)^2 &\leq \left( \int_{x'}^x 1^2 \cdot dt \right) \cdot \left[ \int_{x'}^x \left( \frac{\partial \eta(t, y)}{\partial t} \right)^2 dt \right] \leq \\ &\leq (b-a) \int_{x'}^{x''} \left( \frac{\partial \eta(x, y)}{\partial x} \right)^2 dx. \end{aligned} \quad (15)$$

Из (14) и (15) получим

$$\eta^2(x, y) \leq 2 \left[ \eta^2(x', y) + (b-a) \int_{x'}^{x''} \left( \frac{\partial \eta(x, y)}{\partial x} \right)^2 dx \right]. \quad (16)$$

Принтегрируем обе части (16) по  $x$  в пределах от  $x = x' = x'(y)$  до  $x = x'' = x''(y)$ , воспользовавшись тем, что правая часть от  $x$  не зависит:

$$\begin{aligned} \int_{x'}^{x''} \eta^2(x, y) dx &\leq 2(x'' - x') \left[ \eta^2(x', y) + (b-a) \int_{x'}^{x''} \left( \frac{\partial \eta(x, y)}{\partial x} \right)^2 dx \right] \leq \\ &\leq 2(b-a) \left[ \eta^2(x', y) + (b-a) \int_{x'}^{x''} \left( \frac{\partial \eta(x, y)}{\partial x} \right)^2 dx \right]. \end{aligned} \quad (17)$$

Принтегрируем теперь обе части неравенства (17) по  $y$  в пределах от  $y = c$  до  $y = d$ . Получим

$$\iint_D \eta^2 dx dy \leq 2(b-a) \left[ \int_c^d \eta^2(x', y) dy + (b-a) \iint_D \left( \frac{\partial \eta}{\partial x} \right)^2 dx dy \right]. \quad (18)$$

Очевидно, что

$$\int_c^d \eta^2(x', y) dy \leq \int_{\Gamma} \eta^2(s) ds \leq \frac{1}{\sigma_0} \int_{\Gamma} \sigma(s) \eta^2 ds; \quad (19)$$

$$\iint_D \left( \frac{\partial \eta}{\partial x} \right)^2 dx dy \leq \iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy. \quad (20)$$

Из (18), (19) и (20) следует

$$\begin{aligned} \iint_D \eta^2 dx dy &\leq \\ &\leq 2(b-a) \left[ \frac{1}{\sigma_0} \int_{\Gamma} \sigma \eta^2 ds + (b-a) \iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy \right] \leq \\ &\leq 2(b-a) \max \left[ \frac{1}{\sigma_0}, b-a \right] \cdot \\ &\quad \left\{ \iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy + \int_{\Gamma} \sigma(s) \eta^2 ds \right\}, \end{aligned}$$

г. е. неравенство (12), причем за постоянную  $\bar{\beta}$  можно принять число  $\bar{\beta} = 2(b-a) \max \left[ \frac{1}{\sigma_0}, b-a \right]$ .

Следствие. Пусть  $\xi(x, y) \in W$  и  $\xi|_{\Gamma} = 0$ . Тогда справедливо следующее неравенство Фридрихса:

$$\iint_D \xi^2 dx dy \leq \bar{\alpha} \iint_D \left[ \left( \frac{\partial \xi}{\partial x} \right)^2 + \left( \frac{\partial \xi}{\partial y} \right)^2 \right] dx dy, \quad (21)$$

где  $\bar{\alpha} = 2(b-a)^2$ .

Доказательство. Положим  $\sigma(s) \equiv \sigma_0 = \frac{1}{b-a}$ . Для функций  $\xi(x, y) \equiv \eta(x, y) \in W$ , удовлетворяющих дополнительному условию  $\xi|_{\Gamma} = \eta|_{\Gamma} = 0$ , неравенство (12) примет вид (21), где  $\bar{\alpha} = \bar{\beta} = 2(b-a)^2$ .

Доказательство теоремы 3. Для всякой функции  $w \in W$ ,  $w|_{\Gamma} = \varphi(s)$ , функция  $\xi = w - u$  удовлетворяет условиям следствия из леммы, а потому и неравенству (21). Если учесть равенство (4)

$$\iint_D \left[ \left( \frac{\partial \xi}{\partial x} \right)^2 + \left( \frac{\partial \xi}{\partial y} \right)^2 \right] dx dy = I(w) - I(u), \quad (4')$$

то (21) можно записать в форме

$$\iint_D (w-u)^2 dx dy \leq \bar{\alpha} [I(w) - I(u)]. \quad (21')$$

Прибавляя к неравенству (21') равенство (4'), получим неравенство (10), в котором  $\alpha = \bar{\alpha} + 1$ . Теорема 3 доказана.

Доказательство теоремы 4. Для всякой функции  $w \in W$  функция  $\eta = w - v$  удовлетворяет условию леммы, а значит, и неравенству (12). Если учесть равенство (8)

$$\iint_D \left[ \left( \frac{\partial \eta}{\partial x} \right)^2 + \left( \frac{\partial \eta}{\partial y} \right)^2 \right] dx dy + \int_{\Gamma} \sigma(s) \eta^2 ds = J(w) - J(v), \quad (8')$$

то (12) можно записать в форме

$$\iint_D (\omega - v)^2 dx dy \leq \bar{\beta} [J(\omega) - J(v)]. \quad (12')$$

Складывая (8') и (12') почленно и отбрасывая в левой части неотрицательное слагаемое  $\int_{\Gamma} \sigma(s) \eta^2 ds$ , получим неравенство (11) с постоянной  $\beta = \bar{\beta} + 1$ . Теорема 4 доказана.

**3. Вариационный метод Ритца.** Из теорем 3 и 4 в силу неравенств (10) и (11) следует, что приближениями для решений  $u$  и  $v$  краевых задач (A) и (B) могут служить те функции  $\omega$  из числа допустимых ( $\omega \in W$  и  $\omega|_{\Gamma} = \varphi(s)$  для задачи (A) и  $\omega \in W$  для задачи (B)), на которых функционалы  $I(\omega)$  и  $J(\omega)$  принимают значения, близкие к минимальным значениям этих функционалов  $I(u)$  и  $J(v)$  на соответствующих классах допустимых функций.

Для фактического отыскания приближенных решений Ритцем в 1908 г. был предложен прием, который мы изложим сначала применительно к задаче (A). Для удобства изложения будем считать, что в краевом условии (A)  $\varphi(s) = 0$ , так что  $u|_{\Gamma} = 0$ . К этому приводится общий случай  $\varphi(s) \neq 0$  путем перехода к искомой функции  $\tilde{u}$ ,  $u = \tilde{u} + h$ , где  $h(x, y)$  — какая-нибудь дважды непрерывно дифференцируемая функция, удовлетворяющая краевому условию  $h|_{\Gamma} = \varphi(s)$ . Формальная схема отыскания приближенного решения по методу Ритца состоит в следующем. Обозначим через  $\hat{W}$  линейное подпространство всех функций  $\omega \in W$ , удовлетворяющих граничному условию  $\omega|_{\Gamma} = 0$ . Зададим натуральное  $N$  и фиксируем какие-нибудь  $N$  линейно-независимых функций

$$\omega_1^N(x, y), \omega_2^N(x, y), \dots, \omega_N^N(x, y), \quad (22)$$

удовлетворяющих условию

$$\omega_n^N|_{\Gamma} = 0, \quad n = 1, 2, \dots, N. \quad (23)$$

Рассмотрим теперь  $N$ -мерное линейное пространство  $\hat{W}^N$  всевозможных линейных комбинаций функций (22)

$$\omega_N(x, y, a_1, \dots, a_N) = \sum_{n=1}^N a_n \omega_n^N,$$

где  $a_1, \dots, a_N$  — произвольные вещественные числа.

Будем искать теперь вместо функции  $\omega \equiv u$ , придающей минимум функционалу  $I(\omega)$  на пространстве  $\hat{W}$ , такую функцию  $\omega_N(x, y, a_1, \dots, a_N)$ , которая придает минимум функционалу  $I(\omega)$  на множестве всех функций из  $N$ -мерного пространства

$\bar{W}^N$ . Эту функцию  $\omega_N(x, y, a_1, \dots, a_N) \equiv \bar{\omega}_N(x, y)$  и примем за приближенное решение при сделанном выборе  $N$  базисных функций (22). Задача об отыскании функции  $\bar{\omega}_N(x, y)$  несравненно проще задачи отыскания точного решения  $u(x, y)$ .

Действительно,

$$I[\omega_N(x, y, a_1, \dots, a_N)] \equiv \iint_D \left[ \left( \frac{\partial}{\partial x} \sum_{n=1}^N a_n \omega_n^N \right)^2 + \left( \frac{\partial}{\partial y} \sum_{n=1}^N a_n \omega_n^N \right)^2 \right] dx dy + 2 \iint_D f \left( \sum_{n=1}^N a_n \omega_n^N \right) dx dy, \quad (24)$$

и речь идет об отыскании  $N$  чисел  $a_1, \dots, a_N$ , придающих минимум функции  $I[\omega_N(x, y, a_1, \dots, a_N)]$  от  $N$  переменных. Покажем, что такой набор чисел  $a_1, \dots, a_N$  существует. Первое слагаемое в правой части выражения (24) есть квадратичная форма от  $a_1, \dots, a_N$ . Ввиду линейной независимости системы функций (22) эта форма при  $a_n \neq 0$  строго положительна, так как в противном случае она была бы при некотором наборе чисел  $\bar{a}_n \neq 0$  равна нулю и мы имели бы в силу (21)

$$\begin{aligned} \iint_D \left( \sum_{n=1}^N \bar{a}_n \omega_n^N \right)^2 dx dy &\leq \\ &\leq \bar{\alpha} \iint_D \left[ \frac{\partial}{\partial x} \left( \sum_{n=1}^N \bar{a}_n \omega_n^N \right)^2 + \frac{\partial}{\partial y} \left( \sum_{n=1}^N \bar{a}_n \omega_n^N \right) \right] dx dy = 0, \end{aligned}$$

откуда, вопреки линейной независимости,

$$\sum_{n=1}^N \bar{a}_n \omega_n^N(x, y) \equiv 0.$$

Из-за доказанной положительной определенности квадратичной формы выражение (24) имеет единственный минимум. Этот минимум достигается при тех значениях  $a_n = \bar{a}_n$ ,  $n = 1, \dots, N$ , при которых

$$\frac{\partial I[\omega_N(x, y, a_1, \dots, a_N)]}{\partial a_n} = 0, \quad n = 1, \dots, N. \quad (25)$$

Подробно линейная система уравнений (25) относительно чисел  $a_1, \dots, a_N$  может быть записана в виде

$$\begin{aligned} \sum_{i=1}^N a_i \iint_D \left[ \frac{\partial \omega_i^N}{\partial x} \frac{\partial \omega_n^N}{\partial x} + \frac{\partial \omega_i^N}{\partial y} \frac{\partial \omega_n^N}{\partial y} \right] dx dy = \\ = - \iint_D f \omega_n^N dx dy, \quad n = 1, \dots, N. \quad (26) \end{aligned}$$



Для более краткой записи этой системы и удобства речи в дальнейшем наряду с нормированным пространством  $W$  рассмотрим линейное пространство  $\tilde{W}$ , состоящее из тех же функций, что и  $W$ , но со скалярным умножением  $(w', w'')$

$$(w', w'') \equiv \iint_D \left( \frac{\partial w'}{\partial x} \frac{\partial w''}{\partial x} + \frac{\partial w'}{\partial y} \frac{\partial w''}{\partial y} \right) dx dy + \int_{\Gamma} \sigma(s) w' w'' ds, \quad (27)$$

где  $\sigma(s) \geq \sigma_0 > 0$  — какая-нибудь фиксированная функция. Это скалярное умножение индуцирует норму  $\|w\|_{\tilde{W}}$  в пространстве  $\tilde{W}$  по формуле

$$\|w\|_{\tilde{W}}^2 = (w, w). \quad (28)$$

Обозначим  $\overset{\circ}{W}$  подпространство функций  $w \in \tilde{W}$ , удовлетворяющих условию  $w|_{\Gamma} = 0$ .

После введения скалярного умножения система (26) благодаря условию  $\omega_n^N|_{\Gamma} = 0$  примет вид

$$\sum_{i=1}^N a_i (\omega_i^N, \omega_n^N) = - \iint_D f \omega_n^N dx dy, \quad n = 1, \dots, N. \quad (29)$$

Заметим, что матрица системы (29)

$$\omega^N = \left\| \begin{array}{c} (\omega_1^N, \omega_1^N) \dots (\omega_1^N, \omega_N^N) \\ \dots \dots \dots \dots \dots \dots \dots \\ (\omega_N^N, \omega_1^N) \dots (\omega_N^N, \omega_N^N) \end{array} \right\|$$

есть матрица Грама системы линейно независимых функций (22). Как известно из курса линейной алгебры, ее определитель отличен от нуля.

Решение  $a_n = \bar{a}_n$ ,  $n = 1, \dots, N$ , системы (29) и доставляет функцию

$$\bar{w}_N(x, y) = w_N(x, y, \bar{a}_1, \dots, \bar{a}_N),$$

которую принимают за приближенное решение. Функция  $\bar{w}_N$  допускает простое геометрическое истолкование.

В силу (4) и (27) имеем

$$I(w_N) - I(u) = (w_N - u, w_N - u).$$

Далее

$$\begin{aligned} I(\bar{w}_N) - I(u) &= \min_{w \in \overset{\circ}{W}^N} [I(w) - I(u)] = \\ &= \min_{w_N \in \overset{\circ}{W}^N} (w - u, w - u) = (\bar{w}_N - u, \bar{w}_N - u). \end{aligned}$$

Таким образом,  $\bar{w}_N$  есть тот элемент линейного  $N$ -мерного пространства  $\mathcal{W}^N$ , натянутого на базис (22), который наименее уклоняется от  $u$  в смысле нормы (28), т. е.  $\bar{w}_N$  есть проекция решения  $u$  в подпространство  $\overset{\circ}{\mathcal{W}}^N$  в смысле скалярного произведения (27).

Мы закончили формальное изложение схемы Рунге для отыскания приближенного решения.

Выясним теперь, от чего зависит близость приближенного решения

$$\bar{w}_N = w_N(x, y, \bar{a}_1, \dots, \bar{a}_N),$$

найденного по методу Рунге, к точному решению  $u$  задачи (A), в которой мы условились считать  $\varphi(s) = 0$ . Понятно, что число  $\|\bar{w}_N - u\|_{\mathcal{W}}$  зависит от выбора базисных функций (22). Если бы, например, базисные функции (22) были выбраны так (невероятный случай!), чтобы функция  $u$  оказалась одной из функций  $N$ -мерного пространства  $\overset{\circ}{\mathcal{W}}^N$ , натянутого на базис (22), то приближенное решение  $\bar{w}_N$  совпало бы с точным решением  $u$ . В самом деле,

$$I(\bar{w}_N) - I(u) = \min_{w \in \overset{\circ}{\mathcal{W}}^N} (I(w) - I(u)) = I(u) - I(u) = 0$$

и в силу теоремы 3

$$\|\bar{w}_N - u\|_{\mathcal{W}} \leq \alpha (I(\bar{w}_N) - I(u)) = 0.$$

Однако функция  $u$  нам неизвестна, а известны только некоторые ее свойства, которыми обладает не только она, но и целый класс  $U$  функций. Например, пусть известно, что вторые производные функции  $u$  непрерывны и ограничены постоянной  $M$ . Тогда класс  $U$  состоит из всех дважды непрерывно дифференцируемых функций, вторые производные которых не превосходят  $M$  и которые удовлетворяют условию  $u|_{\Gamma} = 0$ .

Напомним, что для решения  $u$  и любого  $w \in \overset{\circ}{\mathcal{W}}$

$$I(w) - I(u) = (w - u, w - u) = \|w - u\|_{\overset{\circ}{\mathcal{W}}}^2$$

и в силу теоремы 3

$$\|w - u\|_{\overset{\circ}{\mathcal{W}}}^2 \leq \alpha \|w - u\|_{\overset{\circ}{\mathcal{W}}}^2.$$

Поэтому выбор базисных функций надо по возможности осуществить так, чтобы для каждой функции  $v \in U \subset \overset{\circ}{\mathcal{W}}$  нашлась

функция  $\omega_N \in \overset{\circ}{W}^N$ , «близкая» к ней, т. е. такая, для которой  $\|\omega_N - v\|_{\overset{\circ}{W}}$  «мала». Тогда, в частности, будет «мала» величина

$$\|\bar{\omega}_N - u\|_{\overset{\circ}{W}}^2 = \min_{\omega \in \overset{\circ}{W}^N} [I(\omega) - I(u)] = \min_{\omega \in \overset{\circ}{W}^N} (\omega - u, \omega - u),$$

а вместе с тем будет «мала» и величина  $\|\bar{\omega}_N - u\|_{\overset{\circ}{W}}$ :

$$\|\bar{\omega}_N - u\|_{\overset{\circ}{W}}^2 \leq \alpha \|\bar{\omega}_N - u\|_{\overset{\circ}{W}}^2.$$

Говоря точно, наилучшим был бы такой выбор функций (22), при котором число

$$K_N = K_N(U, \overset{\circ}{W}^N) = \sup_{v \in U} \min_{\omega \in \overset{\circ}{W}^N} \|\omega - v\|_{\overset{\circ}{W}} \quad (30)$$

было бы наименьшим возможным. Обозначим  $\kappa_N(U, \overset{\circ}{W})$  число

$$\kappa_N(U, \overset{\circ}{W}) = \inf_{W^0_N \subset W^0} K_N(U, \overset{\circ}{W}^N) = \inf_{\omega_1^N, \dots, \omega_N^N} \sup_{v \in U} \min_{\omega \in \overset{\circ}{W}^N} \|\omega - v\|_{\overset{\circ}{W}}. \quad (31)$$

Это число называется *N-мерным колмогоровским поперечником класса функций U относительно нормированного пространства  $\overset{\circ}{W} \subset \tilde{W}$* . Очевидно, что наилучшим выбором функций (22) был бы такой, при котором число (30) совпадало бы с поперечником А. Н. Колмогорова  $\kappa_N(U, \overset{\circ}{W})$ . При любом  $\varepsilon > 0$  существует, очевидно, набор базисных функций (22), для которых

$$\begin{aligned} I(\bar{\omega}_N) - I(u) &= \|\bar{\omega}_N - u\|_{\overset{\circ}{W}}^2 \leq \\ &\leq \sup_{\bar{u} \in U} \inf_{\omega \in \overset{\circ}{W}_N} \|\omega - \bar{u}\|_{\overset{\circ}{W}}^2 = K_N^2(U, \overset{\circ}{W}^N) \leq \kappa_N^2(U, \overset{\circ}{W}) + \varepsilon. \end{aligned}$$

*N-мерный поперечник  $\kappa_N(X, Y)$  А. Н. Колмогорова множества X, лежащего в линейном нормированном пространстве Y относительно этого пространства определяется формулой*

$$\kappa_N(X, Y) = \inf_{Y^N \subset Y} \sup_{x \in X} \min_{y \in Y^N} \|y - x\|_Y.$$

где  $Y^N$  — произвольное фиксированное *N-мерное линейное многообразие (гиперплоскость)*.

Поперечники сосчитаны во многих случаях. В частности, известно, что для класса всех функций  $v, v \parallel_{\Gamma} = 0$ , имеющих ограниченные некоторой общей константой непрерывные вторые производные

$$\kappa_N(U, \overset{\circ}{W}) = O\left(\frac{1}{\sqrt{N}}\right), \quad (32)$$

$$\kappa_N(U, \overset{\circ}{W}) = O\left(\frac{1}{\sqrt{N}}\right). \quad (33)$$

При учете дополнительных сведений об искомом решении  $u$ , найденных при предварительном анализе задачи или в результате опыта решения близких задач, сужается класс  $U$ , а при этом поперечники  $\kappa_N(U, \overset{\circ}{W})$ ,  $N = 1, 2, \dots$  могут только уменьшаться.

Поэтому искусство и опыт вычислителя состоят в том, чтобы уметь выбрать узкий класс  $U$ , содержащий искомое решение  $u$ , а затем выбрать при заданном  $N$  базисные функции (22) так, чтобы число  $K_N(U, \overset{\circ}{W}^N)$ , введенное равенством (30), не слишком сильно превосходило  $N$ -мерный поперечник  $\kappa_N(U, \overset{\circ}{W})$ . Тогда в правой части неравенства

$$\|\bar{w}_N - u\|_{\overset{\circ}{W}}^2 \leq \alpha [I(\bar{w}_N) - I(u)] = \alpha \|\bar{w}_N - u\|_{\overset{\circ}{W}}^2 \leq \alpha K_N^2(U, \overset{\circ}{W}^N)$$

будет стоять число, близкое к  $\kappa_N^2(U, \overset{\circ}{W})$ , которое с ростом  $N$  стремится к нулю, и притом тем быстрее, чем уже класс  $U$ . Если произвести достаточно полный учет особенностей решения  $u$ , которые удалось выяснить до вычислений, а затем в соответствии с этим хорошо выбрать базисные функции, то достаточно точные приближения получатся уже при малых значениях  $N$ . Но объем вычислительной работы, которая состоит в вычислении коэффициентов и решении системы (26), зависит именно от  $N$ . Таким образом, получится экономный вычислительный алгоритм.

Проиллюстрируем применение метода Ритца еще одним примером: рассмотрим задачу (B). После того, как система базисных функций (22) выбрана, ищем приближенное решение

$$w_N(x, y, a_1, \dots, a_N) = \sum_{n=1}^N a_n \omega_n^N(x, y)$$

в пространстве  $W^N$  всех линейных комбинаций, подбирая постоянные так, чтобы выражение

$$J[w_N(x, y, a_1, \dots, a_N)]$$

приняло наименьшее значение. Для этого числа  $a_1, \dots, a_N$  надо определить из системы уравнений

$$\frac{\partial J[\omega_N(x, y, a_1, \dots, a_N)]}{\partial a_n} = 0, \quad n = 1, \dots, N. \quad (34)$$

Будем считать, что в определении (27) скалярного умножения функция  $\sigma(s)$  совпадает с той, которая входит в краевое условие задачи (B).

Тогда система уравнений (34) примет вид

$$\sum_{i=1}^N a_i (\omega_i^N, \omega_N^N) = - \iint_D f \omega_n^N dx dy + \int_{\Gamma} \varphi(s) \omega_n^N ds, \quad n = 1, \dots, N. \quad (35)$$

Решение этой системы  $a_n = \bar{a}_n$ ,  $n = 1, \dots, N$ , и дает искомое приближенное решение  $\bar{\omega}_N$  задачи

$$\bar{\omega}_N(x, y) = \sum_{n=1}^N \bar{a}_n \omega_n^N(x, y).$$

Для функции  $\bar{\omega}_N \in W$  и решения  $v$  задачи (B) в силу равенства (8)

$$J(\bar{\omega}_N) - J(v) = (\bar{\omega}_N - v, \bar{\omega}_N - v) \leq \max_{\tilde{v} \in U} \min_{\omega \in W^N} (\omega - \tilde{v}, \omega - \tilde{v}),$$

где  $U$  — тот класс функций, которому принадлежит решение  $v$  задачи (B). Из последнего неравенства видно, что базисные функции  $\omega_1^N, \dots, \omega_n^N$  надо выбирать так, чтобы правая часть этого неравенства была возможно меньше. При этом подчинение базисных функций каким-либо граничным условиям не является обязательным, в отличие от того, как это было в предыдущем примере.

**4. Проекционный метод Галеркина.** Б. Г. Галеркин в 1916 г. предложил численный метод решения краевых задач, не требующий знания их вариационной постановки. Изложим этот метод на примере краевой задачи (A), причем, как в п. 3, будем считать, что

$$\left. \begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - f(x, y) &= 0, & (x, y) \in D, \\ u|_{\Gamma} &= 0. \end{aligned} \right\} \quad (36)$$

Вновь выберем систему базисных функций (22), но будем считать (временно!), что функции  $\omega_n^N(x, y)$  имеют непрерывные вторые производные. Вновь будем искать приближенное решение в виде линейной комбинации

$$\omega_N(x, y, a_1, \dots, a_N) = \sum_{n=1}^N a_n \omega_n^N(x, y). \quad (37)$$

Подставим выражение (37) в левые части уравнения и краевого условия (36). Получим

$$\begin{aligned} \frac{\partial^2}{\partial x^2} [\omega_N(x, y, a_1, \dots, a_N)] + \frac{\partial^2}{\partial y^2} [\omega_N(x, y, a_1, \dots, a_N)] - f(x, y) = \\ = \delta_N(x, y, a_1, \dots, a_N), \\ \omega_N|_{\Gamma} = 0, \end{aligned}$$

где  $\delta_N(x, y, a_1, \dots, a_N)$  — возникающая невязка.

Введем в пространстве  $\mathcal{W}$ , наряду со скалярным умножением  $(\omega', \omega'')$ , введенным выше, еще скалярное умножение

$$[\omega', \omega''] = \iint_D \omega' \omega'' dx dy.$$

Если бы  $\delta_N$  оказалась ортогональной ко всем функциям из  $\mathcal{W}$  в смысле этого скалярного умножения, то  $\delta_N(x, y, a_1, \dots, a_N)$  была бы тождественным нулем, а  $\omega_N$  было бы точным решением. Однако параметров  $a_1, \dots, a_N$  слишком мало, чтобы, распорядившись ими, можно было получить точное решение. Поэтому подберем их из условия, чтобы проекции невязки на все  $\omega_n^N$ ,  $n = 1, \dots, N$ , были равны нулю, т. е. чтобы невязка была ортогональна ко всем базисным функциям (22)

$$[\delta_N f, \omega_n^N] = 0, \quad n = 1, \dots, N. \quad (38)$$

В развернутом виде система уравнений (38) относительно чисел  $a_1, \dots, a_N$  запишется так:

$$\begin{aligned} \iint_D \left( \frac{\partial^2 \omega_N}{\partial x^2} + \frac{\partial^2 \omega_N}{\partial y^2} \right) \omega_n^N dx dy = \iint_D f \omega_n^N dx dy, \\ n = 1, \dots, N. \end{aligned} \quad (39)$$

Интегрируя по частям, видим, что благодаря условию  $\omega_n^N|_{\Gamma} = 0$  имеет место равенство

$$\begin{aligned} \iint_D \left( \frac{\partial^2 \omega_N}{\partial x^2} + \frac{\partial^2 \omega_N}{\partial y^2} \right) \omega_n^N dx dy = \\ = - \iint_D \left( \frac{\partial \omega_N}{\partial x} \frac{\partial \omega_n^N}{\partial x} + \frac{\partial \omega_N}{\partial y} \frac{\partial \omega_n^N}{\partial y} \right) dx dy = \\ = - \sum_{i=1}^N a_i \iint_D \left( \frac{\partial \omega_i^N}{\partial x} \frac{\partial \omega_n^N}{\partial x} + \frac{\partial \omega_i^N}{\partial y} \frac{\partial \omega_n^N}{\partial y} \right) dx dy = \\ = - \sum_{i=1}^N a_i (\omega_i^N, \omega_n^N). \end{aligned}$$

Поэтому система (39) переписывается в виде

$$\sum_{i=1}^N a_i (\omega_i^N, \omega_n^N) = - [f, \omega_n^N], \quad n = 1, \dots, N \quad (40)$$

и при сделанном выборе скалярного умножения [,] в точности совпадает с системой (29) метода Ритца.

От дополнительного предположения о наличии вторых производных у базисных функций можно отказаться, поскольку уравнения метода Галеркина (40) сохраняют смысл и без этого дополнительного требования.

**5. Способы решения алгебраической системы.** При не очень больших  $N$  ( $N \sim 100$ ) система уравнений метода Ритца или Галеркина может быть решена точно по существующим стандартным программам для систем линейных уравнений.

Далее, матрица  $\omega^N$  системы Ритца (29) в нашем примере (а это типично) есть матрица Грама для системы базисных функций  $\omega_n^N$ ,  $n = 1, \dots, N$ . Очевидно, что она симметрична, и известно, что она является положительно определенной. Поэтому для вычисления решения системы Ритца (29) можно воспользоваться каким-либо итерационным методом, например методом итераций с чебышевским набором параметров.

Итерационные методы значительно облегчаются, если лишь немногие элементы матрицы  $\omega^N$  отличны от нуля. Мы увидим, что в вариационно-разностных и проекционно-разностных схемах так именно и будет.

**6. Вычислительная устойчивость.** Мы видели, что точность приближенного решения при заданном числе  $N$  базисных функций  $\omega_n^N$ ,  $n = 1, \dots, N$ , зависит от того, насколько хорошо можно приблизить решение элементами  $N$ -мерного линейного пространства, натянутого на базисные функции. Таким образом, точность зависит от выбора аппроксимирующего пространства, но не базиса в нем.

Устойчивость, или свойства обусловленности системы уравнений (29) метода Ритца, или системы (40) метода Галеркина, зависят от того, насколько матрица  $\omega^N$  этой системы хорошо обусловлена. С точки зрения устойчивости было бы идеальным, чтобы базисные функции  $\omega_n^N$ ,  $n = 1, \dots, N$ , образовывали ортонормальный базис. Тогда матрица  $\omega^N$  была бы единичной.

#### ЗАДАЧИ

1. Показать, что решение следующей первой краевой задачи для эллиптического уравнения с переменными коэффициентами

$$\frac{\partial}{\partial x} \left[ a(x, y) \frac{\partial u}{\partial x} \right] + \frac{\partial}{\partial y} \left[ b(x, y) \frac{\partial u}{\partial y} \right] = j(x, y),$$

$$u|_{\Gamma} = \varphi(s),$$

$$a(x, y) \geq a_0 > 0, \quad b(x, y) \geq b_0 > 0$$

придает функционалу

$$I(w) = \iint_D \left[ a \left( \frac{\partial w}{\partial x} \right)^2 + b \left( \frac{\partial w}{\partial y} \right)^2 + 2fw \right] dx dy$$

наименьшее значение на классе всех функций  $w \in W$ , удовлетворяющих дополнительно граничному условию  $w|_{\Gamma} = \varphi(s)$ .

Решение  $u(x, y)$  считать имеющим непрерывные вторые производные.

2. Считая систему базисных функций  $\omega_1^N, \dots, \omega_N^N$  заданной, выпisać систему уравнений Рунта для численного отыскания решения  $u(x, y)$  предыдущей задачи 1 при  $\varphi(s) \equiv 0$ .

3. Показать, что решение следующей третьей краевой задачи для эллиптического уравнения с переменными коэффициентами

$$\begin{aligned} \frac{\partial}{\partial x} \left[ a(x, y) \frac{\partial u}{\partial x} \right] + \frac{\partial}{\partial y} \left[ b(x, y) \frac{\partial u}{\partial y} \right] &= f(x, y) \\ \frac{\partial u}{\partial n} + \sigma(s) u \Big|_{\Gamma} &= \varphi(s) \end{aligned}$$

придает минимум функционалу

$$J(w) = \iint_D \left[ a \left( \frac{\partial w}{\partial x} \right)^2 + b \left( \frac{\partial w}{\partial y} \right)^2 + 2wf \right] dx dy + \int_{\Gamma} (\sigma w^2 - 2\varphi w) ds$$

на множестве всех функций  $w \in W$ .

4. Считая систему базисных функций  $\omega_1^N, \dots, \omega_N^N$  заданной, выпisać систему уравнений Рунта для численного отыскания решения  $u(x, y)$  из задачи 3.

5. Считая систему базисных функций  $\omega_1^N, \dots, \omega_N^N$  заданной, выпisać систему уравнений Галеркина для первой краевой задачи

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + a(x, y) \frac{\partial u}{\partial x} + b(x, y) \frac{\partial u}{\partial y} - c^2(x, y) u &= f, \\ u \Big|_{\Gamma} &= 0. \end{aligned}$$

### § 39. Построение и свойства вариационно-разностных и проекционно-разностных схем

**1. Определение вариационно-разностных и проекционно-разностных схем.** Пусть в замкнутой области  $D$ , где требуется найти решение некоторой вариационной задачи для каждого  $N$  из некоторой монотонно возрастающей последовательности натуральных чисел, указано  $N$  точек  $p_1^N, p_2^N, \dots, p_N^N$ . Совокупность этих точек будем называть сеткой, отвечающей заданному  $N$ . Пусть, далее, для численного решения вариационной задачи по методу Рунта используется какая-нибудь система базисных функций

$$\omega_1^N(x, y), \omega_2^N(x, y), \dots, \omega_n^N(x, y), \dots, \omega_N^N(x, y),$$



$n$ -й член которой  $\omega_n^N(x, y)$  в точке  $P_n^N$  принимает значение 1, а в остальных точках сетки обращается в нуль:

$$\omega_k^N(P_k^N) = \delta_n^k, \quad n, k = 1, 2, \dots, N. \quad (1)$$

В этом случае линейная комбинация

$$\omega_N(x, y, a_1, \dots, a_N) = \sum_{n=1}^N a_n \omega_n^N(x, y)$$

в точке  $P_n^N$  принимает значение  $\omega_N(P_n^N) = a_n$ ,  $n = 1, \dots, N$ . Поэтому можно написать

$$\omega_N(x, y) = \sum_{n=1}^N \omega_N(P_n^N) \omega_n^N(x, y).$$

Система уравнений Рунта для определения таких значений коэффициентов  $a_1, \dots, a_N$  линейной комбинации, при которых вариационный функционал достигает минимума на линейном пространстве, натянутом на базисные функции  $\omega_1^N, \dots, \omega_N^N$ , будет связывать таким образом значения  $\omega_N(P_n^N)$ ,  $n = 1, \dots, N$ , самой искомой функции  $\omega_N$  в точках выбранной нами сетки  $P_1^N, \dots, P_N^N$ , т. е. окажется некоторой разностной схемой.

Эта разностная схема в соответствии со способом ее построения называется вариационно-разностной.

Точно также, если воспользоваться базисными функциями  $\omega_1^N, \dots, \omega_N^N$ , удовлетворяющими условиям (1), для реализации проекционного метода Галеркина, то уравнения Галеркина превратятся в некоторую разностную схему, которую естественно назвать проекционно-разностной.

Для наглядности полезно заметить следующее. При заданных значениях  $\omega_N(P_n^N)$ ,  $n = 1, \dots, N$ , линейную комбинацию

$$\omega_N(x, y) = \sum_{n=1}^N \omega_N(P_n^N) \omega_n^N(x, y)$$

можно понимать как некоторую формулу, доопределяющую, или *восполняющую*, функцию  $\omega_N(x, y)$  всюду в области  $D$  по ее значениям  $\omega_N(P_n^N)$ ,  $n = 1, \dots, N$ , в точках сетки. Очевидно, что выбор сетки  $P_n^N$ ,  $n = 1, \dots, N$ , при заданном числе  $N$ , а также выбор системы базисных функций  $\omega_1^N, \dots, \omega_N^N$ , удовлетворяющих условию (1) и определяющих способ восполнения сеточной функции, неоднозначны. Так, например, в одномерном случае функцию можно было бы восполнять на отрезке по ее значениям на сетке кусочно-линейно, квадратически, строя интерполяционный многочлен Лагранжа и т. д. От выбора сетки  $P_n^N$  и базисных функций зависят вид и свойства возникающей

вариационно-разностной или проекционно-разностной схемы для данной вариационной или дифференциальной краевой задачи.

Рассмотрим примеры вариационно-разностных схем для задач (А) и (В) из § 38. При этом будем считать, что область  $D$ , где надо найти решение, выпуклая. (Область  $D$  называется выпуклой, если вместе с любыми двумя точками  $P$  и  $P'$ , принадлежащими области  $D$ , каждая точка отрезка  $PP'$  с концами  $P$  и  $P'$  также принадлежит  $D$ ).

Предположение о выпуклости области  $D$  не связано с существом дела, но облегчит изложение.

**2. Пример вариационно-разностной схемы для первой краевой задачи.** Фиксируем натуральное  $N$ . Впишем в контур  $\Gamma$  ограничивающий область  $D$ , замкнутую несамопересекающуюся ломаную  $Q_1^N Q_2^N \dots Q_m^N Q_1^N$  с вершинами в некоторых точках  $Q_1^N, \dots, Q_m^N$ . Обозначим многоугольник  $Q_1^N \dots Q_m^N Q_1^N$  через  $D_N$ .

Разобьем многоугольник  $D_N$  на треугольники так, чтобы каждое звено ломаной  $Q_1^N \dots Q_m^N Q_1^N$  оказалось стороной одного из треугольников и чтобы каждые два треугольника этого разбиения либо не пересекались, либо имели общую вершину, либо имели общую сторону и чтобы общее число вершин  $P_1^N, \dots, P_N^N$  этих треугольников, лежащих внутри многоугольника  $D_N$ , равнялось  $N$ . Совокупность точек  $P_1^N, \dots, P_N^N$  и примем за

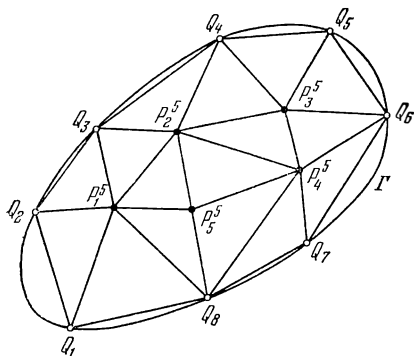


Рис. 47

сетку (рис. 47). Теперь построим базисные функции  $\omega_1^N, \dots, \omega_N^N$ . Определим базисную функцию  $\omega_n^N(x, y)$ ,  $n = 1, \dots, N$ . Сначала зададим ее в точках сетки по формуле (1)

$$\omega_n^N(P_k^n) = \delta_n^k, \quad n, k = 1, \dots, N.$$

Затем зададим ее в точках  $Q_1^N, \dots, Q_m^N$ , положив ее равной нулю в этих точках. Таким образом, функция  $\omega_n^N(x, y)$  уже определена во всех вершинах треугольников, образующих разбиение  $D_N$ . В каждом из этих треугольников доопределим ее линейно. Осталось определить ее в области  $D \setminus D_N$ , где мы положим ее равной нулю.

Заметим, что в тех треугольниках, для которых  $P_n^N$  не является одной из вершин, функция  $\omega_n^N(x, y)$  при нашем построе-

нии окажется равной нулю. В треугольнике с вершиной в точке  $P_n^N$  функция  $\omega_n^N = \omega_n^N(x, y)$  в пространстве  $xuy$  изобразится куском плоскости (рис. 48), проходящей через сторону, лежащую против вершины  $P_n^N$  и приподнятой на единицу над точкой  $P_n^N$ . Система уравнений Рунца (29), § 38, для определения коэффициентов  $a_n = \bar{\omega}_n(P_n^N)$ , задающих приближенное решение

$$\omega_N = \sum_{i=1}^N \omega_N(P_i^N) \omega_i^N(x, y),$$

имеет вид

$$\sum_{i=1}^N \omega_N(P_i^N) (\omega_i^N, \omega_n^N) = - \iint_D f \omega_n^N dx dy, \\ n = 1, \dots, N. \quad (2)$$

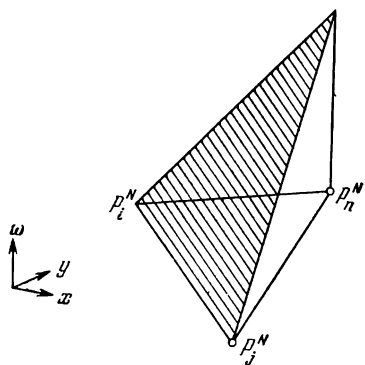


Рис. 48.

Это и есть вариационно-разностная схема, возникающая при сделанном выборе сетки и базисных функций.

Матрица этой разностной схемы

$$\omega^N = \|(\omega_n^N, \omega_i^N)\|, \quad n, i = 1, \dots, N,$$

имеет элементами числа

$$(\omega_n^N, \omega_i^N) = \iint_D \left( \frac{\partial \omega_n^N}{\partial x} \frac{\partial \omega_i^N}{\partial x} + \frac{\partial \omega_n^N}{\partial y} \frac{\partial \omega_i^N}{\partial y} \right) dx dy. \quad (3)$$

Очевидно, что только те числа  $(\omega_n^N, \omega_i^N)$  могут отличаться от нуля, для которых точки  $P_n^N$  и  $P_i^N$  являются вершинами одного и того же треугольника разбиения. Действительно, если  $P_n^N$  и  $P_i^N$  не являются соседними точками сетки в этом смысле, то области, в которых  $\omega_n^N \neq 0$  и  $\omega_i^N \neq 0$  не пересекаются, а потому подынтегральное выражение в формуле (3) всюду в области интегрирования  $D$  тождественно обращается в нуль.

Таким образом,  $n$ -е уравнение из числа образующих вариационно-разностную схему (2), связывает значения искомой функции в точке  $P_n^N$  со значениями этой функции только в тех точках сетки, которые являются для нее соседними.

Вычисление коэффициентов по формуле (3) трудностей не вызывает. Действительно, коэффициент  $(\omega_n^N, \omega_i^N)$  представляет

собой интеграл от величины

$$\frac{\partial \omega_n^N}{\partial x} \frac{\partial \omega_i^N}{\partial x} + \frac{\partial \omega_n^N}{\partial y} \frac{\partial \omega_i^N}{\partial y}, \quad (4)$$

распространенный по паре треугольников разбиения, имеющих отрезок  $P_n^N P_i^N$  своей общей стороной. Далее, интеграл по любому одному из этих треугольников полностью определяется длинами его сторон, но не зависит от поворотов и сдвигов этого треугольника. В самом деле, постоянная внутри треугольника величина (4) подынтегрального выражения представляет собой произведение длин векторов  $\text{grad } \omega_n^N$  и  $\text{grad } \omega_i^N$  на косинус угла между этими векторами и выражается формулой

$$\frac{1}{h_n h_i} \cos(\bar{h}_n, \bar{h}_i). \quad (5)$$

Здесь  $h_n$  и  $h_i$  — длины высот, выходящих из вершин  $P_n^N$  и  $P_i^N$  соответственно, а  $\bar{h}_n$  и  $\bar{h}_i$  — векторы, направленные вдоль соответствующих высот в сторону вершин, как и векторы  $\text{grad } \omega_n^N$  и  $\text{grad } \omega_i^N$ . Интеграл по треугольнику получается умножением величины (5) на площадь треугольника.

Построение вариационно-разностной схемы (2) при сделанном выборе точек  $Q_i^N$  и  $P_i^N$  закончено. Однако очевидно, что не при всяком выборе этих точек, по которому однозначно определяется и система базисных функций  $\omega_n^N$ ,  $n = 1, \dots, N$ , можно ожидать, что полученное с помощью схемы (2) приближенное решение

$$\bar{w}_N = w_N(x, y, \bar{w}_n^N(P_1^N), \dots, \bar{w}^N(P_N^N))$$

будет «хорошим приближением» для точного решения  $u(x, y)$  задачи. В самом деле, если, например, все точки  $Q_1^N, \dots, Q_m^N$  и  $P_1^N, \dots, P_N^N$  разместить в одной «половине» области, а в другой «половине» не поместить ни одной точки сетки, то приближение  $P_1^N, \dots, P_N^N$  не должно оказаться хорошим. Целесообразный выбор точек  $Q_1^N, \dots, Q_m^N$  и сетки  $P_1^N, \dots, P_N^N$  должен быть осуществлен с учетом соотношений § 38, которые мы сейчас воспроизведем:

$$\begin{aligned} \|\omega_N - u\|_W^2 &\leq \alpha \|\bar{w}_N - u\|_{\bar{W}}^2 \leq \alpha K_N^2(U, \overset{\circ}{W}^N) = \\ &= \alpha \sup_{v \in U} \inf_{w \in \overset{\circ}{W}^N} \|w - v\|_{\bar{W}}^2. \end{aligned} \quad (6)$$

Здесь  $\overset{\circ}{W}^N$  —  $N$ -мерное линейное пространство всевозможных линейных комбинаций базисных функций, а  $U$  — множество функций, которому принадлежит точное решение.

Именно, из соотношений (6) видно, что целесообразный выбор точек  $Q_1^N, \dots, Q_m^N$  и  $P_1^N, \dots, P_N^N$  должен быть осуществлен с учетом класса  $U$  таким образом, чтобы число  $K_N(U, \overset{\circ}{W}^N)$  оказалось «возможно меньшим» и чтобы последовательность  $K_N(U, \overset{\circ}{W}^N)$  при  $N \rightarrow \infty$  «возможно быстрее» стремилась к нулю.

Всегда  $K_N(U, \overset{\circ}{W}^N) \geq \kappa_N(U, \overset{\circ}{W})$ , где  $\kappa_N(U, \overset{\circ}{W})$  —  $N$ -мерный колмогоровский поперечник множества  $U$  относительно нормированного пространства  $\overset{\circ}{W}$  (см. п. 3 § 38). Поэтому достаточным условием «хорошего» выбора точек является «близость» числа  $K_N(U, \overset{\circ}{W}^N)$  к  $\kappa_N(U, \overset{\circ}{W})$ .

Однако, вообще говоря, не для всякого множества функций  $U$  существует выбор сетки, при котором  $K_N(U, \overset{\circ}{W}^N)$  «не сильно» превосходит  $\kappa_N(U, \overset{\circ}{W})$ , так, чтобы при  $N' \rightarrow \infty$  величины  $K_N(U, \overset{\circ}{W}^N)$  и  $\kappa_N(U, \overset{\circ}{W})$  имели одинаковый порядок малости относительно  $N^{-1}$ . Дело в том, в частности, что кусочно-линейные базисные функции, которыми мы пользуемся в этом пункте для восполнения сеточных функций, при любом выборе точек порождают пространства  $\overset{\circ}{W}^N$  кусочно-линейных функций, которые не исчерпывают всех возможных  $N$ -мерных подпространств пространства  $\overset{\circ}{W}$  и среди которых может и не оказаться подпространств  $\overset{\circ}{W}^N$ , реализующих хорошую аппроксимацию множества  $U$ .

Разберем подробно случай, когда имеющаяся об искомом решении  $u$  предварительная информация позволяет заключить лишь, что это решение  $u$  принадлежит классу всех функций  $U$ , вторые производные которых не превосходят некоторого числа  $M$  и которые обращаются в нуль на границе.

В этом случае мы укажем, как надо расположить точки  $Q_1^N, \dots, Q_m^N, P_1^N, \dots, P_N^N$ , чтобы с ростом  $N$  величина  $K_N(U, \overset{\circ}{W}^N)$  имела порядок  $O\left(\frac{1}{\sqrt{N}}\right)$ . При этом благодаря (6) для погрешности  $\bar{w}_N$  —  $u$  приближенного решения  $\bar{w}_N$  гарантируются оценки

$$\left. \begin{aligned} \|\omega_N - u\|_{\overset{\circ}{W}} &\leq \frac{c}{\sqrt{N}}, \\ \|\omega_N - u\|_{\overset{\circ}{W}} &\leq \frac{ac}{\sqrt{N}}, \end{aligned} \right\} \quad (7)$$

где  $c$  — некоторая постоянная.

Заметим, что эти оценки из-за равенств (32) и (33), § 38 для поперечников

$$\left. \begin{aligned} \kappa_N(U, \overset{\circ}{W}) &= O\left(\frac{1}{\sqrt{N}}\right), \\ \kappa_N(U, \overset{\circ}{W}) &= O\left(\frac{1}{\sqrt{N}}\right), \end{aligned} \right\} \quad (8)$$

неулучшаемы в следующем смысле. Если искать приближенное решение в виде линейной комбинации каких-либо фиксированных функций  $\psi_1^N(x, y), \dots, \psi_N^N(x, y)$ ,

$$\omega_N = \sum_{k=1}^N c_k \psi_k^N,$$

то ни при каком выборе функций  $\psi_k^N(x, y)$  и ни при каком способе вычисления коэффициентов  $c_k$  по правой части  $f(x, y)$ , нельзя получить оценки вида  $\|\omega_N - u\|_{\overset{\circ}{W}} = o\left(\frac{1}{\sqrt{N}}\right)$  и  $\|\omega_N - u\|_{\overset{\circ}{W}} = o\left(\frac{1}{\sqrt{N}}\right)$ , справедливые при любой  $f(x, y)$ , для которой решение  $u$  принадлежит нашему классу  $U$ .

**Теорема 1.** Пусть  $U$  — множество всех функций, вторые производные которых непрерывны и не превосходят по модулю некоторого числа  $M$ , и которые обращаются в нуль на границе  $\Gamma$ . Пусть для каждого  $N$  из некоторой возрастающей последовательности натуральных чисел осуществлен выбор точек  $Q_1^N, Q_2^N, \dots, Q_m^N$ ,  $m = m(N)$ , разбиение многоугольника  $D_N = Q_1^N \dots Q_m^N Q_1^N$  на треугольники, порождающее сетку  $P_1^N, P_2^N, \dots, P_N^N$ , как описано выше. Пусть при этом выполнены следующие три условия:

1°. Длина  $l$  любой стороны треугольников разбиения удовлетворяет неравенству

$$l \leq C_1 h,$$

где

$$h = \left[ \frac{\text{площадь } D}{N} \right]^{1/2},$$

$a$   $C_1$  — некоторое положительное число, не зависящее от  $h$ .

2°. Площадь  $S_N$  области  $D \setminus D_N$  удовлетворяет оценке

$$S_N \leq C_2 h^2, \quad C_2 = \text{const}. \quad (9)$$

3°. Каждый угол  $\alpha$  любого из треугольников разбиения области  $D_N$  удовлетворяет оценке

$$\alpha > \alpha_0 = \text{const} > 0. \quad (10)$$

При сформулированных условиях для величины  $K_N(U, \overset{\circ}{W}^N)$ :

$$K_N^2(U, \overset{\circ}{W}^N) = \sup_{v \in U} \inf_{\omega \in \overset{\circ}{W}^N} \iint_D \left\{ \left[ \frac{\partial(\omega - v)}{\partial x} \right]^2 + \left[ \frac{\partial(\omega - v)}{\partial y} \right]^2 \right\} dx dy \quad (11)$$

имеет место оценка

$$K_N(U, \overset{\circ}{W}^N) \leq C_3 h, \quad (12)$$

где  $C_3$  — некоторая постоянная.

Доказательство. Достаточно установить, что для каждой функции  $v \in U$  следующая функция

$$w(x, y) = \sum_{k=1}^N v(P_k^V) \omega_k^N(x, y), \quad w \in W, \quad (13)$$

удовлетворяет оценке

$$\iint_D \left\{ \left[ \frac{\partial(w-v)}{\partial x} \right]^2 + \left[ \frac{\partial(w-v)}{\partial y} \right]^2 \right\} dx dy \leq C_4 h^2, \quad (14)$$

так как, очевидно, в таком случае выполнена и оценка (12).

Интеграл (14) можно разбить на сумму (неотрицательных) интегралов по многоугольнику  $D_N$ , вписанному в область  $D$ , и по его дополнению  $D \setminus D_N$  до всей области  $D$ :

$$\begin{aligned} \iint_D \left\{ \left[ \frac{\partial(w-v)}{\partial x} \right]^2 + \left[ \frac{\partial(w-v)}{\partial y} \right]^2 \right\} dx dy = \\ = \iint_{D_N} \left\{ \left[ \frac{\partial(w-v)}{\partial x} \right]^2 + \left[ \frac{\partial(w-v)}{\partial y} \right]^2 \right\} dx dy + \\ + \iint_{D \setminus D_N} \left\{ \left[ \frac{\partial(w-v)}{\partial x} \right]^2 + \left[ \frac{\partial(w-v)}{\partial y} \right]^2 \right\} dx dy. \end{aligned} \quad (15)$$

Оценим каждое из двух слагаемых в правой части (15) и установим оценки

$$\iint_{D_N} \{ \dots \} dx dy \leq A_1 h^2, \quad (16)$$

$$\iint_{D \setminus D_N} \{ \dots \} dx dy \leq A_2 h^2, \quad (17)$$

где  $A_1$  и  $A_2$  — некоторые постоянные, не зависящие от  $v \in U$  и от  $h$ . Очевидно, что из (16) и (17) в силу (15) вытекает (14) с постоянной  $C_4 = A_1 + A_2$ .

Для доказательства оценки (16) достаточно показать, что внутри каждого из треугольников, образующих разбиение области  $D_N$ , выполнены неравенства

$$\left| \frac{\partial(w-v)}{\partial x} \right| \leq B h, \quad \left| \frac{\partial(w-v)}{\partial y} \right| \leq B h, \quad (18)$$

где  $B$  — некоторая постоянная, не зависящая ни от  $v \in U$ , ни от  $h$ . Тогда, очевидно, оценка (16) имеет место, если за  $A_1$  принять число  $A_1 = 2B^2 \times \times$  (площадь  $D$ ). Итак, для завершения доказательства оценки (16) надо установить оценки (18), к чему мы и переходим. Доказательство оценок (18) мы разобьем на два этапа. Сначала покажем, что производная  $\frac{d(w-v)}{dt}$

функции по любому направлению  $l$  может изменяться внутри треугольника не более чем на величину  $A_3 h$ , где  $A_3 = \text{const}$ , так что для любых двух точек  $(x', y')$  и  $(x'', y'')$ , принадлежащих треугольнику, можно написать

$$\left| \left[ \frac{d(w-v)}{dl} \right]_{(x'', y'')} - \left[ \frac{d(w-v)}{dl} \right]_{(x', y')} \right| \leq A_3 h. \quad (19)$$

Выберем затем какие-либо две стороны этого треугольника, образующие острый угол  $\alpha$ , и покажем, что всюду в треугольнике производные  $\frac{d(w-v)}{dl_1}$

и  $\frac{d(w-v)}{dl_2}$  вдоль направлений  $l_1$  и  $l_2$  этих сторон удовлетворяют оценкам

$$\left| \frac{d(w-v)}{dl_1} \right| \leq A_3 h, \quad \left| \frac{d(w-v)}{dl_2} \right| \leq A_3 h. \quad (20)$$

Далее воспользуемся формулами

$$\left. \begin{aligned} \frac{d(w-v)}{dl_1} &= \frac{\partial(w-v)}{\partial x} \cos \alpha_1 + \frac{\partial(w-v)}{\partial y} \sin \alpha_1, \\ \frac{d(w-v)}{dl_2} &= \frac{\partial(w-v)}{\partial x} \cos \alpha_2 + \frac{\partial(w-v)}{\partial y} \sin \alpha_2, \end{aligned} \right\} \quad (21)$$

где  $\alpha_1$  и  $\alpha_2$  — углы, образованные направлениями  $l_1$  и  $l_2$  с осью  $Ox$ . Рассматривая формулы (21) как систему уравнений относительно  $\frac{\partial(w-v)}{\partial x}$  и  $\frac{\partial(w-v)}{\partial y}$ , найдем

$$\left. \begin{aligned} \frac{\partial(w-v)}{\partial x} &= \frac{\sin \alpha_2}{\sin(\alpha_2 - \alpha_1)} \frac{d(w-v)}{dl_1} - \frac{\sin \alpha_1}{\sin(\alpha_2 - \alpha_1)} \frac{d(w-v)}{dl_2}, \\ \frac{\partial(w-v)}{\partial y} &= -\frac{\cos \alpha_2}{\sin(\alpha_2 - \alpha_1)} \frac{d(w-v)}{dl_1} + \frac{\cos \alpha_1}{\sin(\alpha_2 - \alpha_1)} \frac{d(w-v)}{dl_2}. \end{aligned} \right\} \quad (22)$$

Но благодаря условию (10) угол  $\alpha = \alpha_2 - \alpha_1 \geq \alpha_0 > 0$  и  $\alpha \geq \pi - 2\alpha_0$ , так что  $\sin \alpha \geq \sin \alpha_0 = \text{const} > 0$ .

Из равенств (22) и неравенств (20) следуют оценки

$$\left| \frac{\partial(w-v)}{\partial x} \right| \leq \frac{2}{\sin \alpha_0} A_3 h, \quad \left| \frac{\partial(w-v)}{\partial y} \right| \leq \frac{2}{\sin \alpha_0} A_3 h,$$

которые примут вид (18), если обозначить  $B = \frac{2}{\sin \alpha_0} A_3$ .

Для завершения доказательства оценок (18), а вместе с тем и (16) осталось доказать оценки (19) и (20), на которые мы опирались. Докажем (19). Обозначим через  $s$  направление от точки  $(x', y')$  к точке  $(x'', y'')$ . На отрезке, соединяющем эти точки, любая функция  $\psi(x, y)$  может рассматриваться как функция от  $s$ , где  $s$  — расстояние от точки  $(x', y')$ . По теореме о конечных приращениях

$$\psi(x'', y'') - \psi(x', y') = \sqrt{(x'' - x')^2 + (y'' - y')^2} \frac{d\psi(\xi, \eta)}{ds},$$

где  $(\xi, \eta)$  — некоторая точка на отрезке, соединяющем точки  $(x', y')$  и  $(x'', y'')$ . Если

$$\psi(x, y) \equiv \frac{dv(x, y)}{dl},$$



то

$$\frac{dv(x'', y'')}{dl} - \frac{dv(x', y')}{dl} = \sqrt{(x'' - x')^2 + (y'' - y')^2} \frac{d}{ds} \left( \frac{dv(\xi, \eta)}{dl} \right). \quad (23)$$

Обозначим углы между направлениями  $l$  и  $s$  с осью  $Ox$  соответственно через  $\alpha$  и  $\beta$ . Тогда имеют место символические равенства

$$\begin{aligned} \frac{d}{dl} &= \cos \alpha \frac{\partial}{\partial x} + \sin \alpha \frac{\partial}{\partial y}, \\ \frac{d}{ds} &= \cos \beta \frac{\partial}{\partial x} + \sin \beta \frac{\partial}{\partial y}. \end{aligned}$$

Очевидно, что

$$\begin{aligned} \frac{d}{ds} \left( \frac{d}{dl} \right) &= \cos \alpha \cos \beta \frac{\partial^2}{\partial x^2} + [\cos \alpha \sin \beta + \sin \alpha \cos \beta] \frac{\partial^2}{\partial x \partial y} + \\ &+ \sin \alpha \sin \beta \frac{\partial^2}{\partial y^2}. \end{aligned}$$

Поэтому

$$\begin{aligned} \left| \frac{d}{ds} \left[ \frac{dv(\xi, \eta)}{dl} \right] \right| &= \\ = \left| \cos \alpha \cos \beta \frac{\partial^2 v(\xi, \eta)}{\partial x^2} + \sin(\alpha + \beta) \frac{\partial^2 v(\xi, \eta)}{\partial x \partial y} + \sin \alpha \sin \beta \frac{\partial^2 v(\xi, \eta)}{\partial y^2} \right| &\leq 3M \end{aligned}$$

Поскольку  $\sqrt{(x'' - x')^2 + (y'' - y')^2} \leq 2c_1 h$ , то из (23) получим неравенство

$$\left| \frac{dv(x'', y'')}{dl} - \frac{dv(x', y')}{dl} \right| < 6Mc_1 h,$$

совпадающее с (19), если принять  $A_3 = 6Mc_1$ . Для доказательства первого неравенства (20) заметим, что на стороне треугольника, имеющей направле-

ние  $l_1$ , есть точка, где  $\frac{d(w - v)}{dl_1} = 0$ . В самом деле, на концах этой стороны  $w - v$  обращается в нуль по построению, а значит, по теореме Ролля в промежуточной точке производная обращается в нуль. Обозначим координаты этой точки  $(x', y')$  и воспользуемся неравенством (19), в котором примем направление  $l$  совпадающим с направлением  $l_1$ . Получим первое неравенство (20). Второе доказывается аналогично. Завершив доказательство неравенств (19) и (20), мы завершили тем самым и доказательство неравенства (16). Для завершения доказательства всей теоремы осталось установить неравенство (17).

Заметим прежде всего, что каждая функция  $v \in U$  удовлетворяет условиям

$$\left| \frac{dv}{dx} \right| \leq ML, \quad \left| \frac{\partial v}{\partial y} \right| \leq ML, \quad (24)$$

где  $M$  — максимум модулей вторых производных функции  $v(x, y)$  в области  $D$ , а  $L$  — диагональ какого-либо квадрата, содержащего  $D$ . Пусть прямая  $y = \text{const}$  пересекает область  $D$ . Поскольку в концах отрезка пересечения этой прямой с  $\Gamma$  по условию  $v(x, y)$  обращается в нуль, то в некоторой внутренней точке  $(x^0, y)$  этого отрезка по теореме Ролля будет  $\frac{\partial v(x_0, y)}{\partial x} = 0$ . В любой другой точке этого отрезка

$$\left| \frac{\partial v(x, y)}{\partial x} \right| = \left| \frac{\partial v(x, y)}{\partial x} - \frac{\partial v(x_0, y)}{\partial x} \right| = |x - x_0| \cdot \left| \frac{\partial^2 v(\xi, y)}{\partial x^2} \right| \leq LM.$$

Второе неравенство (24) доказывается аналогично. Из конструкции базисных функций  $\omega_n^N$  следует, что функция  $w(x, y) = \sum_{n=1}^N \sigma(P_n^N) \omega_n^N(x, y)$  в области  $D \setminus D_N$ , по которой ведется интегрирование в левой части (17), есть тождественный нуль. Таким образом, благодаря оценкам (24) подынтегральная функция в левой части неравенства (17) не превосходит числа  $2M^2L^2$ , а сам интеграл не превосходит числа

$$2M^2L^2 \cdot S_N \leq 2M^2L^2C_2h^2$$

Таким образом, неравенство (17) справедливо, если принять  $A_4 = 2M^2L^2C_2$ . Теорема доказана.

**3. Пример вариационно-разностной схемы для третьей краевой задачи.** Рассмотрим третью краевую задачу (B) § 38:

$$\left. \begin{aligned} \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} &= f(x, y), \\ \frac{\partial u}{\partial n} + \sigma(s) u &= \varphi(s). \end{aligned} \right\} \quad (25)$$

Пусть при некотором  $N$ , принадлежащем заданной возрастающей последовательности натуральных чисел выбрана сетка  $P_1^N, \dots, P_N^N$  и система базисных функций  $\omega_1^N, \dots, \omega_N^N$ , удовлетворяющих условию (1)

$$\omega_n^N(P_k^N) = \delta_n^k, \quad n, k = 1, 2, \dots, N.$$

Тогда для коэффициентов  $w_N(P_n^N)$  линейной комбинации

$$w_N(x, y) = \sum_{n=1}^N w_N(P_n^N) \omega_n^N(x, y),$$

придающей минимум функционалу  $J(w)$  на классе всех функций вида  $w = a_0 + a_1 \omega_1^N + \dots + a_N \omega_N^N$ , система уравнений Рунца (35) § 38 запишется как следующая вариационно-разностная схема:

$$\sum_{i=1}^N w_N(P_i^N) (w_n^N, \omega_i^N) = - \iint_D f \omega_n^N dx dy + \int_{\Gamma} \varphi(s) \omega_n^N ds, \quad n = 1, \dots, N. \quad (26)$$

Несколько конкретизируем выбор сетки и базисных функций. Для заданного натурального  $N$  впишем в контур  $\Gamma$  замкнутую несамопересекающуюся ломаную  $Q_1^N Q_2^N \dots Q_m^N Q_1^N$  с вершинами в точках  $Q_1^N, \dots, Q_m^N$  и ограничивающую многоугольник  $D_N$ . Разобьем этот многоугольник на треугольники так, чтобы каждые два из них либо не пересекались, либо имели общую вершину, либо имели общую сторону и чтобы общее число вер-

шин этих треугольников, включая вершины  $Q_1^N, \dots, Q_m^N$ , оказалось равно  $N$ . Совокупность всех этих вершин примем за сетку. Обозначим точки сетки через  $P_1^N, P_2^N, \dots, P_m^N$ , причем положим для определенности  $P_n^N = Q_n^N$  для  $n = 1, 2, \dots, \dots, m$ . Определим теперь базисную функцию  $\omega_n^N(x, y)$ ,  $n = 1, \dots, \dots, N$ , следующим образом. Сначала зададим ее в точках сетки в соответствии с условием (1):

$$\omega_n^N(P_k^N) = \delta_n^k, \quad n, k = 1, 2, \dots, N. \quad (27)$$

Затем определим ее в каждом треугольнике разбиения так, чтобы она оказалась в этом треугольнике линейной функцией, принимающей в его вершинах значения по формуле (27). Таким образом, функция  $\omega_n^N(x, y)$

уже определена всюду в многоугольнике  $D_N$ . Определим ее теперь в области  $D \setminus D_N$  и на границе  $\Gamma$ . Область  $D \setminus D_N$  состоит из лунок, каждая из которых ограничена одним из звеньев ломаной  $Q_1^N \dots Q_m^N Q_1^N$  и стягиваемой этим звеном как хордой дугой контура  $\Gamma$ . Фиксируем произвольно одну из этих лунок и рассмотрим треугольник из разбиения  $D_N$ , для которого хорда этой лунки является одной из его сторон. В этом треугольнике функция  $\omega_n^N(x, y)$  уже определена и является линейной функцией (быть может, тождественным нулем). Определим теперь  $\omega_n^N(x, y)$  внутри лунки и на ее границе так, чтобы  $\omega_n^N(x, y)$  осталась линейной функцией в области, образованной присоединением лунки к треугольнику (эта область заштрихована на рис. 49). Пропедев такое доопределение в каждой из лунок, мы завершим построение функции  $\omega_n^N(x, y)$ .

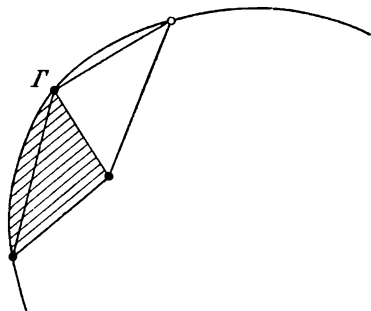


Рис 49.

Теперь коэффициенты и правые части вариационно-разностной схемы (26) приняли определенные численные значения. Заметим, что если точки  $P_n^N$  и  $P_i^N$  не являются вершинами одного и того же треугольника разбиения, то соответствующий коэффициент  $(\omega_n^N, \omega_i^N)$  схемы (26) обращается в нуль.

Обсудим теперь вопрос о точности приближенного решения, найденного с помощью схемы (26). В силу теоремы 4 § 38

$$\|\omega_N - v\|_W^2 \leq \beta [J(\omega_N) - J(v)]. \quad (28)$$

Далее, в силу равенства (8) § 38 и (28)

$$\begin{aligned} \|\omega_N - v\|_{\tilde{W}}^2 &\leq \beta [J(\omega_N) - J(v)] = \\ &= \beta \left( \iint_D \left\{ \left[ \frac{\partial(\omega_N - v)}{\partial x} \right]^2 + \left[ \frac{\partial(\omega_N - v)}{\partial y} \right]^2 \right\} dx dy + \right. \\ &\quad \left. + \int_{\Gamma} \sigma(s) (\omega_N - v)^2 ds \right) \equiv \beta \|\omega_N - v\|_{\tilde{W}}. \end{aligned} \quad (29)$$

Пусть о точном решении  $v(x, y)$  известно лишь, что оно принадлежит некоторому множеству функций  $V$ . Тогда для разности  $\omega_N - v$  в силу равенства (29) гарантированы лишь следующие оценки:

$$\left. \begin{aligned} \|\bar{\omega}_N - v\|_{\tilde{W}}^2 &\leq K_N^2(V, \tilde{W}^N), \\ \|\bar{\omega}_N - v\|_{\tilde{W}} &\leq \beta K_N^2(V, \tilde{W}^N), \end{aligned} \right\} \quad (30)$$

где

$$K_N^2(V, \tilde{W}^N) = \sup_{v \in V} \inf_{\omega \in \tilde{W}^N} (\omega - v, \omega - v) \quad (31)$$

и  $\tilde{W}^N$  — линейное  $N$ -мерное пространство, натянутое на выбранную нами систему базисных функций  $\omega_1^N(x, y), \dots, \omega_N^N(x, y)$ . Рассмотрим теперь случай, когда  $V$  состоит из всех функций, имеющих непрерывные вторые производные, не превосходящие по модулю некоторого числа.

В следующей теореме 2 сформулированы дополнительные требования к сетке  $P_1^N, \dots, P_N^N$ , при выполнении которых

$$K_N(V, \tilde{W}^N) \leq \frac{A}{\sqrt{N}}. \quad (32)$$

**Теорема 2.** Пусть  $V$  — множество всех функций, имеющих непрерывные вторые производные, не превосходящие некоторого числа  $M$  по модулю. Пусть, далее, построенная выше сетка  $P_n^N, n = 1, \dots, N$ , подчинена следующим двум дополнительным требованиям:

1°. Длина  $l$  каждой стороны любого из треугольников разбиения  $D_N$  удовлетворяет оценке

$$l \leq c_1 h, \quad h = \left( \frac{\text{площадь } D}{N} \right)^{1/2},$$

где  $c_1$  — некоторая постоянная.

2°. Каждый угол  $\alpha$  любого треугольника разбиения удовлетворяет оценке

$$\alpha > \alpha_0 > 0,$$

где  $\alpha_0$  — некоторая постоянная, не зависящая от  $N$ .

Тогда справедлива оценка (32).

Доказательство. Из определения (31) величины  $K_N(V, \mathbb{W}^N)$  следует, что для доказательства оценки (32) достаточно для каждой функции  $u(x, y) \in V$  построить такую функцию  $w_N(x, y)$ , для которой имеет место неравенство

$$\iint_D \left\{ \left[ \frac{\partial(w-u)}{\partial x} \right]^2 + \left[ \frac{\partial(w-u)}{\partial y} \right]^2 \right\} dx dy + \int_{\Gamma} \sigma(s)(w-u)^2 ds \leq \frac{A^2}{N} \quad (33)$$

с постоянной  $A$ , не зависящей от  $u$  и от  $h$ . Покажем, что такой функцией может служить функция

$$w(x, y) = \sum_{n=1}^N u(P_n^N) \omega_n^N(x, y). \quad (34)$$

В силу структуры левой части неравенства (33) достаточно показать, что имеют место следующие неравенства:

$$\left| \frac{\partial(w-u)}{\partial x} \right| \leq B_1 h, \quad \left| \frac{\partial(w-u)}{\partial y} \right| \leq B_1 h \quad \text{всюду на } D, \quad (35)$$

$$|w-u| \leq B_2 h \quad \text{на } \Gamma, \quad (36)$$

где  $B_1$  и  $B_2$  — некоторые постоянные. Неравенства (35) доказываются почти дословно так же, как неравенства (18), установленные выше в многоугольнике  $D_N$ . Для доказательства неравенства (36) заметим, что в силу неравенств (35), имеющих место и на границе  $\Gamma$ , производная

$$\frac{d(w-u)}{ds} = \cos \gamma \frac{\partial(w-u)}{\partial x} + \sin \gamma \frac{\partial(w-u)}{\partial y}$$

функции  $w-u$  вдоль границы не превосходит по модулю числа  $2B_1 h$ . Здесь  $\gamma$  — угол между направлением границы в данной точке и осью  $Ox$ . Далее в точках  $P_n^N = Q_n^N$ ,  $n = 1, 2, \dots, m$ , имеют место равенства  $w-u = 0$ . Поэтому в произвольной точке  $Q$  границы

$$|w-u|_Q = \left| \int_{Q_n^N}^Q \frac{d(w-u)}{ds} ds \right| \leq s_{QQ_n^N} \cdot 2B_1 h \leq 2(\text{длина } \Gamma) \cdot B_1 \cdot h,$$

где  $s_{QQ_n^N}$  — расстояние от точки  $Q$  до ближайшей к ней точке  $Q_n^N$  сетки, измеренное вдоль границы  $\Gamma$ . Теорема доказана.

**4. О методике доказательства сходимости.** При использовании вариационно-разностных схем нам не пришлось разбивать доказательство сходимости на исследование устойчивости и аппроксимации, как мы это делали во всех остальных главах. При реализации вычислений устойчивость, которую надо понимать как хорошую обусловленность возникающих систем уравнений, по-прежнему играет важную роль, но не как фактор, обеспечивающий сходимость, а лишь как свойство, позволяющее не учитывать влияние ошибок округления на окончательный результат. Понятие же аппроксимации в том смысле, как мы его понимали всюду в других главах, перестает играть роль. Его заменяет аппроксимация функциональных множеств  $U$  линейными комбинациями базисных функций.

Впрочем, вариационно-разностная схема на регулярной сетке может оказаться некоторой обычной разностной схемой (см. задачу в конце этого параграфа), и тогда вариационный подход к ее исследованию может быть дополнен подходами, принятыми при исследовании обычных разностных схем для получения дополнительной информации о свойствах приближенных решений.

**5. Сопоставление вариационно-разностных схем с общими вариационными и обычными разностными.** Вариационно-разностные схемы — синтез вариационных и обычных разностных схем. Одним из основных достоинств метода Галеркина—Ритца является большая свобода в выборе базисных функций. Если заранее известно, что искомое решение  $u$  принадлежит некоторому конкретному узкому функциональному классу  $U$ , имеющему быстроубывающую последовательность  $N$ -мерных поперечников  $\chi_N$ , то в принципе можно так выбрать базисные функции, чтобы получить хорошую точность уже при малых  $N$  и, следовательно, при малом объеме вычислений. Это давало возможность хорошим вычислителям численно решать избранные задачи еще до появления быстродействующих вычислительных машин. Однако фактическое построение базисных функций с хорошими свойствами — трудная задача.

В вариационно-разностном методе свобода выбора базисных функций ограничена структурой, которая определяется разбиением области на множество многоугольников, вершины которых служат точками сетки, и выбором способа доопределения функций с сетки на всю область. Это ограничение в свободе выбора базисных функций приносит зато некоторый автоматизм в их построении. При этом остается определенная возможность учета особенностей класса функций  $U$ , которому принадлежит решение, за счет использования разбиений на неравные многоугольники и за счет выбора способа восполнений,

который в каждом из многоугольников разбиения может быть осуществлен, как и само разбиение, с использованием априорной информации о поведении решения в этом многоугольнике.

С другой стороны, вариационно-разностные схемы сохраняют удобство обычных разностных схем, состоящее в простой структуре матриц, которые содержат много нулевых элементов. Это достигается за счет использования базисных функций, каждая из которых отлична от нуля лишь в малой области, примыкающей к одной из точек сетки. Сохраняется наглядность разностных схем, в которых искомыми являются значения интересующей нас функции в точках сетки, а не какая-то вспомогательная система чисел, не имеющая непосредственного наглядного истолкования. В то же время вариационно-разностный подход позволяет преодолеть трудности, которые возникают при использовании разностных схем на нерегулярных сетках и при учете граничных условий в криволинейных областях.

### ЗАДАЧА

Пусть разбиение области  $D_N$  на треугольники осуществлено так, что точка  $P_n^N$  сетки при данном фиксированном  $N$  оказалась вершиной прямоугольных

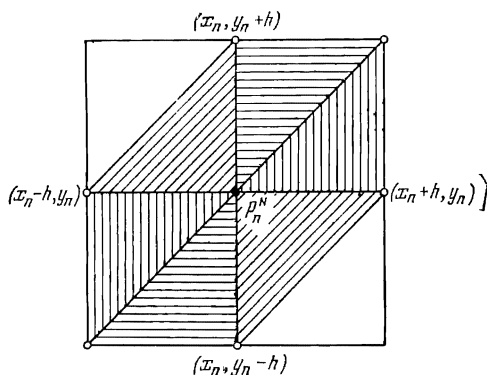


Рис. 50.

равнобедренных треугольников с катетами длины  $h$ , заштрихованными на рис. 50.

Показать, что уравнение

$$\sum_{i=1}^N w_{iN}(\rho_i^N)(\omega_n^N, \omega_i^N) = - \iint_D f \omega_n^N dx dy,$$

отвечающее узлу  $P_n^N$  в вариационно-разностной схеме (2), примет в этом случае вид

$$-h^2 \left[ \frac{\omega_N(x_n+h, y_n) - 2\omega_N(x_n, y_n) + \omega_N(x_n-h, y_n)}{h^2} + \frac{\omega_N(x, y+h) - 2\omega_N(x, y) + \omega_N(x, y-h)}{h^2} \right] = - \iint_D f \omega_n^N dx dy,$$

где  $(x_n, y_n)$  — координаты точки  $P_n^N$ .



## УСТОЙЧИВОСТЬ ЭВОЛЮЦИОННЫХ РАЗНОСТНЫХ КРАЕВЫХ ЗАДАЧ КАК ОГРАНИЧЕННОСТЬ НОРМ СТЕПЕНЕЙ НЕКОТОРОГО ОПЕРАТОРА

В предыдущих частях книги много внимания было уделено исследованию устойчивости разностных краевых задач  $L_h u^{(h)} = f^{(h)}$ . Была исследована, в частности, устойчивость некоторых разностных схем, аппроксимирующих задачу Дирихле для уравнения Пуассона. Это — стационарная задача: ее решение не зависит от времени. Однако основное внимание мы уделили эволюционным задачам, которые отвечают процессам, протекающим во времени, например таким, как распространение тепла или распространение волн. Приемы исследования эволюционных разностных краевых задач лучше развиты. Отчасти это объясняется возможностью рассматривать в ряде случаев стационарные состояния как результат стабилизации процессов, протекающих во времени.

При исследовании устойчивости эволюционных разностных задач мы пользовались принципом максимума, энергетическими неравенствами, спектральными признаками и другими приемами. Во всех этих приемах неявно используется специальная структура эволюционных разностных схем, в силу которой решение  $u^{(h)}$  задается на одном или нескольких начальных временных слоях сетки, а затем шаг за шагом вычисляется на последующих временных слоях. Здесь мы отразим слоистый характер эволюционных разностных схем в самой их записи, поставив в соответствие разностной схеме некоторый линейный оператор  $R_h$ , с помощью которого осуществляется переход от уже известных на данном временном слое значений решения к еще неизвестным значениям этого решения на следующем временном слое. Выбор этого оператора можно осуществить по-разному. Мы будем выбирать его так, чтобы устойчивость разностной схемы оказалась равносильна ограниченности норм его степеней. Это позволит с единой точки зрения взглянуть на уже встречавшиеся приемы исследования устойчивости эволюционных разностных краевых задач, как на способы исследования свойств степеней оператора  $R_h$ , а также сформулировать понятие спектра семейства разностных операторов и спектральный признак устойчивости несамосопряженных разностных краевых задач.

## ГЛАВА 13

## КОНСТРУКЦИЯ ОПЕРАТОРА ПЕРЕХОДА

## § 40. Слоистая структура решений эволюционных задач

Во всех рассмотренных выше примерах эволюционных разностных схем

$$L_h u^{(h)} = f^{(h)} \quad (1)$$

задавались значения решения  $u^{(h)}$  на одном или нескольких начальных слоях сетки. Значения  $u^{(h)}$  на последующих слоях шаг за шагом определялись из уравнений, составляющих разностную краевую задачу (1). Под *слоем* мы понимаем совокупность точек сетки  $D_h$ , лежащих на прямой (или плоскости)  $t = \text{const}$ . В дальнейшем будем считать, что рассматриваемые разностные схемы (1) обладают указанной слоистой структурой.

Пример 1. Рассмотрим разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} &= \Lambda_{xx} u_m^p + \varphi(x_m, t_p), \\ u_0^{p+1} &= \psi_1(t_{p+1}), \quad u_M^{p+1} = \psi_2(t_{p+1}), \quad u_m^0 = \psi(x_m), \\ m &= 0, 1, \dots, M; \quad p = 0, 1, \dots, [T/\tau] - 1, \end{aligned} \right\} \quad (2)$$

аппроксимирующую задачу о распространении тепла

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \varphi(x, t), \quad 0 < x < 1, \quad 0 < t < T, \\ u(0, t) &= \psi_1(t), \quad u(1, t) = \psi_2(t), \quad u(x, 0) = \psi(x), \\ 0 < x < 1. \end{aligned} \right\} \quad (3)$$

Зная значения решения  $u^{(h)}$  в точках слоя  $t = t_p = p\tau$ , т. е. зная сеточную функцию

$$u^p = \{u_m^p\}, \quad m = 0, 1, \dots, M, \quad (4)$$

аргумента  $m$ , мы можем последовательно вычислить значения сеточных функций  $u^{p+1} = \{u_m^{p+1}\}$ ,  $u^{p+2}$  и т. д., пользуясь формулой

$$u_m^{p+1} = (1 - 2r) u_m^p + r(u_{m-1}^p + u_{m+1}^p) + \tau\varphi(x_m, t_p). \quad (5)$$

Сеточная функция  $u^0 = \{u_m^0\} = \{\psi(x_m)\}$  задана.

Итак, решение  $u^{(h)}$ , определенное на двумерной сетке

$$(x_m, t_p) = (mh, p\tau), \quad m = 0, 1, \dots, M; \quad p = 0, 1, \dots, [T/\tau] \quad (6)$$

в плоскости  $Oxt$ , естественным образом расслоилось и заменилось последовательностью функций

$$u^0, u^1, \dots, u^p, \quad p = [T/\tau], \quad (7)$$

определенных на одномерных сетках. Одномерные сетки, на которых определены  $u^p$ ,  $p = 0, 1, \dots, p$ , при всех  $p$  одинаковы (рис. 51, а), так что их можно считать различными экземплярами одной и той же сетки. Мы изобразили эту одномерную сетку на рис. 51, б.

Рассмотрим линейное пространство  $U'$  функций, определенных на одномерной сетке (см. рис. 51, б). К нему принадлежат, в частности, сеточные функции  $u^p$ ,  $p = 0, 1, \dots, p$ . Это линейное пространство будем считать нормированным. Норма элемента  $u = \{u_0, u_1, \dots, u_M\}$  может быть задана, например, одним из равенств

$$\left. \begin{aligned} \|u\| &= \max_m |u_m|, \\ \|u\| &= \left( h \sum_{m=0}^M |u_m|^2 \right)^{1/2}. \end{aligned} \right\} \quad (8)$$

В определении устойчивости и сходимости участвует норма  $\|u^{(h)}\|_{U_h}$  решения разностной краевой задачи (1). Мы будем пользоваться только такими нормами  $\|u^{(h)}\|_{U_h}$ , при выборе которых учтен слоистый характер решения  $u^{(h)}$ , а именно выполнено равенство

$$\|u^{(h)}\|_{U_h} = \max_p \|u^p\|, \quad (9)$$

где  $p$  пробегает значения  $p = 0, 1, \dots, [T/\tau]$ , т. е. все те значения, при которых области определения сеточных функций  $u^p = \{u_m^p\}$  принадлежат двумерной области определения решения  $u^{(h)}$ .

Пример 2. Рассмотрим разностное уравнение

$$\left. \begin{aligned} \frac{u_m^{p+1} - 2u_m^p + u_m^{p-1}}{\tau^2} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= \varphi(x_m, t_p), \\ m = 0, \pm 1, \dots; \quad p = 1, 2, \dots, [T/\tau] - 1, \end{aligned} \right\} \quad (10)$$

которое является разностным аналогом дифференциального уравнения

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = \varphi(x, t), \quad 0 < t < T, \quad -\infty < x < \infty. \quad (11)$$

В отличие от примера 1, решение  $u^{(h)}$  этого разностного уравнения еще не определяется по своим значениям в точках сетки одного слоя  $t = p\tau$ . Здесь необходимо знать значения  $u^{(j)}$

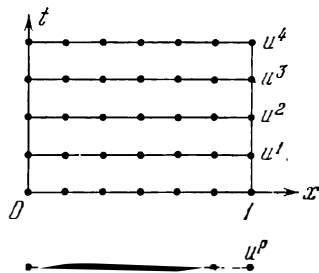
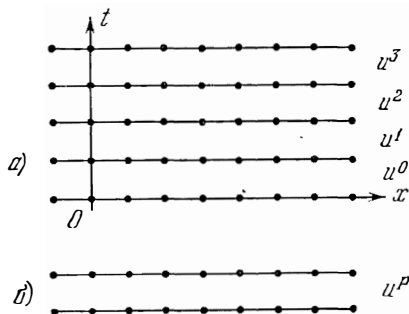


Рис. 51.

в точках сетки двух слоев:  $t = p\tau$  и  $t = (p+1)\tau$ , т. е. вектор-функцию (рис. 52, а)

$$u^p = \begin{pmatrix} \dots u_{-1}^{p+1} & u_0^{p+1} & u_1^{p+1} & \dots \\ \dots u_{-1}^p & u_0^p & u_1^p & \dots \end{pmatrix}.$$

По значениям  $u^p$  в силу уравнения (10) последовательно определяются  $u^{p+1}$ ,  $u^{p+2}$  и т. д. В связи с этим за пространство  $U'_h$  примем пространство вектор-функций (рис. 52, б)



$$u = \begin{pmatrix} \dots b_{-1} & b_0 & b_1 & \dots \\ \dots a_{-1} & a_0 & a_1 & \dots \end{pmatrix}$$

с некоторой нормой  $\|u\|$ . Относительно этой нормы заметим следующее.

Решение дифференциального уравнения (11) определяется двумя функциями:

$$u(x, t_0) \text{ и } \frac{\partial u(x, t)}{\partial t},$$

Рис. 52.

разностными аналогами которых являются соответственно сеточные функции

$$\dots, u_{-1}^p, u_0^p, u_1^p, \dots$$

и

$$\dots, \frac{u_{-1}^{p+1} - u_{-1}^p}{\tau}, \frac{u_0^{p+1} - u_0^p}{\tau}, \frac{u_1^{p+1} - u_1^p}{\tau}, \dots$$

Поэтому всякая естественная норма в пространстве  $U'_h$  должна зависеть от обеих этих сеточных функций. Можно положить, например,

$$\|u\| = \sup_m |a_m| + \sup_m \left| \frac{b_m - a_m}{\tau} \right|$$

или

$$\|u\| = \left[ h \sum_m \left( |a_m| + \frac{1}{\tau^2} |b_m - a_m|^2 \right) \right]^{1/2}.$$

После введения нормы в пространстве  $U'_h$  автоматически по формуле (9) вводится норма в пространстве  $U_h$  сеточных функций, определенных на двумерной сетке:

$$\|u^{(h)}\|_{U_h} = \max_p \|u^p\|.$$

Номер  $p$  пробегает все те значения  $p = 0, 1, \dots, [T/\tau]$ , при которых области определения сеточных вектор-функций принадлежат двумерной области определения сеточной функции  $u^{(h)}$ .

При нашем условии всегда пользоваться нормами вида (9) неравенство

$$\|u^{(i)}\|_{U_h} \leq c \|f^{(h)}\|_{F_h},$$

означающее в случае линейности оператора  $L_h$  устойчивость, равносильно выполнению неравенства

$$\|u^p\| \leq c \|f^{(i)}\|_{F_h}$$

при всех тех  $p$ , при которых функции  $u^p$  определены. Это оказывается удобным при исследовании устойчивости.

### ЗАДАЧИ

1. Указать пространство  $U'_h$  для разностной схемы

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - u_{mn}^p}{\tau} &= \Lambda_{xx} u_{mn}^p + \Lambda_{yy} u_{mn}^p + \Phi(x_m, y_n, t_p), \\ u_{mn}^p|_{\Gamma} &= 0, \quad u_{mn}^0 = \Psi(x_m, y_n), \\ m, n &= 1, 2, \dots, M-1; \quad Mh = 1, \\ p &= 0, 1, \dots, [T/\tau] - 1; \end{aligned} \right\}$$

$(x_m, y_n, t_p) = (mh, nh, p\tau)$ ,  $\Gamma$  состоит из тех точек сетки, которые лежат на боковой границе параллелепипеда  $0 \leq x, y \leq 1, 0 \leq t \leq T$ .

2. Указать пространство  $U'_h$  для разностной схемы расщепления

$$\left. \begin{aligned} \frac{u_{mn}^{p+1} - \tilde{u}_{mn}}{\tau} &= \Lambda_{xx} u_{mn}^{p+1} + \Phi(x_m, y_n, t_p), \\ \frac{\tilde{u}_{mn} - u_{m\tau}^p}{\tau} &= \Lambda_{yy} \tilde{u}_{mn}, \\ u_{mn}^p|_{\Gamma} &= \tilde{u}_{mn}|_{\Gamma} = 0, \quad u_{mn}^0 = \Psi(x_m, y_n), \\ m, n &= 1, 2, \dots, M-1; \quad Mh = 1; \quad p = 0, 1, \dots, [T/\tau] - 1; \end{aligned} \right\}$$

$\Gamma$  — боковая граница параллелепипеда  $0 \leq x, y \leq 1, 0 \leq t \leq T$ .

### § 41. Запись разностных краевых задач в виде $u^{p+1} = R_h u^p + \tau \rho^p$

1. Канонический вид. Разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= \Phi_m^p, \quad m = 0, \pm 1, \dots, \\ u_m^0 &= \psi_m, \quad p = 0, 1, \dots, [T/\tau] - 1, \end{aligned} \right\}$$

Для задачи

$$\begin{aligned} u_t - u_x &= \Phi(x, t), & 0 < t < T, \\ u(x, 0) &= \psi(x), & -\infty < x < \infty, \end{aligned}$$

запишем в виде

$$\left. \begin{aligned} u^{p+1} &= \{(1-r)u_m^p + ru_{m+1}^p\} + \tau\Phi_m^p, \\ u_m^0 &= \psi_m, \quad r = \tau/h. \end{aligned} \right\} \quad (1)$$

Положив

$$v_m^{p+1} = (1-r)u_m^p + ru_{m+1}^p, \quad \rho_m^p = \Phi_m^p, \quad (2)$$

перепишем ее в виде

$$u^{p+1} = v^{p+1} + \tau\rho^p, \quad u_m^0 = \psi_m.$$

Слагаемое  $v^{p+1}$  полностью определяется по  $u^p = \{u_m^p\}$ , так что его можно записать в форме

$$v^{p+1} = R_h u^p,$$

где  $R_h$  — оператор, который каждой сеточной функции  $u^p \in U'_h$  ставит в соответствие сеточную функцию  $v^{p+1} \in U'_h$  по формуле (2). Запись (1) примет тогда вид

$$\left. \begin{aligned} u^{p+1} &= R_h u^p + \tau\rho^p, \\ u^0 &\text{ задано.} \end{aligned} \right\} \quad (3)$$

В этом параграфе мы покажем и на других примерах, как осуществляется приведение эволюционных разностных краевых задач

$$L_h u^{(h)} = f^{(h)} \quad (4)$$

к виду (3). Мы установим, что если при таком приведении удовлетворены некоторые естественные требования, то устойчивость задачи (4) на отрезке  $0 \leq t \leq T$  равносильна выполнению неравенств

$$\|R_h^p\| < K, \quad p = 1, 2, \dots, [T/\tau], \quad (5)$$

где  $K$  — некоторая постоянная, не зависящая от  $h$ , и сведем, таким образом, исследование устойчивости к оценкам величин  $\|R_h^p\|$ , т. е. норм степеней оператора перехода  $R_h$ .

Аналогичные построения и рассуждения были проведены в §§ 15 и 16. Напомним, что в § 41 изучение устойчивости было сведено к рассмотрению неравенств

$$\|u^p\| \leq c \|f^{(h)}\|_{F_h}. \quad (6)$$

Именно было установлено, что устойчивость равносильна существованию числа  $c$ , не зависящего от  $h$  и  $f^{(h)} \in F_h$  и такого, что неравенство (6) выполняется при всех  $p$ ,  $p = 1, 2, \dots, [T/\tau]$ .

Теперь приступаем к реализации намеченного плана, причем начнем с примера приведения разностной схемы к виду (3).

Рассмотрим разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= \varphi_m^p, \\ p = 0, 1, \dots, [T/\tau] - 1; \quad m = 1, 2, \dots, M - 1, \\ u_0^p &= \psi_1(t_p), \quad u_M^p = \psi_2(t_p), \quad p = 0, 1, \dots, [T/\tau], \\ u_m^0 &= \psi(mh), \quad m = 0, 1, \dots, M. \end{aligned} \right\} \quad (7)$$

Ясно, что должны быть выполнены условия согласования  $\psi_1(0) = \psi(0)$ ,  $\psi_2(0) = \psi(1)$ . В силу условий задачи  $u^0 = \{u_m^0\}$  задано, а функции  $u^1, u^2, \dots$  можно последовательно вычислить. Для этого следует переписать разностное уравнение из схемы (7) в виде

$$u_m^{p+1} = (1 - 2r) u_m^p + r (u_{m-1}^p + u_{m+1}^p) + \tau \varphi_m^p,$$

$$r = \frac{\tau}{h^2}, \quad m = 1, 2, \dots, M - 1; \quad p = 0, 1, \dots, [T/\tau] - 1,$$

и использовать равенства

$$u_0^{p+1} = \psi_1(t_{p+1}), \quad u_M^{p+1} = \psi_2(t_{p+1}).$$

Поэтому примем за  $U'_h$  пространство сеточных функций

$$u = \{u_0, u_1, \dots, u_M\}$$

с нормой

$$\|u\| = \max_m |u_m|.$$

Запишем теперь разностную краевую задачу в виде

$$\left. \begin{aligned} u^{p+1} &= R_h u^p + \tau \varphi^p, \\ u^0 &\text{ задано,} \end{aligned} \right\} \quad (8)$$

обозначив через  $R_h$  оператор, который каждому элементу  $a = \{a_m\}$  пространства  $U'_h$  ставит в соответствие некоторый элемент  $b = \{b_m\}$  того же пространства по формулам

$$\left. \begin{aligned} b_0 &= a_0, \\ b_m &= (1 - 2r) a_m + r (a_{m-1} + a_{m+1}), \quad m = 1, 2, \dots, M - 1, \\ b_M &= a_M. \end{aligned} \right\} \quad (9)$$

При таком выборе оператора  $R_h$  сеточная функция  $\rho^p$  из  $U'_h$ ,

$$\rho^p = \{\rho_0^p, \rho_1^p, \dots, \rho_M^p\},$$

определится формулой

$$\rho^p = \left( \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \quad \varphi_1^p, \dots, \varphi_{M-1}^p, \quad \frac{\psi_2(t_{p+1}) - \psi_2(t_p)}{\tau} \right),$$

$$p = 0, 1, \dots, [T/\tau] - 1.$$

Этим мы закончили приведение рассматриваемой разностной схемы (7) к виду (3). Мы собираемся использовать запись разностных краевых задач в этом виде для исследования устойчивости. Но для того, чтобы неравенства (6), означающие устойчивость, имели смысл, должна быть определена норма  $\|f^{(h)}\|_{F_h}$ . В нашем примере разностная краевая задача (7) записывается в виде (4), если положить

$$L_h u^{(h)} \equiv \begin{cases} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2}, & m = 1, 2, \dots, M-1, \\ u_0^{p+1}, & p = 0, 1, \dots, [T/\tau] - 1, \\ u_M^{p+1}, & p = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0, & m = 0, 1, \dots, M, \end{cases}$$

$$f^{(h)} \equiv \begin{cases} \varphi(x_m, t_p) & m = 1, 2, \dots, M-1, \\ \psi_1(t_{p+1}), & p = 0, 1, \dots, [T/\tau] - 1, \\ \psi_2(t_{p+1}), & p = 0, 1, \dots, [T/\tau] - 1, \\ \psi(x_m), & m = 0, 1, \dots, M. \end{cases}$$

Норму  $\|f^{(h)}\|_{F_h}$  определим равенством

$$\|f^{(h)}\|_{F_h} = \max_{m, p} |\varphi(x_m, t_p)| + \max_m |\psi(x_m)| +$$

$$+ \max_p \left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right| + \max_p \left| \frac{\psi_2(t_{p+1}) - \psi_2(t_p)}{\tau} \right|. \quad (10)$$

**2. Устойчивость как равномерная ограниченность норм степеней  $R_h$ .** Сформулируем теперь два условия, соблюдение которых при приведении какой-либо разностной схемы (4) к виду (3) позволяет утверждать, что в случае справедливости неравенств (5) есть устойчивость.

Условие 1°. *Имеют место неравенства*

$$\|\rho^p\| \leq K_1 \|f^{(h)}\|_{F_h},$$

где  $K_1$ , не зависит от  $h$  и  $f^{(h)}$ , а  $p$  пробегает все те значения, при которых  $\rho^p$  имеет смысл.



Условие 2°. Имеет место оценка

$$\|u^0\| \leq K_2 \|f^{(h)}\|_{F_h},$$

где  $K_2$  — некоторая постоянная, не зависящая от  $h$  и  $f^{(h)}$ .

Условия 1° и 2° требуют определенного согласования выбора норм в пространствах  $U'_h$  и  $F_h$ , и выбора оператора  $R_h$  (вид вектора  $\rho^p$  однозначно зависит от выбора оператора  $R_h$ ). Отметим, что в рассмотренном нами примере приведения разностной схемы (7) к виду (3) эти условия выполнены. Чтобы убедиться в этом, достаточно сопоставить нормы сеточных функций  $\rho^p$  и  $u^0$

$$\|\rho^p\| = \max_m |\rho_m^p| = \max \left( \left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right|, |\varphi(x_1, t_p)|, \dots \right. \\ \left. \dots, |\varphi(x_{M-1}, t_p)|, \left| \frac{\psi_2(t_{p+1}) - \psi_2(t_p)}{\tau} \right| \right),$$

$$\|u^0\| = \max_m |u_m^0| = \max_m |\psi(x_m)|$$

с равенством (10), определяющим норму  $\|f^{(h)}\|_{F_h}$ .

Числа  $K_1$  и  $K_2$  в этом примере можно считать равными единице.

Докажем теперь, что если при приведении разностной краевой задачи (4) к виду (3) выполнены условия 1° и 2°, то для устойчивости разностной схемы (4) достаточно, чтобы выполнялись оценки (5). Нам надо показать, что оценки

$$\|u^p\| \leq K_3 \|f^{(h)}\|_{F_h},$$

где  $K_3$  — некоторое число, не зависящее от  $h$  и  $f^{(h)}$ , справедливы при всех  $p$ ,  $p = 0, 1, \dots, p_0$ , при которых области определения сеточных функций  $u^p$  принадлежат области определения решения  $u^{(h)}$ .

Из уравнения

$$u^{p+1} = R_h u^p + \tau \rho^p, \quad p = 0, 1, \dots, p_0 - 1,$$

следует, что

$$u^p = R_h u^{p-1} + \tau \rho^{p-1} = R_h (R_h u^{p-2} + \tau \rho^{p-2}) + \tau \rho^{p-1} = \\ \dots \\ = R_h^p u^0 + \tau (R_h^{p-1} \rho^0 + R_h^{p-2} \rho^1 + \dots + \rho^{p-1}), \quad p = 0, 1, \dots, p_0. \quad (11)$$

По условию

$$\|R_h^l\| \leq K, \quad l = 0, 1, \dots, [T/\tau].$$

Из этого неравенства и равенства (11) вытекает оценка

$$\|u^p\| \leq K \|u^0\| + \tau K (\|\rho^0\| + \|\rho^1\| + \dots + \|\rho^{p-1}\|). \quad (12)$$

Учитывая условия 1° и 2°, в силу которых

$$\begin{aligned} \| \rho^l \| &\leq K_1 \| f^{(h)} \|_{F_h}, \quad l = 0, 1, \dots, p_0 - 1, \\ \| u^0 \| &\leq K_2 \| f^{(h)} \|_{F_h}, \end{aligned}$$

оценку (12) можно заменить следующей:

$$\| u^p \| \leq (KK_2 + \tau p K K_1) \| f^{(h)} \|_{F_h} \leq K(K_1 + TK_1) \| f^{(h)} \|_{F_h} = c \| f^{(h)} \|_{F_h},$$

где число  $c = K(K_2 + TK_1)$  не зависит от  $h$  и  $f^{(h)}$ .

Устойчивость доказана.

Воспользуемся установленным достаточным признаком устойчивости и покажем, что разностная схема (7) при  $r = \tau/h^2 \leq 1/2$  устойчива.

Именно убедимся, что для оператора  $R_h$ , который мы ввели по формулам (9) при приведении разностной схемы (7) к каноническому виду (3), выполнено неравенство  $\| R_h \| \leq 1$ , а значит, и неравенство  $\| R_h^p \| \leq \| R_h \|^p \leq 1$ . При  $r \leq 1/2$  имеют место оценки

$$\begin{aligned} |b_0| &= |a_0| \leq \max_m |a_m| = \|a\|, \\ |b_m| &= |(1-2r)a_m + r(a_{m-1} + a_{m+1})| \leq \\ &\leq (1-2r + 2r) \max_m |a_m| = \|a\|, \\ |b_M| &= |a_M| \leq \max_m |a_m| = \|a\|. \end{aligned}$$

Из этих оценок вытекает, что  $\|b\| = \|R_h a\| \leq \|a\|$ , т. е.  $\|R_h\| \leq 1$ . Итак, при  $r \leq 1/2$  достаточный признак устойчивости выполнен. Можно показать, что если постоянная  $r = \tau/h^2 > 1/2$ , то достаточный признак устойчивости не выполнен. Возникает вопрос, не теряется ли устойчивость и в общем случае, когда неравенства  $\| \dot{R}_h^p \| < K$ ,  $p = 1, 2, \dots, [T/\tau]$ , перестают быть верными. Оказывается, что действительно справедливость неравенств  $\| \dot{R}_h^p \| < K$  необходима для устойчивости, если только выполнено некоторое условие 3°, которое мы сформулируем сейчас в общем виде и которое выполнено для рассматриваемого примера.

Условие 3°. Пусть разностная краевая задача (4) приведена к виду (3). Возьмем какую-нибудь функцию  $\bar{u}^0$  из  $U'_h$  и построим сеточные функции  $\bar{u}^1, \bar{u}^2, \dots, \bar{u}^p, \dots$  по рекуррентным формулам  $\bar{u}^{p+1} = R_h \bar{u}^p$ . Совокупность сеточных функций  $\{\bar{u}^p\}$ ,  $p = 0, 1, \dots, [T/\tau]$ , каждая из которых принадлежит  $U'_h$ , образует некоторую сеточную функцию  $\bar{u}^{(h)}$  из пространства  $U_h$ . Вычислим для нее  $\bar{f}^{(h)}$ ,

$$\bar{f}^{(h)} \equiv L_h \bar{u}^{(h)}.$$

Будем говорить, что *при приведении разностной схемы (4) к каноническому виду (3) выполнено условие  $3^\circ$ , если справедлива оценка вида*

$$\|\bar{f}^{(h)}\|_{F_h} \leq K_3 \|\bar{u}^0\|,$$

где постоянная  $K_3$  не зависит от  $\bar{u}^0$  из  $U'_0$  и не зависит от  $h$ .

Убедимся, что в описанном выше приведении разностной схемы (7) к каноническому виду (3) условие  $3^\circ$  выполнено. Действительно, задавшись произвольной функцией  $\bar{u}^0 = \{\bar{u}_m^0\}$ , получим

$$\bar{\Phi}_m^p \equiv 0, \quad \bar{\Psi}_1^p = u_0^0, \quad \bar{\Psi}_2^p = u_M^0, \quad \bar{\Psi}_m \equiv \bar{u}_m^0.$$

При нашем выборе норм имеет место равенство

$$\|\bar{f}^{(h)}\|_{F_h} = \|\bar{u}^0\|.$$

Докажем теперь, что если при приведении разностной схемы (4) к каноническому виду (3) выполнено условие  $3^\circ$ , то для ее устойчивости на отрезке  $0 < t < T$  необходимо, чтобы выполнялись оценки (5):

$$\|R_h^\rho\| < K, \quad \rho = 1, 2, \dots, [T/\tau],$$

где  $K$  — какое-нибудь число, не зависящее от  $h$ .

Если сформулированный признак не выполнен, то при любом  $K$  можно указать такие  $h$ ,  $\rho_0$  и сеточную функцию  $\bar{u}^0$ , что

$$\|R_h^{\rho_0} \bar{u}^0\| > K \|\bar{u}^0\| \quad (\|\bar{u}^{\rho_0}\| > K \|\bar{u}^0\|).$$

Построив по  $\bar{u}^0$  векторы  $\bar{u}^\rho$  ( $\|\bar{u}^{\rho_0}\| > K \|\bar{u}^0\|$ ) и образовав из  $\bar{u}^\rho$  сеточную функцию  $\bar{u}^{(h)}$ , мы из условия  $3^\circ$  заключаем, что для нее

$$\|\bar{f}^{(h)}\| \leq K_3 \|\bar{u}^0\| \quad (\bar{f}^{(h)} = L_h \bar{u}^{(h)}).$$

В то же время

$$\|\bar{u}^{(h)}\|_{U_h} = \max_p \|\bar{u}^p\| \geq \|\bar{u}^{\rho_0}\| > K \|\bar{u}^0\|.$$

Отсюда ясно, что

$$\|\bar{u}^{(h)}\|_{U_h} > \frac{K}{K_3} \|\bar{f}^{(h)}\|_{F_h}.$$

Это неравенство из-за произвольности  $K$  и означает неустойчивость.

Теперь подведем итог рассмотрению этого параграфа. Мы показали, что, приводя разностную схему  $L_h u^{(h)} = f^{(h)}$  к виду (3),

$$\left. \begin{aligned} u^{\rho+1} &= R_h u^\rho + \tau \rho^\rho, \\ u^0 &\text{ задано,} \end{aligned} \right\}$$

можно использовать затем оператор  $R_h$  для исследования устойчивости. Именно, доказана следующая

**Теорема.** Если при приведении разностной схемы (4) к виду (3) соблюдено условие 3°, то для устойчивости необходимо, чтобы выполнялось неравенство

$$\|R_h^p\| < K, \quad p = 1, 2, \dots, [T/\tau], \quad (13)$$

где  $K$  — некоторое число, не зависящее от  $h$ . Если приведение к виду (3) проведено с соблюдением условий 1° и 2°, то оценки (13) достаточны для устойчивости.

Мы должны обратить внимание читателя на то, что обычно расслоение сеточной функции  $u^{(h)}$  и приведение разностной схемы к каноническому виду (3) могут быть проведены несколькими различными способами. Однако мы не будем на этом подробно останавливаться (см. § 14, где этот вопрос обсуждался в случае разностных схем для обыкновенных дифференциальных уравнений).

**3. Пример.** В заключение параграфа рассмотрим неявную разностную схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m-1}^{p+1} - 2u_m^{p+1} + u_{m+1}^{p+1}}{h^2} &= \Phi_m^p, \\ m = 1, 2, \dots, M-1; \quad p = 0, 1, \dots, [T/\tau] - 1, \\ u_m^0 &= \psi_m, \quad m = 0, \dots, M, \\ u_0^{p+1} &= \psi_1(t_{p+1}), \quad u_M^{p+1} = \psi_2(t_{p+1}), \quad p = 0, 1, \dots, [T/\tau] - 1, \end{aligned} \right\} \quad (14)$$

для задачи о теплопроводности

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} &= \Phi(x, t), \\ u(x, 0) &= \psi(x), \\ u(0, t) &= \psi_1(t), \quad u(1, t) = \psi_2(t), \\ 0 < t < T, \quad 0 < x < 1. \end{aligned} \right\} \quad (15)$$

Мы подробно рассматривали эту схему в § 28.

Примем за  $u^p$  вектор  $u^p = (u_0^p, u_1^p, \dots, u_M^p)$  с нормой  $\|u^p\| = \max_m |u_m^p|$ . Решение на  $(p+1)$ -м слое запишем в виде суммы

$$u^{p+1} = v^{p+1} + \tau \rho^p,$$

где

$$v^{p+1} = (v_0^{p+1}, v_1^{p+1}, \dots, v_M^{p+1}) \quad \text{и} \quad \rho^p = (\rho_0^p, \rho_1^p, \dots, \rho_M^p)$$

в свою очередь являются решениями вспомогательных систем

уравнений

$$\left. \begin{aligned} v_0^{p+1} &= u_0^p = \psi_1(t_p), \\ r v_{m+1}^{p+1} - (1 + 2r) v_m^{p+1} + r v_{m-1}^{p+1} &= -u_m^p, \\ m &= 1, 2, \dots, M-1, \\ v_M^{p+1} &= u_M^p = \psi_2(t_p), \end{aligned} \right\} \quad (16)$$

$$\left. \begin{aligned} \rho_0^p &= \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \\ r \rho_{m+1}^p - (1 + 2r) \rho_m^p + r \rho_{m-1}^p &= \varphi_m^p, \quad m = 1, 2, \dots, M-1, \\ \rho_M^p &= \frac{\psi_2(t_{p+1}) - \psi_2(t_p)}{\tau}. \end{aligned} \right\} \quad (17)$$

Первую из этих систем можно считать определением оператора  $R_h$  и положить

$$v^{p+1} = R_h u^p.$$

Если  $\|f^{(h)}\|_{F_h}$  определить по-прежнему равенством (10), то выполнение условия 1°

$$\|\rho^p\| \leq K_1 \|f^{(h)}\|_{F_h}$$

вытекает из оценки

$$\|\rho^p\| \leq \max \left[ \left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right|, \left| \frac{\psi_2(t_{p+1}) - \psi_2(t_p)}{\tau} \right|, \max_m |\varphi_m^p| \right],$$

справедливой для решения  $\{\rho_m^p\}$  системы (17) в силу оценки (7) § 4. При этом  $K_1 = 1$ .

Условие 2°

$$\|u^0\| \leq K_2 \|f^{(h)}\|_{F_h}$$

также выполнено в силу (10), ибо  $u_m^0 = \psi_m$ , причем, очевидно, можно положить  $K_2 = 1$ . Далее, в § 28 была доказана оценка

$$|v_m^{p+1}| \leq \max_m |u_m^p|,$$

которую можно истолковать как неравенства

$$\|R_h u\| \leq \|u\|, \quad \|R_h\| \leq 1,$$

откуда

$$\|R_h^p\| \leq K = 1.$$

Мы сейчас повторили схему доказательства устойчивости из § 28, показав, что она совпадает с предлагаемой здесь. Рассмотренный пример интересен еще и в том отношении, что в нем используется довольно сложный способ конструирования вектора  $\rho^p$ .

## ЗАДАЧИ

1. Для системы уравнений акустики

$$\frac{\partial u}{\partial t} + A \frac{\partial u}{\partial x} = \varphi(x, t), \quad -\infty < x < \infty, \quad 0 < t < T,$$

$$u(x, 0) = \psi(x), \quad -\infty < x < \infty,$$

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad u(x, t) = \begin{pmatrix} v(x, t) \\ w(x, t) \end{pmatrix},$$

$$\varphi(x, t) = \begin{pmatrix} \varphi_1(x, t) \\ \varphi_2(x, t) \end{pmatrix}, \quad \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix}$$

привести к каноническому виду (3) схему

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} + A \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \\ - \frac{\tau}{2h^2} A^2 (u_{m+1}^p - 2u_m^p + u_{m-1}^p) = \varphi(x_m, t_p), \\ u_m^0 = \psi(x_m), \end{aligned} \right\} (*)$$

приняв  $u^p = \{u_m^p\}$ . Проверить, что при выборе норм по формулам

$$\|u^{(h)}\|_{U_h} = \max_p \|u^p\|, \quad \|f^{(h)}\|_{F_h} = \max_p [\|\psi\|, \max_p \|\varphi^{(p)}\|],$$

где

$$\|u^p\|^2 = \sum_m (|v_m^p|^2 + |w_m^p|^2),$$

$$\|\psi\|^2 = \sum_m (|\psi_1(mh)|^2 + |\psi_2(mh)|^2),$$

$$\|\varphi^{(p)}\|^2 = \sum_m (|\varphi_1(x_m, t_p)|^2 + |\varphi_2(x_m, t_p)|^2).$$

выполнены условия приведения 1° — 3°.

Доказать, что при  $r \leq 1$  разностная схема (\*) устойчива, а при  $r > 1$  неустойчива.

Указание. Для оценки норм  $\|R_h^p\|$  перейти к переменным

$$I_m^{(1)} = v_m + w_m, \quad I_m^{(2)} = v_m - w_m,$$

называемым римановыми инвариантами, и воспользоваться спектральным признаком п. 4 § 25.

2. Разностную схему

$$\left. \begin{aligned} \frac{v_m^{p+1} - 2v_m^p + v_m^{p-1}}{\tau^2} - \frac{v_{m+1}^p - 2v_m^p + v_{m-1}^p}{h^2} = \varphi_m^p, \\ p = 1, \dots, [T/\tau] - 1, \\ v_m^0 = \psi_0(x_m), \quad v_m^1 = \tau\psi_1(x_m) + \psi_0(x_m), \quad m = 0, \pm 1, \dots \end{aligned} \right\}$$

аппроксимирующую задачу Коши

$$\frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} = \varphi(x, t), \quad -\infty < x < \infty, \quad 0 < t < T,$$

$$v(x, 0) = \psi_0(x), \quad \frac{\partial v(x, 0)}{\partial t} = \psi_1(x), \quad -\infty < x < \infty,$$

привести к каноническому виду (3), положив

$$u^p = \left\{ \begin{array}{l} v_m^{p+1} \\ v_m^p \end{array} \right\},$$

$$\|u^p\|^2 = \sum_m |v_m^p|^2 + \sum_m \left| \frac{v_m^{p+1} - v_m^p}{\tau} \right|^2,$$

$$\|\varphi^p\|^2 = \sum_m |\varphi_m^p|^2, \quad \|\psi_k\|^2 = \sum |\psi_k(x_m)|^2,$$

$$\|u^{(h)}\|_{U_h} = \max_p \|u^p\|,$$

$$\|f^{(h)}\|_{F_h} = \|\psi_0\| + \|\psi_1\| + \max_p \|\varphi^p\|.$$

а) Проверить, что выполнены условия 1° — 3°.

б) Доказать, что при  $\tau/h = r \leq 1$  схема устойчива, а при  $\tau/h = r > 1$  неустойчива.

## § 42. Использование частных решений при конструировании оператора перехода

В § 41 рассказывалось о приведении разностной краевой задачи

$$L_h u^{(h)} = f^{(h)} \quad (1)$$

к виду

$$\left. \begin{array}{l} u^{p+1} = R_h u^p + \tau \rho^p, \\ u^0 \text{ задано.} \end{array} \right\} \quad (2)$$

При этом оператор  $R_h$  можно выбирать по-разному. Цель приведения к виду (2) состоит в том, чтобы по оценкам величин  $\|R_h^p\|$  можно было судить об устойчивости. Было показано, что оценка

$$\|R_h^p\| < K, \quad p = 1, 2, \dots, [T/\tau], \quad (3)$$

обеспечивает устойчивость, если только оператор  $R_h$  и нормы выбраны так, что выполняются условия:

$$1^\circ \quad \|\rho^p\| \leq K_1 \|f^{(h)}\|_{F_h},$$

где  $p$  пробегает все значения, при которых  $\rho^h$  определено;

$$2^\circ \quad \|u^0\| \leq K_2 \|f^{(h)}\|_{F_h}.$$

В примерах, рассмотренных в § 41, оператор  $R_h$  удавалось выбрать достаточно простым и в то же время обеспечить выполнение условий  $1^\circ$  и  $2^\circ$ . Однако могут встретиться и такие примеры — один из них мы рассмотрим в этом параграфе, — когда условие  $1^\circ$  является слишком жестким, так что оператор  $R_h$  с учетом этого условия нельзя взять достаточно простым.

В этом параграфе будет показано, что оценка (3) остается достаточной для устойчивости, если условие  $1^\circ$  заменить менее ограничительным условием  $1^*$ . Благодаря замене условия  $1^\circ$  условием  $1^*$  оператор  $R_h$  при приведении разностной схемы к виду (2) можно выбрать тем проще, чем больше сведений имеется о решениях рассматриваемой разностной задачи (1). В частности, структура оператора  $R_h$ , пригодного при приведении разностной схемы к виду (2), упрощается, если известны некоторые частные решения разностного уравнения, входящего в состав разностной краевой задачи (1). Соответственно упрощается доказательство неравенства (3), из которого вытекает устойчивость.

Все это мы поясним позже примерами. Перейдем к формулированию условия  $1^*$ .

Пусть  $z^{(h)}$  — какая-нибудь сеточная функция на  $U_h$ , зависящая, вообще говоря, от  $f^{(h)}$ . При приведении разностной схемы (1) к каноническому виду (2) мы расслаиваем сеточную функцию  $u^{(h)}$  на функции  $u^p \in U'_h$ . Проведем такое же расслоение функции  $z^{(h)}$  на функции  $z^p \in U'_h$  и предположим, что  $z^p$  удовлетворяют неравенствам вида

$$\|z^p\| \leq \tilde{K} \|f^{(h)}\|_{F_h}, \quad (4)$$

где  $\tilde{K}$  — некоторая постоянная, а  $p$  пробегает те значения  $p = 0, 1, \dots, p_0$ , при которых  $z^p$  определено.

Условие  $1^*$ . Существует функция  $z^{(h)}$ , удовлетворяющая неравенствам (4) и такая, что справедливы оценки

$$\left\| \rho^p - \frac{1}{\tau} (z^{p+1} - R_h z^p) \right\| \leq K_1 \|f^{(h)}\|_{F_h}.$$

Если в качестве  $z^{(h)}$  можно взять  $z^{(h)} \equiv 0$ , то выполняется не только условие  $1^*$ , но и более жесткое условие  $1^\circ$ .

**Теорема.** Если разностная задача (1) записывается в каноническом виде (2) с соблюдением условий  $1^*$  и  $2^\circ$ , то из оценки (3) вытекает неравенство

$$\|u^p\| \leq c \|f^{(h)}\|_{F_h}, \quad (5)$$

означающее устойчивость. Постоянная  $c$  может быть выбрана по формуле

$$c = K(2K_2 + 2\tilde{K} + TK_1) + \tilde{K}$$



Доказательство. Введем функцию  $w^{(h)} = u^{(h)} - z^{(h)}$ , для которой из уравнения

$$u^{p+1} = R_h u^p + \tau \rho^p$$

вытекает равенство

$$w^{p+1} = R_h w^p + \tau \tilde{\rho}^p, \quad (6)$$

где

$$\tilde{\rho}^p = \rho^p - \frac{1}{\tau} (z^{p+1} - R_h z^p).$$

В силу условия 1\* имеем

$$\|\tilde{\rho}^p\| \leq K_1 \|f^{(h)}\|_{F_h}.$$

Пользуясь равенством (6) и неравенством (3), без труда находим, как мы много раз делали,

$$\|w^p\| \leq K \|w^0\| + TK \max_p \|\tilde{\rho}^p\| \leq K \|w^0\| + TKK_1 \|f^{(h)}\|_{F_h}. \quad (7)$$

Далее,

$$\|w^0\| \leq 2(K_2 + \tilde{K}) \|f^{(h)}\|_{F_h}. \quad (8)$$

Это следует из неравенств

$$\begin{aligned} \|w^0\| &= \|u^0 - z^0\| \leq \|u^0\| + \|z^0\|, \\ \|u^0\| &\leq K_2 \|f^{(h)}\|_{F_h}, \quad \|z^0\| \leq \tilde{K} \|f^{(h)}\|_{F_h}, \end{aligned}$$

второе из которых совпадает с условием 2°, а третье — с неравенством (4) при  $p = 0$ .

Подставив оценку (8) для  $\|w^0\|$  в (7), находим:

$$\|w^p\| \leq [2(K_2 + \tilde{K})K + TKK_1] \|f^{(h)}\|_{F_h}.$$

Остается только заметить, что

$$\begin{aligned} \|u^p\| &= \|w^p + z^p\| \leq \|w^p\| + \|z^p\| \leq \\ &\leq \{[2K(K_2 + \tilde{K}) + TKK_1] + \tilde{K}\} \|f^{(h)}\|_{F_h} \leq \\ &\leq [K(2K_2 + 2\tilde{K} + TK_1) + \tilde{K}] \|f^{(h)}\|_{F_h} = c \|f^{(h)}\|_{F_h}. \end{aligned}$$

Благодаря замене условия 1° условием 1\* при исследовании устойчивости можно распределить трудности между построением такого оператора  $R_h$ , нормы степеней которого не слишком трудно оценить, и доказательством существования функции  $z^{(h)}$ . Потребовав с самого начала, чтобы условие 1\* выполнялось при  $z^{(h)} = 0$ , т. е. чтобы было выполнено условие 1°, мы накладываем тем самым жесткие ограничения на выбор оператора  $R_h$ . Может оказаться, что любой оператор  $R_h$ , который нам удастся подобрать с выполнением условия 1°, имеет сложный вид,

так что оценка нормы его степеней слишком трудна. С другой стороны, выбрав оператор  $R_h$  предельно простым, скажем единичным, и не увязав его с разностной задачей, мы перенесем всю трудность на проверку выполнения условия  $1^*$ , т. е. на получение оценок для функции  $z^{(h)}$ , которую в этом случае наиболее естественно выбрать совпадающей с  $u^{(h)}$ . Введение такого оператора  $R_h$  и такой функции  $z^{(h)}$  нисколько не продвинуло бы нас в исследовании устойчивости.

В качестве оператора  $R_h$  надо стараться брать как можно более простой оператор. Однако  $R_h$  должен настолько полно учитывать свойства разностной задачи  $L_h u^{(h)} = f^{(h)}$ , чтобы выполнение условия  $1^*$ , т. е. существование функции  $z^{(h)}$ , было достаточно очевидным. Часто удается воспользоваться свободой в выборе  $R_h$ , которая возникает благодаря тому, что вместо условия  $1^0$  должно выполняться лишь менее ограничительное условие  $1^*$ , для облегчения доказательства устойчивости. В качестве функции  $z^{(h)}$  при этом используются функции, которые строятся из решений разностных задач при правой части  $f^{(h)}$  того или иного специального вида.

Мы сейчас покажем на примерах, как пользоваться предлагаемым приемом.

**Пример 1.** Рассмотрим разностную краевую задачу (1) вида

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= \varphi_m^p, \quad m = 0, 1, \dots, M-1; \quad Mh = 1, \\ u_m^0 &= \psi(x_m), \quad m = 0, 1, \dots, M. \\ u_M^p &= \psi_1(t_p), \quad p = 1, 2, \dots, [T/\tau]. \end{aligned} \right\} \quad (9)$$

Это разностная задача аппроксимирует задачу

$$\left. \begin{aligned} u_t - u_x &= \varphi(x, t), \quad 0 < x < 1, \quad 0 < t < T, \\ u(x, 0) &= \psi(x), \quad 0 < x < 1, \\ u(1, t) &= \psi_1(t), \quad 0 < t < T. \end{aligned} \right\} \quad (10)$$

при следующем выборе норм:

$$\|u^{(h)}\|_{U_h} = \max_p \max_m |u_m^p|,$$

$$\|f^{(h)}\|_{F_h} = \max_p \max_m |\varphi_m^p| + \max_m |\psi(x_m)| + \max_p |\psi_1(t_p)| + \max_p \left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right|.$$

Для приведения задачи (9) к каноническому виду (2) положим

$$u^p = (u_0^p, u_1^p, \dots, u_M^p), \quad \|u\| = \max_m |u_m|.$$

Оператор  $R_h$ ,  $b = R_h a$ , переводящий элемент  $a = (a_0, a_1, \dots, a_M)$  пространства  $U'_h$  в элемент  $b = (b_0, b_1, \dots, b_M)$  того же пространства, определим равенствами

$$\left. \begin{aligned} b_m &= (1-r)a_m + ra_{m+1}, & m &= 0, 1, \dots, M-1, \\ b_M &= 0, & r &= \tau/h. \end{aligned} \right\}$$

Тогда, очевидно,

$$\rho^p = \left( \Phi_0^p, \Phi_1^p, \dots, \Phi_{M-1}^p, \frac{\Psi_1(t_{p+1})}{\tau} \right).$$

Ясно, что условие 1°

$$\|\rho^p\| \leq K_1 \|f^{(h)}\|_{F_h}$$

не выполнено из-за того, что последняя компонента вектора  $\rho^p$  есть  $\frac{\Psi_1(t_{p+1})}{\tau}$  и растет при  $\tau \rightarrow 0$ . (В этой задаче легко было бы задать оператор  $R_h$  так, чтобы условие 1° выполнялось. Для этого достаточно в определении оператора  $R_h$  равенство  $b_M = 0$  заменить равенством  $b_M = a_M$ .) Между тем нетрудно указать такое  $z^p$ , что условие 1\*

$$\left\| \rho^p - \frac{1}{\tau} (z^{p+1} - R_h z^p) \right\| \leq \tilde{K} \|f^{(h)}\|_{F_h}$$

окажется выполненным. Левая часть этого неравенства записывается в виде

$$\left\| \left( \Phi_0^p, \Phi_1^p, \dots, \Phi_{M-1}^p, \frac{\Psi_1(t_{p+1})}{\tau} \right) - \frac{1}{\tau} (z^{p+1} - R_h z^p) \right\|.$$

Поэтому для доказательства выполнения условия 1\* достаточно построить функцию  $z^{(i)}$ , удовлетворяющую уравнению

$$\frac{1}{\tau} (z^{p+1} - R_h z^p) = \left( 0, 0, \dots, 0, \frac{\Psi_1(t_{p+1})}{\tau} \right),$$

которое можно записать в виде

$$z^{p+1} = R_h z^p + \tau \left( 0, 0, \dots, 0, \frac{\Psi_1(t_{p+1})}{\tau} \right)$$

или

$$\left. \begin{aligned} \frac{z_m^{p+1} - z_m^p}{\tau} - \frac{z_{m+1}^p - z_m^p}{h} &= 0, & m &= 0, 1, \dots, M-1, \\ z_M^{p+1} &= \Psi_1(t_{p+1}). \end{aligned} \right\} \quad (11)$$

В случае, если  $\Psi_1(t_{p+1})$  не зависит от  $t$ , эта задача имеет стационарное (не зависящее от  $p$ ) решение

$$z_m^p \equiv \psi_1 = \text{const.}$$

В общем случае  $\psi_1^p = \psi_1(t_p)$  зависит от  $p$ , но при ограниченной норме  $\|f^{(h)}\|_{F_h}$ , содержащей член  $\left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right|$ , не может быстро меняться. Поэтому функция  $z^{(h)}$ , определенная равенством

$$z_m^p \equiv \psi_1(t_p),$$

хоть и не будет стационарным решением (да и вообще решением) задачи (11), но «почти» удовлетворяет соотношениям (11). Именно

$$\frac{1}{\tau}(z^{p+1} - R_h z^p) = \left( \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \dots, \dots, \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \frac{\psi_1(t_{p+1})}{\tau} \right).$$

Поэтому

$$\begin{aligned} \left\| \rho^p - \frac{1}{\tau}(z^{p+1} - R_h z^p) \right\| &= \left\| \Phi_0^p - \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \right. \\ \Phi_1^p - \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, \dots, \Phi_{M-1}^p - \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau}, 0 \left. \right\| &\leq \\ &\leq \|(\Phi_0^p, \Phi_1^p, \dots, \Phi_{M-1}^p, 0)\| + \\ &+ \left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right| \|(1, 1, \dots, 1, 0)\| \leq \|f^{(h)}\|_{F_h} \end{aligned}$$

Условие 1\* выполнено;  $\tilde{K} = 1$  и  $z^{(h)}$  удовлетворяют неравенству

$$\|z^p\| = \max_m |z_m^p| \leq \|f^{(h)}\|_{F_h}.$$

Условие 2°

$$\|u^0\| \leq K_2 \|f^{(h)}\|_{F_h}$$

также выполнено:

$$\|u^0\| = \max_m \|u_m^0\| = \max_m |\psi(x_m)| \leq \|f^{(h)}\|_{F_h}.$$

Для доказательства устойчивости, которая имеет место при  $\tau \leq h$ , достаточно показать, что  $\|R_h^p\| \leq 1$ . Справедливость этого вытекает из оценки  $\|R_h\| \leq 1$ :

$$\|R_h a\| = \max_m |a_m(1-r) + r a_{m+1}| \leq \max_m |a_m| = \|a\|$$

Пример 2. Рассмотрим в качестве более сложного примера другую разностную схему для той же дифференциальной краевой задачи (10):

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \frac{\tau}{2} \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} &= \Phi_m^p, \\ \rho = 0, 1, \dots, [T/\tau] - 1; \quad m = 1, 2, \dots, M - 1, \\ u_m^0 &= \psi(x_m), \quad m = 0, 1, \dots, M, \\ u_{M+1}^{p+1} &= \psi_1(t_{\rho+1}), \quad \rho = 0, 1, \dots, [T/\tau] - 1, \\ \frac{u_0^{p+1} - u_0^p}{\tau} - \frac{u_1^p - u_0^p}{h} &= \Phi_0^p, \quad \rho = 0, 1, \dots, [T/\tau] - 1. \end{aligned} \right\} (12)$$

Порядок по  $x$  разностного уравнения, входящего в эту схему, — второй, а порядок соответствующего дифференциального уравнения (10) — первый. Поэтому на левой границе  $x = 0$  ( $m = 0$ ) добавлено условие

$$\frac{u_0^{p+1} - u_0^p}{\tau} - \frac{u_1^p - u_0^p}{h} = \Phi_0^p,$$

которое мы будем использовать в форме

$$u_0^{p+1} = (1 - r)u_0^p + ru_1^p + \tau\Phi_0^p,$$

Разностную схему (12) мы уже рассматривали в § 23, где обсуждались вопросы аппроксимации. Норма в пространстве  $F_h$  там была введена следующим образом: если

$$f^{(h)} = \begin{cases} a_m & \text{в точках } (mh, 0), \quad m = 0, 1, \dots, M, \\ b^p & \text{в точках } (1, p\tau), \quad p = 0, 1, \dots, [T/\tau], \\ c^p & \text{в точках } (0, p\tau), \quad p = 0, 1, \dots, [T/\tau] - 1, \\ \Phi_m^p & \text{в точках } (mh, p\tau), \quad m = 1, 2, \dots, M - 1; \\ & \rho = 0, 1, \dots, [T/\tau] - 1, \end{cases}$$

то

$$\|f^{(h)}\|_{F_h} = h \max_p |c^p| + h \max_p \left| \frac{c^{p+1} - c^p}{\tau} \right| + \\ + \left( h \sum_{m=0}^M |a_m|^2 \right)^{1/2} + \max_p |b^p| + \max_p \left| \frac{b^{p+1} - b^p}{\tau} \right| + \max_p \max_m |\Phi_m^p|.$$

Как было показано в § 23, аппроксимация в этом случае имеет порядок  $h^2$ . Покажем, что при выборе нормы  $\|u^{(h)}\|_{U_h}$  по формуле

$$\|u^{(h)}\|_{U_h} = \max_r \left( \frac{h}{2} |u_0^r|^2 + h \sum_{m=1}^M |u_m^r|^2 \right)^{1/2}$$

где  $r \leq 1$  наряду с аппроксимацией имеет место и устойчивость.

Мы проверим устойчивость, приведя разностную схему (12) к каноническому виду (2). Положим для этого

$$u^p = (u_0^p, u_1^p, \dots, u_M^p)$$

с нормой

$$\|u^p\| = \left( \frac{h}{2} |u_0^p|^2 + h \sum_{m=1}^M |u_m^p|^2 \right)^{1/2}.$$

Оператор  $R_h$  определим следующими формулами:

Если  $a = (a_0, a_1, \dots, a_M)$ ,  $b = (b_0, b_1, \dots, b_M)$  и  $b = R_h a$ , то

$$\left. \begin{aligned} b_0 &= (1-r)a_0 + ra_1, \\ b_m &= \left(-\frac{r}{2} + \frac{r^2}{2}\right)a_{m-1} + (1-r^2)a_m + \left(\frac{r}{2} + \frac{r^2}{2}\right)a_{m+1}, \\ b_M &= 0. \end{aligned} \right\} \quad (13) \quad m = 1, 2, \dots, M-1,$$

В таком случае

$$\rho^p = \left[ \varphi_0^p, \varphi_1^p, \dots, \varphi_{M-1}^p, \frac{\psi_1(t_{p+1})}{\tau} \right].$$

Условие 1°, очевидно, при нашем выборе норм не выполняется. В самом деле, если, например,

$$f^{(h)} = \begin{cases} 0 & \text{в точках } (mh, 0), & m = 0, 1, \dots, M, \\ 0 & \text{в точках } (1, p\tau), & p = 0, 1, \dots, [T/\tau], \\ 1 & \text{в точках } (0, p\tau), & p = 0, 1, \dots, [T/\tau] - 1, \\ 0 & \text{в точках } (mh, p\tau) & m = 1, 2, \dots, M-1, \\ & & p = 0, 1, \dots, [T/\tau] - 1, \end{cases}$$

то

$$\rho^p = (1, 0, \dots, 0), \quad \|\rho^p\| = \frac{\sqrt{h}}{\sqrt{2}}, \quad \|f^{(h)}\|_{F_h} = h,$$

так что ни при каком  $K_1$  не может выполняться для всех  $h$  неравенство

$$\|\rho^p\| \leq K_1 \|f^{(h)}\|_{F_h}.$$

При сделанном нами выборе пространства  $U'_h$ , состоящего из векторов  $u^p = (u_0^p, u_1^p, \dots, u_M^p)$ , и при нашем выборе норм, по-видимому, нельзя указать оператор  $R_h$  так, чтобы условие  $1^\circ$  выполнялось, но условию  $1^*$  удовлетворить можно.

Прежде чем доказать последнее утверждение, заметим, что если изменить норму  $\|f^{(h)}\|_{F_h}$ , положив

$$\|f^{(h)}\|_{F_h} = \sqrt{h} \max |c^p| + \left( h \sum_{m=0}^M |a_m|^2 \right)^{1/2} + \max \left| \frac{b^p}{\tau} \right| + \max_p \max_m |\varphi_m^p|,$$

то оператор  $R_h$ , определенный формулами (13), будет удовлетворять условию  $1^\circ$ , но порядок аппроксимации будет не  $h^2$ , а только  $h^{3/2}$ . Мы умеем, не меняя норм, привести разностную краевую задачу (12) к каноническому виду (2) с соблюдением условий  $1^\circ$  и  $2^\circ$ , если за  $U'_h$  принять совокупность сеточных вектор-функций

$$u^p = \begin{bmatrix} u_m^{p+1} \\ u_m^p \end{bmatrix}, \quad m = 0, 1, \dots, M.$$

При этом усложнится конструкция оператора  $R_h$  и оценка нормы его степеней. Поэтому мы не будем рассматривать такое приведение.

Покажем, что при сделанном нами выборе (13) оператора  $R_h$  существуют  $z^p$ , удовлетворяющие условию  $1^*$ :

$$\left\| \rho^p - \frac{1}{\tau} (z^{p+1} - R_h z^p) \right\| < K_1 \|f^{(h)}\|_{F_h}.$$

Для построения функции  $z^{(h)}$  поступим аналогично тому, как мы делали в примере 1, и выпишем стационарное (не зависящее от  $p$ ) решение задачи

$$\left. \begin{aligned} \frac{z_m^{p+1} - z_m^p}{\tau} - \frac{z_{m+1}^p - z_{m-1}^p}{2h} - \frac{\tau}{2} \frac{z_{m+1}^p - 2z_m^p + z_{m-1}^p}{h^2} &= 0, \\ m &= 1, 2, \dots, M-1, \\ \frac{z_0^{p+1} - z_0^p}{\tau} - \frac{z_1^p - z_0^p}{h} &= \Phi_0^p, \\ z_M^p &= \Psi_1^p, \end{aligned} \right\} \quad (14)$$

в предположении, что  $\Phi_0^p$  и  $\Psi_1^p$  фиксированы и от  $p$  не зависят.

Это решение имеет вид

$$z_m^p = \Phi_0^p h \frac{1+r}{2} \left[ \left( \frac{r-1}{r+1} \right)^m - \left( \frac{r-1}{r+1} \right)^M \right] + \Psi_1^p.$$

Функция  $\{z_m^p\}$  удовлетворяет оценке

$$\|z^p\| \leq \tilde{K} \|f^{(h)}\|_{F_h}, \quad K = 2.$$

Действительно,

$$\|z^p\| = \left( \frac{h}{2} |z_0^p|^2 + h \sum_{m=1}^M |z_m^p|^2 \right)^{1/2} \leq \\ \leq 2h \max_m |\varphi_m^p| + \max_p |\psi_1(t_p)| \leq 2 \|f^{(h)}\|_{F_h}.$$

Пусть

$$\xi^p \equiv z^{p+1} - R_h z^p.$$

Поскольку  $z^p$  — решение стационарной задачи, можно написать

$$z^p = R_h z^p + \tau \delta^p,$$

где

$$\delta^p = \left( \varphi_0^p, 0, 0, \dots, 0, \frac{\psi_1(t_p)}{\tau} \right).$$

Поэтому

$$\xi^p = z^{p+1} - z^p + \tau \delta^p,$$

так что

$$\xi_m^p = \begin{cases} h(\varphi_0^{p+1} - \varphi_0^p) \frac{1+r}{2} \left[ 1 - \left( \frac{r-1}{r+1} \right)^M \right] + \\ \quad + (\psi_1(t_{p+1}) - \psi_1(t_p)) + \tau \varphi_0^p, & m=0, \\ h(\varphi_0^{p+1} - \varphi_0^p) \frac{1+r}{2} \left[ \left( \frac{r-1}{r+1} \right)^m - \left( \frac{r-1}{r+1} \right)^M \right] + \\ \quad + (\psi_1(t_{p+1}) - \psi_1(t_p)), & 1 < m < M, \\ 0, & m=M. \end{cases}$$

Следовательно, координаты вектора  $\tilde{\rho}^p = \rho^p - \frac{1}{\tau} \xi^p$  имеют вид

$$\tilde{\rho}_m^p = \begin{cases} -\frac{\psi_1(t_{p+1}) + \psi_1(t_p)}{\tau} - \frac{h}{\tau} (\varphi_0^{p+1} - \varphi_0^p) \frac{1+r}{2} \left[ 1 - \left( \frac{r+1}{r+1} \right)^M \right], & m=0, \\ -\frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} - \frac{h}{\tau} (\varphi_0^{p+1} - \varphi_0^p) \frac{1+r}{2} \left[ \left( \frac{r-1}{r+1} \right)^m - \right. \\ \quad \left. - \left( \frac{r-1}{r+1} \right)^M \right] + \varphi_m^p, & 0 < m < M, \\ 0, & m=M. \end{cases}$$

Неравенство, означающее условие 1\*

$$\|\tilde{\rho}^p\| = \left\| \rho^p - \frac{1}{\tau} (z^{p+1} - R_h z^p) \right\| \leq K_1 \|f^{(h)}\|_{F_h},$$



выполнено:

$$\begin{aligned} \|\bar{\rho}^p\| &= \left( \frac{h}{2} |\bar{\rho}_0^p|^2 + h \sum_{m=1}^{M-1} |\bar{\rho}_m^p|^2 \right)^{1/2} \leq \\ &\leq \left| \frac{\psi_1(t_{p+1}) - \psi_1(t_p)}{\tau} \right| + \frac{h |\varphi_0^{p+1} - \varphi_0^p|}{\tau} + \max_m |\varphi_m^p| \leq \|f^{(h)}\|_{F_h}. \end{aligned}$$

Выполнение условия 2°

$$\|u^0\| \leq K_2 \|f^{(h)}\|_{F_h}$$

очевидно:

$$\|u^0\| = \left( \frac{h}{2} |u_0^0|^2 + h \sum_{m=1}^M |u_m^0|^2 \right)^{1/2} \leq \left( h \sum_{m=0}^M |u_m^0|^2 \right)^{1/2} \leq \|f^{(h)}\|_{F_h}.$$

Для доказательства устойчивости, которая имеет место при  $r \leq 1$ , остается показать, что при этом условии

$$\|R_h^p\| \leq K, \quad p = 1, 2, \dots, [T/\tau], \quad (15)$$

где  $K$  — некоторая постоянная, не зависящая от  $h$ . Докажем сначала, что для любого вектора  $u = (u_0, u_1, \dots, u_M)$ , компонента  $u_M$  которого равна нулю, справедливо неравенство

$$\|R_h u\| \leq \|u\|. \quad (16)$$

Далее, применяя оператор  $R_h$  к вектору  $(0, 0, \dots, 0, 1)$ , получим вектор  $(0, 0, \dots, 0, \frac{r}{2} + \frac{r^2}{2}, 0)$ , норма которого не превосходит  $\sqrt{h}$ . Поэтому для произвольного вектора  $u = (u_0, u_1, \dots, u_M)$ , компонента которого  $u_M$  не обязательно равна нулю, имеем с учетом неравенства (16), справедливого для вектора  $u$  вида  $u = (u_0, u_1, \dots, u_{M-1}, 0)$ ,

$$\begin{aligned} \|R_h u\| &= \|R_h(u_0, u_1, \dots, u_{M-1}, 0) + u_M R_h(0, 0, \dots, 0, 1)\| \leq \\ &\leq \|R_h(u_0, u_1, \dots, u_{M-1}, 0)\| + |u_M| \sqrt{h} \leq \\ &\leq \|(u_0, u_1, \dots, u_{M-1}, 0)\| + \|(0, 0, \dots, u_M)\| \leq 2\|u\|, \\ &\|R_h\| \leq 2. \end{aligned} \quad (17)$$

Теперь докажем неравенство (15). В силу определения оператора  $R_h$  вектор  $v = R_h u$  имеет нулевую компоненту  $v_M$ ,  $v_M = 0$ . Поэтому, используя (16) и (17), получаем

$$\begin{aligned} \|R_h^p u\| &= \|R_h^{p-1}(R_h u)\| = \|R_h^{p-1} v\| \leq \|v\| = \\ &= \|R_h u\| \leq 2\|u\|, \quad \|R_h^p\| \leq 2 \equiv K. \end{aligned}$$

Остается обосновать неравенство (16), на которое мы опирались, т. е. доказать следующее предложение.

Пусть  $u = (u_0, u_1, \dots, u_{M-1}, 0)$  — произвольный вектор, компонента  $u_M$  которого равна нулю, и пусть  $v \equiv R_h u$ . Тогда

$\|v\| \leq \|u\|$ , т. е.

$$\left(\frac{h}{2}v_0^2 + h \sum_{m=1}^M v_m^2\right)^{1/2} \leq \left(\frac{h}{2}u_0^2 + h \sum_{m=1}^M u_m^2\right)^{1/2}. \quad (18)$$

Напомним, что в силу определения (13) оператора  $R_h$

$$\left. \begin{aligned} v_0 &= (1-r)u_0 + ru_1, \\ v_m &= \left(-\frac{r}{2} + \frac{r^2}{2}\right)u_{m-1} + (1-r^2)u_m + \left(\frac{r}{2} + \frac{r^2}{2}\right)u_{m+1}, \\ & \quad m = 1, 2, \dots, M-1, \\ v_M &= 0. \end{aligned} \right\}$$

Отметим неравенство

$$\begin{aligned} v_m^2 &\leq \left[\left(-\frac{r}{2} + \frac{r^2}{2}\right)u_{m-1} + (1-r^2)u_m + \left(\frac{r}{2} + \frac{r^2}{2}\right)u_{m+1}\right]^2 + \\ & \quad + \frac{r^2(1-r^2)}{4}(u_{m-1} - 2u_m + u_{m+1})^2 \equiv \\ &\equiv \frac{r^2(1-r)}{2}u_{m-1}^2 + (1-r^2)u_m^2 + \frac{r^2(1+r)}{2}u_{m+1}^2 - \\ & \quad - r(1-r^2)u_{m-1}u_m + r(1-r^2)u_mu_{m+1}, \end{aligned}$$

которое выполнено при  $r \leq 1$ , а также очевидное тождество

$$\frac{r^2(1-r)}{2} + 1 - r^2 + \frac{r^2(1+r)}{2} = 1.$$

Теперь при  $r \leq 1$  легко проследить справедливость следующей цепочки неравенств, не считая при этом, что  $u_M = 0$ :

$$\begin{aligned} \frac{1}{2}v_0^2 + \sum_{m=1}^M v_m^2 &\leq \frac{1}{2}[u_0(1-r) + u_1r]^2 + \\ & + \sum_{m=1}^{M-1} \left[ \frac{r^2(1-r)}{2}u_{m-1}^2 + (1-r^2)u_m^2 + \frac{r^2(1+r)}{2}u_{m+1}^2 - \right. \\ & \quad \left. - r(1-r^2)u_{m-1}u_m + r(1-r^2)u_mu_{m+1} \right] = \\ & = \frac{1}{2}[u_0(1-r) + u_1r]^2 + \sum_{m=1}^M u_m^2 + \\ & + \left[ \frac{r^2(1-r)}{2}u_0^2 - \frac{r^2(1+r)}{2}u_1^2 - r(1-r)u_0u_1 \right] - \frac{r^2(1-r)}{2}u_{M-1}^2 = \\ & = \left[ \frac{1}{2}u_0^2 + \sum_{m=1}^M u_m^2 \right] - \frac{r}{2}\{u_0^2 + [(1-r)u_0 + ru_1]^2\} - \frac{r^2(1-r)}{2}u_{M-1}^2 \leq \\ & \leq \frac{1}{2}u_0^2 + \sum_{m=1}^M u_m^2. \end{aligned}$$

Полученное энергетическое неравенство

$$\frac{1}{2} v_0^2 + \sum_{m=1}^M v_m^2 \leq \frac{1}{2} u_0^2 + \sum_{m=1}^M u_m^2$$

сильнее неравенства (18), которое мы доказываем.

Итак, устойчивость схемы (12) при  $r \leq 1$  установлена. При  $r > 1$  устойчивости нет ни при каком разумном выборе норм, так как нарушено необходимое условие устойчивости Куранта, Фридрикса и Леви.

### § 43. Некоторые способы оценки норм степеней операторов

В §§ 41, 42 было показано, что эволюционные разностные схемы

$$L_h u^{(h)} = f^{(h)} \quad (1)$$

обычно можно привести к виду

$$\left. \begin{aligned} u^{p+1} &= R_h u^p + \tau \rho^p, \\ u^0 &\text{ задано} \end{aligned} \right\} \quad (2)$$

так, чтобы устойчивость была равносильна равномерной по  $h$  ограниченности норм степеней оператора перехода

$$\|R_h^p\| < K, \quad p = 1, 2, \dots, [T/\tau]. \quad (3)$$

Поскольку условия (3) равносильны устойчивости, то любой способ исследования устойчивости есть также способ проверки того, выполняются ли неравенства (3).

Здесь мы изложим с точки зрения оценки норм степеней операторов некоторые подходы к исследованию устойчивости, встречающиеся уже в гл. 8, дополнив эти подходы новыми аспектами.

#### 1. Необходимые спектральные условия ограниченности $\|R_h^p\|$ .

Пусть  $\lambda_h$  — какое-нибудь собственное число оператора  $R_h$ , а  $u^{(h)}$  — соответствующий собственный вектор,  $R_h u^{(h)} = \lambda_h u^{(h)}$ . Тогда

$$\|R_h^p u^{(h)}\| = |\lambda_h|^p \|u^{(h)}\|,$$

и поэтому  $\|R_h^p\| \geq |\lambda_h|^p$ . Поскольку  $\lambda_h$  — произвольное собственное число, то

$$\|R_h^p\| \geq [\max |\lambda_h|]^p, \quad p = 1, 2, \dots, [T/\tau], \quad (4)$$

где  $\max |\lambda_h|$  — наибольшее из собственных чисел оператора  $R_h$ . В силу (4) очевидно (см. § 15), что для выполнения условия (3) должен существовать круг

$$|\lambda| \leq 1 + c\tau \quad (5)$$

на комплексной плоскости с постоянной  $\epsilon$ , не зависящей от  $\hbar$ , в котором лежат все собственные числа оператора  $R_h$ .

Проведенные рассуждения лишь несущественно усложнятся, а результат не изменится, если в качестве  $\lambda_h$  рассматривать не только собственные значения оператора  $R_h$ , но и все его точки спектра. В случае, если  $U'_h$  — конечномерное пространство, спектр оператора  $R_h$  не зависит от выбора нормы и целиком состоит из собственных значений. Это — важнейший случай, естественно возникающий при аппроксимации дифференциальных краевых задач в ограниченной области разностными задачами на сетке  $D_h$ , состоящей из конечного числа узлов. В этом случае условие (5) необходимо для выполнения условия (3) независимо от выбора нормы. Если необходимый спектральный признак устойчивости не выполнен, задача безнадежно неустойчива — этого нельзя поправить никаким разумным выбором норм. Аналогичную ситуацию мы подробно разбирали для случая обыкновенных разностных уравнений в § 15.

Выявим связь между спектральным признаком Неймана устойчивости разностной задачи Коши, рассмотренным в § 25, и спектральным признаком (5) равномерной ограниченности (3) норм степеней оператора  $R_h$ . Воспользуемся для этого, например, разностной схемой

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= \varphi(x_m, t_p), \\ u_m^0 &= \psi(x_m), \\ m = 0, \pm 1, \dots; \quad p &= 0, 1, \dots, [T/\tau] - 1, \end{aligned} \right\} \quad (6)$$

аппроксимирующей задачу Коши

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} &= \varphi(x, t), \quad -\infty < x < \infty, \quad 0 < t < T, \\ u(x, 0) &= \psi(x). \end{aligned}$$

Мы исследовали ее устойчивость с помощью признака Неймана в § 25.

Перепишем рассматриваемую схему в каноническом виде (2), определив  $R_h$ ,  $v = R_h u$  и  $\rho^p$  формулами

$$\left. \begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, \quad r = \tau/h, \\ \rho^p &= \varphi^p \{ \equiv \varphi(x_m, t_p) \}. \end{aligned} \right\}$$

Определим норму в  $U'_h$  формулой  $\|u\| = \sup_m |u_m|$ . Тогда функции  $u = \{u_m\} = \{e^{i\alpha m}\}$  при любом вещественном  $\alpha$  принадлежат пространству  $U'_h$  и являются собственными функциями оператора  $R_h$ :

$$R_h u = (1-r)e^{i\alpha m} + re^{i\alpha(m+1)} = [(1-r) + re^{i\alpha}]e^{i\alpha m} = \lambda(\alpha)u,$$

где

$$\lambda(\alpha) = 1 - r + re^{i\alpha} \quad (7)$$

является собственным числом. Условие устойчивости (5) ввиду независимости  $\lambda(\alpha)$  от  $\tau$  сводится к требованию  $|\lambda(\alpha)| \leq 1$ , которое выполняется при всех вещественных  $\alpha$  при  $r \leq 1$ .

Как показано в § 25, условие (5) в случае задачи Коши для двуслойных разностных схем относительно одной искомой сеточной функции не только необходимо, но и достаточно для устойчивости, если норма определена равенством

$$\|u\| = \left( h \sum_{m=-\infty}^{\infty} u_m^2 \right)^{1/2}.$$

(В этом случае  $\{e^{iam}\}$  не принадлежат пространству  $U'_h$  и, следовательно, не являются собственными функциями, но точки (7) все равно принадлежат спектру оператора  $R_h$ .)

**2. Спектральный критерий ограниченности степеней самосопряженного оператора.** Допустим, что  $M$ -мерное линейное пространство  $U'_h$  состоит из функций, определенных в точках  $P_1, P_2, \dots, P_m$  сетки (эта сетка может лежать на отрезке, на плоскости или в пространстве — безразлично), и что в  $U'_h$  введено скалярное произведение, которое для произвольной пары функций  $u, v \in U'_h$  обозначим  $(u, v)$ . Пусть, далее, оператор  $R_h$  равномерно по  $h$  ограничен некоторой постоянной  $c_1$ :

$$\|R_h\| < c_1, \quad (8)$$

и отображает пространство  $U'_h$  на некоторое подпространство  $\tilde{U}'_h \subset U'_h$  размерности  $N \leq M$ , причем на подпространстве  $\tilde{U}'_h$  оператор  $R_h$  является самосопряженным, т. е.  $(R_h u, v) = (u, R_h v)$  для любой пары функций  $u, v \in \tilde{U}'_h$ . Как известно из курса линейной алгебры, в этом случае в подпространстве  $\tilde{U}'_h$  существует ортонормальный базис

$$\psi^{(1)}, \psi^{(2)}, \dots, \psi^{(N)}, \quad (9)$$

состоящий из собственных векторов оператора  $R_h$ . Обозначим через  $\lambda_1, \lambda_2, \dots, \lambda_N$  соответствующие (вещественные) собственные числа:

$$R_h \psi^{(k)} = \lambda_k \psi^{(k)}, \quad k = 1, 2, \dots, N. \quad (10)$$

**Теорема 1.** Для выполнения оценки (3) необходимо и достаточно, чтобы выполнялось неравенство

$$\max_k |\lambda_k| \leq 1 + c_2 \tau, \quad c_2 = \text{const}. \quad (11)$$

Доказательство. Необходимость доказана в п. 1. Докажем достаточность. Пусть  $u \in U'_h$ . Разложим вектор  $R_h u = v \in \tilde{U}'_h$  по базису (9):

$$v = \sum \alpha_k \psi^{(k)}.$$

Тогда в силу (10)

$$R_h^p u = R_h^{p-1} v = \sum_{k=1}^M (\lambda_k^{p-1} \alpha_k) \psi^{(k)},$$

$$\|R_h^p u\| = \left( \sum_{k=1}^M |\alpha_k \lambda_k^{p-1}|^2 \right)^{1/2} \leq$$

$$\leq \max_k |\lambda_k^{p-1}| \left( \sum_{k=1}^M \alpha_k^2 \right)^{1/2} = \max_k |\lambda_k^{p-1}| \cdot \|v\|. \quad (12)$$

Принимая во внимание, что благодаря (8)

$$\|v\| = \|R_h u\| \leq c_1 \|u\|,$$

а также условие (11), из оценки (12) выведем (3):

$$\|R_h^p\| \leq c_1 \max_k |\lambda_k^{p-1}| \leq c_1 (1 + c_2 \tau)^{p-1} \leq c_1 (1 + c_2 \tau)^{T/\tau} < c_1 e^{c_2 T} = K.$$

Ниже мы установим некоторые признаки самосопряженности оператора  $R_h$  и укажем некоторые способы оценки собственных чисел.

**3. Признаки самосопряженности.** Введем обозначение

$$[u, v] = \frac{1}{M} \sum u(P_k) v(P_k) \quad (13)$$

и будем считать, что скалярное произведение в пространстве  $U'_h$  определено равенством

$$(u, v) = [u, v]. \quad (14)$$

Пусть, далее, оператор  $R_h$ ,  $b = R_h a$ , задан формулами

$$b(P_k) = \sum_{P_s} \alpha_{ks} a(P_s),$$

где  $P_s$  и  $P_h$  пробегают все множество точек сетки. Оператор  $R_h$  является самосопряженным в том и только том случае, если

$$\alpha_{ks} \equiv \alpha_{sk}. \quad (15)$$

В не зависящей от нумерации точек форме этот признак означает, что при вычислении  $b(P)$  в произвольной точке  $P$  сетки значение  $a(Q)$  в другой произвольной точке  $Q$  должно входить с тем же коэффициентом, с каким значение  $a(P)$  входит в выражение для  $b(Q)$ .

Если среди точек сетки  $\{P_k\}$  выделено некоторое подмножество  $\Gamma_h$  (граница области) и оператор  $R_h$  задан формулами

$$\left. \begin{aligned} b(P_k) &= \sum_{P_s} \alpha_{ks} a(P_s), & P_k \in \Gamma_h, \\ b(P_k) &= 0, & P_k \in \Gamma_h, \end{aligned} \right\} \quad (16)$$

то на подпространстве  $\tilde{U}'_h$  формулы (16) равносильны следующим:

$$\left. \begin{aligned} b(P_k) &= \sum_{P_s \in \Gamma_h} \alpha_{ks} a(P_s), & \text{если } P_k \in \Gamma_h, \\ b(P_k) &= 0, & \text{если } P_k \in \Gamma_h. \end{aligned} \right\}$$

Условие самосопряженности оператора  $R_h$  на подпространстве  $\tilde{U}'_h$  тогда, как легко видеть, состоит в равенствах

$$\alpha_{ks} = \alpha_{sk}, \quad P_s \in \Gamma_h, \quad P_k \in \Gamma_h. \quad (17)$$

Так, например, оператор  $b = R_h a$ ,

$$\left. \begin{aligned} b_k &= (1 - 2r)a_k + r(a_{k-1} + a_{k+1}), & k = 1, 2, \dots, M-1, \\ b_0 &= b_M = 0, \end{aligned} \right\}$$

возникающий при приведении разностного аналога уравнения теплопроводности на отрезке к каноническому виду (2), удовлетворяет условию (17), но не удовлетворяет условию (15).

**4. Оценки собственных значений оператора  $R_h$ .** В некоторых случаях собственные значения можно указать точно, как это было сделано в § 27 для оператора  $\Lambda_{xx}$  над функциями на сеточном отрезке, обращающимися в нуль на его концах, а также для оператора  $\Lambda_{xx} + \Lambda_{yy}$  над функциями на сеточном квадрате, обращающимися в нуль на его сторонах.

Для самосопряженных разностных операторов можно пользоваться вариационными методами. Известно, что в этом случае

$$\min_{u' \in U'_h} \frac{(R_h u', u')}{(u', u')} = \lambda'_{\min}, \quad \max_{u' \in U'_h} \frac{(R_h u', u')}{(u', u')} = \lambda'_{\max}. \quad (18)$$

Пусть, например, оператор  $\Lambda_{xx} + \Lambda_{yy}$  действует на сеточные функции пространства  $U'_h$ , определенные не на квадрате, а в более сложной области, составленной из квадратов, и обращающиеся в нуль на границе этой области. Поместим эту область в достаточно большую квадратную сеточную область и рассмотрим оператор  $\Lambda_{xx} + \Lambda_{yy}$  над функциями из  $U'_h$ , определенными на сеточном квадрате и обращающимися в нуль на его границе.

Доопределим каждую функцию  $u'$  из  $U'_h$  до некоторой функции  $u''$  из  $U''_h$ , положив ее тождественно равной нулю во всех тех точках сеточного квадрата, которые не принадлежат

первоначальной области. Легко видеть, что для каждой такой функции благодаря ее обращению в нуль на границе первоначальной области выполнено равенство

$$\frac{(R_h u', u')}{(u', u')} = \frac{(R_h u'', u'')}{(u'', u'')}, \quad R_h = \Lambda_{xx} + \Lambda_{yy}.$$

Поэтому при переходе от формул (18) к формулам

$$\min_{u'' \in U''_h} \frac{(R_h u'', u'')}{(u'', u'')} = \lambda''_{\min},$$

$$\max_{u'' \in U''_h} \frac{(R_h u'', u'')}{(u'', u'')} = \lambda''_{\max}$$

получим числа  $\lambda''_{\min}$  и  $\lambda''_{\max}$ , которые удовлетворяют оценкам

$$\lambda''_{\min} \leq \lambda'_{\min}, \quad \lambda'_{\max} \leq \lambda''_{\max}. \quad (19)$$

Но в случае квадратной области собственные значения известны, так что  $\lambda''_{\min}$  и  $\lambda''_{\max}$  известны, и мы получили оценки (19) границ спектра оператора  $\Lambda_{xx} + \Lambda_{yy}$  над функциями из  $U''_h$ , определенными в первоначальной области.

Во многих случаях можно применять вариационные методы оценки собственных значений, аналогичные вариационным методам для дифференциальных уравнений. Например, первое собственное значение задачи

$$\tilde{R}_h u^{(h)} = \lambda u^{(h)}, \quad u^{(h)}|_{\Gamma_h} = 0,$$

где  $\Gamma_h$  — граница сеточной области  $D_h$  и во внутренних точках задано выражение

$$\begin{aligned} \tilde{R}_h u_{mn} \equiv & \frac{1}{h^2} \left\{ a \left( x_m + \frac{h}{2}, y_n \right) (u_{m+1, n} - u_{mn}) - \right. \\ & \left. - a \left( x_m - \frac{h}{2}, y_n \right) (u_{mn} - u_{m-1, n}) \right\} + \\ & + \frac{1}{h^2} \left\{ b \left( x_m, y_n + \frac{h}{2} \right) (u_{m, n+1} - u_{mn}) - \right. \\ & \left. - b \left( x_m, y_n - \frac{h}{2} \right) (u_{mn} - u_{m, n-1}) \right\}, \quad a(x, y) > 0, \quad b(x, y) > 0, \end{aligned}$$

может лишь уменьшиться от замены переменных коэффициентов  $a(x, y)$ ,  $b(x, y)$  на постоянные

$$a = \max_{x, y} a(x, y), \quad b = \max_{x, y} b(x, y).$$

Это доказывается так же, как аналогичный факт для дифференциальных уравнений, см. [19].



В случае постоянных коэффициентов можно перейти от исходной области к квадратной и получить оценки, подобные оценкам (19). Собственные числа оператора  $a\Lambda_{xx} + b\Lambda_{yy}$  в квадратной области легко вычислить точно.

**5. Выбор скалярного умножения.** Пусть оператор  $R_h, v = R_h u$ , задан равенством

$$B_h v = \tilde{A}_h u, \quad (20)$$

причем при некотором фиксированном выборе скалярного умножения  $(u, v) \equiv [u, v]$  не обязательно по формулам (13), (14) операторы  $\tilde{A}_h$  и  $B_h$  — самосопряженные:

$$[\tilde{A}_h u, v] \equiv [u, \tilde{A}_h v], \quad [B_h u, v] \equiv [u, B_h v].$$

Пусть, далее,  $B_h > 0$ :

$$[B_h u, u] > 0, \quad \text{если } u \neq 0.$$

Тогда оператор  $R_h = B_h^{-1} \tilde{A}_h$  является самосопряженным в смысле скалярного умножения

$$(u, v)_{B_h} \equiv [B_h u, v]. \quad (21)$$

В самом деле,

$$\begin{aligned} (B_h^{-1} \tilde{A}_h u, v)_{B_h} &= [B_h (B_h^{-1} \tilde{A}_h) u, v] = [\tilde{A}_h u, v] = \\ &= [u, \tilde{A}_h v] = [B_h^{-1} B_h u, \tilde{A}_h v] = [B_h u, B_h^{-1} \tilde{A}_h v] = (u, B_h^{-1} \tilde{A}_h v)_{B_h}. \end{aligned}$$

Доказанное тождество по  $u$  и  $v$

$$(B_h^{-1} \tilde{A}_h u, v)_{B_h} \equiv (u, B_h^{-1} \tilde{A}_h v)_{B_h}$$

и означает самосопряженность оператора  $R_h$ .

Таким образом, выбор скалярного умножения по формуле (21) позволяет воспользоваться спектральным критерием п. 2 ограниченности норм степеней самосопряженных операторов. Именно, можно утверждать, что оператор  $R_h$ , определенный равенством (20), имеет вещественные собственные значения  $\lambda_k$  и полную ортонормальную систему собственных векторов  $\psi^{(k)}$ :

$$\lambda_k B_h \psi^{(k)} = \tilde{A}_h \psi^{(k)}, \quad (22)$$

и что расположение всех собственных чисел  $\lambda_k$  на отрезке  $-1 \leq \lambda \leq 1$  необходимо и достаточно для выполнения неравенств

$$\|R_h^n\|_{B_h} \leq 1, \quad (23)$$

где норма оператора задается с помощью скалярного умножения (21).

**6. Критерии устойчивости Самарского.** В предложенной А. А. Самарским [23], [24] теории устойчивости широкого класса разностных схем в гильбертовом пространстве указаны необходимые и достаточные условия устойчивости в терминах линейных неравенств между операторными коэффициентами этих схем, а также получены другие результаты.

Приведем здесь лишь два результата из этой теории.

Пусть  $U'_h$  — евклидово пространство с некоторым скалярным умножением  $(u, v) \equiv [u, v]$ , и пусть оператор  $R_h$ ,  $v = R_h u$ ,  $u, v \in U'_h$ , задан равенством

$$B_h \frac{v - u}{\tau} + A_h u = 0, \quad (24)$$

где  $A_h$  и  $B_h$  — самосопряженные операторы, причем  $B_h > 0$ . Определим энергетическую норму  $\|u\|_{B_h}$  в пространстве  $U'_h$ , положив

$$\|u\|_{B_h}^2 = [B_h u, u] \equiv (u, u)_{B_h}. \quad (25)$$

Тогда справедлива следующая

**Теорема 2. Условия**

$$0 \leq A_h \leq \frac{2}{\tau} B_h \quad (26)$$

*необходимы и достаточны для выполнения неравенств*

$$\|R_h^p\| \leq 1, \quad p \geq 0. \quad (27)$$

**Доказательство.** Определим самосопряженный оператор  $\tilde{A}_h$  равенством  $\tilde{A}_h \equiv B_h - \tau A_h$ . Тогда (24) равносильно равенству (20), а условия (26) равносильны условиям  $-B_h \leq \tilde{A}_h \leq B_h$ , т. е. условиям

$$-[B_h u, u] \leq [\tilde{A}_h u, u] \leq [B_h u, u]. \quad (28)$$

Как показано в п. 5, оператор  $R_h$  является самосопряженным в смысле скалярного умножения (21) и утверждение теоремы равносильно утверждению, что все собственные числа  $\lambda_k$  оператора  $R_h$  лежат на отрезке  $-1 \leq \lambda \leq 1$  в том и только том случае, если выполнены условия (28). Докажем это последнее утверждение.

Пусть выполнены условия (28). Умножим равенство (22) скалярно на собственную функцию  $\psi^{(k)}$  оператора  $R_h$ . Получим  $\lambda_k [B_k \psi^{(k)}, \psi^{(k)}] = [\tilde{A}_h \psi^{(k)}, \psi^{(k)}]$ , откуда

$$|\lambda_k| = \left| \frac{[\tilde{A}_h \psi^{(k)}, \psi^{(k)}]}{[B_h \psi^{(k)}, \psi^{(k)}]} \right| \leq 1.$$

Обратно, пусть  $\max |\lambda_k| \leq 1$ . Покажем, что выполнены условия (28). Пусть  $u = \sum c_k \psi^{(k)}$  — разложение произвольного элемента  $u \in U'_h$  по ортонормальному в смысле скалярного умножения (21) базису  $\{\psi^{(k)}\}$ . Тогда

$$\begin{aligned} |[\tilde{A}_h u, u]| &= |[\tilde{A}_h \sum c_k \psi^{(k)}, \sum c_k \psi^{(k)}]| = |[\sum c_k \tilde{A}_h \psi^{(k)}, \sum c_k \psi^{(k)}]| = \\ &= |[\sum c_k \lambda_k B_h \psi^{(k)}, \sum c_k \psi^{(k)}]| = |[B_h \sum c_k \lambda_k \psi^{(k)}, \sum c_k \psi^{(k)}]| = \\ &= |(\sum c_k \lambda_k \psi^{(k)}, \sum c_k \psi^{(k)})_{B_h}| = \sum c_k^2 |\lambda_k| \leq \\ &\leq \sum c_k^2 = (u, u)_{B_h} = [B_h u, u]. \end{aligned}$$

Отсюда  $[B_h u, u] \geq |[\tilde{A}_h u, u]|$ , что равносильно условиям (28). Теорема доказана.

Заметим, что проверка условий (28) равносильна проверке того, будут ли неотрицательны все собственные числа самосопряженных в смысле скалярного умножения  $[u, v]$  операторов  $B_h - \tilde{A}_h$  и  $B_h + \tilde{A}_h$ .

Приведем без доказательства еще один критерий устойчивости, применимый к разностным схемам (24), для которых  $B_h > 0$ ,  $A_h = A_h^* > 0$ . Введем в пространстве  $U'_h$  энергетическую норму  $\|u\|_{A_h}$ , положив  $\|u\|_{A_h}^2 = [A_h u, u]$ .

**Теорема 3.** *Выполнение условия  $B_h \geq \frac{\tau}{2} A_h$  необходимо и достаточно для того, чтобы имело место неравенство  $\|R_h\|_{A_h} \leq 1$ .*

Теорема 3 содержится в п. 4 § 1 гл. VI книги [23] и доказывается без помощи спектрального подхода, который здесь не удастся применить из-за несамосопряженности оператора  $B_h$ .

### ЗАДАЧИ

1. Пусть оператор  $R_h$ ,  $b = R_h a$ , задан формулами

$$\left. \begin{aligned} b_m &= (1-r)a_m + ra_{m+1}, \quad m = 0, 1, \dots, M-1, \\ b_M &= 0. \end{aligned} \right\}$$

Показать, что в пространстве  $U'_h$  сеточных функций  $\{a_m\}$ ,  $m = 0, 1, \dots, M$ , нельзя задать скалярное произведение так, чтобы оператор  $R_h$  стал самосопряженным.

## СПЕКТРАЛЬНЫЙ ПРИЗНАК УСТОЙЧИВОСТИ НЕСАМОСОПРЯЖЕННЫХ ЭВОЛЮЦИОННЫХ КРАЕВЫХ ЗАДАЧ

Здесь мы покажем, что по спектру несамосопряженного оператора  $R_h$  нельзя судить об устойчивости разностной краевой задачи в ограниченной области, введем понятие спектра *семейства* операторов  $\{R_h\}$  и рассмотрим спектральную постановку вопроса об устойчивости, остающуюся разумной и в случае несамосопряженных разностных краевых задач в ограниченных областях. Будет указан необходимый и близкий к достаточному спектральный признак устойчивости.

### § 44. Спектр семейства операторов $\{R_h\}$

**1. Необходимость усовершенствования спектрального признака устойчивости.** В гл. 13 было показано, что обычно эволюционные разностные краевые задачи можно привести к виду

$$\left. \begin{aligned} u^{p+1} &= R_h u^p + \tau \rho^p, \\ u^0 &\text{ задано} \end{aligned} \right\} \quad (1)$$

так, чтобы устойчивость на интервале времени  $0 \leq t \leq T$  была равносильна равномерной по  $h$  ограниченности норм степеней оператора перехода  $R_h$ , т. е. оценке

$$\|R_h^p\| < K, \quad p = 1, 2, \dots, [T/\tau], \quad (2)$$

где  $\lambda$  — шаг сетки по времени,  $\tau = \tau(h)$ .

Было установлено, что расположение собственных значений оператора  $R_h$  внутри круга

$$|\lambda| < 1 + c\tau \quad (3)$$

на комплексной плоскости необходимо для выполнения условия (2), т. е. для устойчивости. В § 43 было показано, что в случае самосопряженного оператора  $R_h$  условие (3) является не только необходимым, но и достаточным условием равномерной

ограниченности (2) норм степеней оператора  $R_h$ . Этот же факт установлен в § 25 и для разностной задачи Коши с постоянными коэффициентами в случае двуслойных разностных схем относительно одной неизвестной функции независимо от того, имеет ли место самосопряженность. Однако в общем случае не-самосопряженных разностных краевых задач в ограниченных областях необходимым признаком (3) очень далек от достаточного и совершенно неадекватен вопросу о равномерной ограниченности (2) норм  $\|R_h^p\|$  степеней оператора  $R_h$ .

Это показывает следующий пример.

Пример. Для разностной краевой задачи

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= \varphi(x_m, t_p), \\ u_M^p &= 0, \quad p = 0, 1, \dots, [T/\tau], \\ u_m^0 &= \psi(x_m), \quad m = 0, 1, \dots, M; Mh = 1, \end{aligned} \right\} \quad (4)$$

аппроксимирующей задачу

$$\begin{aligned} u_t - u_x &= \varphi(x, t), \\ u(1, t) &= 0, \quad 0 < x < 1, \quad 0 < t < T, \\ u(x, 0) &= \psi(x) \end{aligned}$$

при естественном приведении к каноническому виду (1) оператор  $R_h$ ,  $v = R_h u$ , задается формулами

$$\begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, \quad m = 0, 1, \dots, M-1, \\ v_M &= 0, \quad r = \tau/h. \end{aligned}$$

Его матрица имеет вид

$$R_h = \begin{bmatrix} 1-r & r & 0 & \dots & 0 & 0 \\ 0 & 1-r & r & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1-r & r \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \quad (5)$$

Спектр матрицы состоит из ее собственных значений, т. е. из корней уравнения

$$\det(R_h - \lambda E) = 0 \quad \text{или} \quad -\lambda(1-r-\lambda)^M = 0.$$

Корни этого уравнения  $\lambda = 0$  и  $\lambda = 1-r$  образуют спектр оператора  $R_h$  при любом  $h$ . Этот спектр лежит в единичном круге  $|\lambda| \leq 1$  при  $0 < r \leq 2$ . Между тем для схемы (4) при  $1 < r \leq 2$  не выполнено условие Куранта, Фридрихса и Леви, так что устойчивости  $\|R_h^p\| < K$  ни в какой разумной норме быть не может.

В самом деле, покажем, что в случае  $T \geq 1$  и нормы  $\|u\| = \max_m |u_m|$  справедливо неравенство

$$\max_{\rho=1, 2, \dots, [T/\tau]} \|R_h^\rho\| \geq |1 - 2r|^{1/h} = \rho^{1/h}. \quad (6)$$

При  $r > 1$  также  $\rho > 1$ , так что при  $h \rightarrow 0$  и  $\tau = rh \rightarrow 0$  величина  $\max_\rho \|R_h^\rho\|$  экспоненциально возрастает и условие  $\|R_h^\rho\| < K$  грубо нарушается. Для доказательства неравенства (6) заметим, что в случае  $u_m^i = (-1)^m$ ,  $m = 0, 1, \dots, M$ , значения  $u_m^\rho$  функции

$$u^\rho = R_h^\rho u^0, \quad \rho = 1, 2, \dots, M \text{ при } m = 0, 1, \dots, M - \rho$$

задаются формулой

$$u_m^\rho = (-1)^m (1 - 2r)^\rho, \quad m = 0, 1, \dots, M - \rho.$$

Поэтому

$$\|R_h^\rho u^0\| \geq |1 - 2r|^\rho \|u^0\|, \quad \rho \leq M,$$

так что при этих значениях  $\rho$ ,  $\rho = 1, 2, \dots, M$ ,

$$\|R_h^\rho\| \geq |1 - 2r|^\rho, \quad \rho = 1, 2, \dots; \quad M = 1/h,$$

и неравенство (6) доказано.

Итак, установлено, что необходимым спектральный признак (3) равномерной ограниченности  $\|R_h^\rho\| < K$ , использующий собственные значения операторов  $R_h$ , слишком груб в случае несамосопряженных операторов  $R_h$ : в нашем примере он не улавливает неустойчивость, имеющую место при  $1 < r \leq 2$ .

**2. Определение спектра семейства операторов.** Пусть линейный оператор  $R_h$  определен на линейном нормированном пространстве  $U_h$ .

Будем обозначать через  $\{R_{hi}\}$  совокупность операторов  $R_h$  при всех тех значениях  $h$ , которые принимает параметр  $h$ , характеризующий густоту сетки. По самой природе разностных схем шаг сетки  $h$  может принимать сколь угодно малые положительные значения.

Комплексное число  $\lambda$  будем называть *точкой спектра семейства операторов*  $\{R_h\}$ , если для любых положительных  $h_0$  и  $\varepsilon$  можно указать такое  $h$ ,  $h < h_0$ , что неравенство

$$\|R_h u - \lambda u\| < \varepsilon \|u\|$$

имеет некоторое решение  $u$ ,  $u \in U_h$ .

Совокупность всех таких чисел  $\lambda$  будем называть *спектром семейства операторов*  $\{R_h\}$ .

### 3. Необходимое условие устойчивости.

**Теорема 1.** Пусть хотя бы одна точка  $\lambda_0$  спектра семейства операторов  $\{R_h\}$  лежит вне единичного круга комплексной плоскости, так что  $|\lambda_0| > 1$ . В таком случае нельзя указать общую для всех  $h$  постоянную  $K$  такую, чтобы выполнялось неравенство

$$\|R_h^p\| < K, \quad (7)$$

в котором  $p$  пробегает целые значения от 0 до  $p = p_0(h)$ , где  $p_0(h) \rightarrow \infty$  при  $h \rightarrow 0$ .

**Доказательство.** Допустим сначала, что не существует чисел  $h_0 > 0$  и  $c > 0$  таких, что при всех  $h < h_0$  справедлива оценка

$$\|R_h\| < c. \quad (8)$$

При этом допущении доказываемое утверждение очевидно. Поэтому остается рассмотреть случай, когда существуют  $h_0 > 0$  и  $c > 0$  такие, что при  $h < h_0$  неравенство (8) справедливо.

Положим  $|\lambda_0| = 1 + \delta$ , где  $\lambda_0$  — та точка спектра, для которой  $|\lambda_0| > 1$ . Задав произвольно числом  $K$ , выберем  $p$  и  $\varepsilon$  так, чтобы выполнялись неравенства

$$(1 + \delta)^p > 2K, \\ 1 - (1 + c + c^2 + \dots + c^{p-1})\varepsilon > \frac{1}{2}.$$

По определению точки спектра семейства операторов  $\{R_h\}$ , можно указать сколь угодно малые положительные  $h$ , при которых существует вектор  $u \in U_h$ , являющийся решением неравенства

$$\|R_h u - \lambda_0 u\| < \varepsilon \|u\|. \quad (9)$$

Положим

$$R_h u = \lambda_0 u + z. \quad (10)$$

Ясно, что  $\|z\| < \varepsilon \|u\|$ . Далее, из (10) можно вывести, что

$$R_h^p u = \lambda_0^p u + (\lambda_0^{p-1} z + \lambda_0^{p-2} R_h z + \dots + R_h^{p-1} z).$$

Поскольку  $|\lambda_0| > 1$ , то

$$\|\lambda_0^{p-1} z + \lambda_0^{p-2} R_h z + \dots + R_h^{p-1} z\| \leq \\ \leq |\lambda_0|^p (1 + \|R_h\| + \|R_h^2\| + \dots + \|R_h^{p-1}\|) \varepsilon \|u\|,$$

а следовательно,

$$\|R_h^p u\| \geq |\lambda_0|^p [1 - \varepsilon (1 + c + c^2 + \dots + c^{p-1})] \|u\| > \\ > (1 + \delta)^p \frac{1}{2} \|u\| \geq 2K \frac{1}{2} \|u\| = K \|u\|.$$

Число  $h$  во всем этом построении можно считать настолько малым, чтобы  $p$  было меньше, чем  $p_0(h)$ .

Ввиду произвольности  $K$  доказано наше утверждение о том, что *расположение всех точек спектра семейства операторов  $\{R_h\}$  внутри или на границе единичного круга  $|\lambda| \leq 1$  необходимо для выполнения оценки  $\|R_h^p\| < K$ .*

**4. Обсуждение понятия спектра семейства операторов  $\{R_h\}$ .** Начнем с того, что обратим внимание на аналогию между определением точки спектра семейства операторов  $\{R_h\}$  и следующим определением точки спектра какого-либо оператора  $R$ , которое приводится в курсах функционального анализа. Мы будем в качестве  $R$  брать оператор  $R_h$  при некотором фиксированном  $h$ .

Точка  $\lambda$  на комплексной плоскости называется *точкой спектра оператора  $R_h$* , если при любом положительном  $\varepsilon$  неравенство

$$\|R_h u - \lambda u\| < \varepsilon \|u\|$$

имеет решение  $u$ , принадлежащее пространству  $U'_h$ , на котором определен оператор  $R_h$ .

При сравнении определений точки спектра *семейства операторов  $\{R_h\}$*  и точки спектра оператора  $R_h$  может возникнуть мысль, что спектр семейства  $\{R_h\}$  состоит из тех точек комплексной плоскости, которые получаются путем предельного перехода при  $h \rightarrow 0$  из точек спектра оператора  $R_h$ , когда  $h \rightarrow 0$  по всевозможным подпоследовательностям. Но, вообще говоря, это предположение ошибочно.

Рассмотрим оператор  $R_h$ ,  $v = R_h u$ , задаваемый равенствами

$$\left. \begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, & m=0, 1, \dots, M-1, \\ v_M &= 0, & Mh=1. \end{aligned} \right\} \quad (11)$$

Оператор (11) действует в  $(M+1)$ -мерном линейном пространстве и описывается матрицей (5). Известно, что спектр матрицы состоит из ее собственных значений, т. е. из корней  $\lambda$  уравнения  $\det(R_h - \lambda E) = 0$ . Мы вычислили эти собственные значения в п. 1. Это  $\lambda = 0$  и  $\lambda = 1 - r$ . Таким образом, спектр оператора  $R_h$  при любом  $h$  состоит из двух точек  $\lambda = 0$  и  $\lambda = 1 - r$ , не зависящих от  $h$ . Однако спектр семейства операторов  $\{R_h\}$ , как будет показано в § 45, состоит не только из этих двух точек, чего, казалось бы, можно было ожидать, а еще и из всех точек круга  $|\lambda - 1 + r| < r$  радиуса  $r$  с центром в точке  $\lambda = 1 - r$  (рис. 27, стр. 248). При  $r \leq 1$  спектр семейства операторов  $\{R_h\}$  лежит в единичном круге  $|\lambda| \leq 1$ , а при  $r > 1$  этот необходимый спектральный признак устойчивости не выполнен: неравенство  $\|R_h^p\| < K$  не может выполняться равномерно по  $h$ .



На рис. 53 приведены графики зависимости величин  $\|R_h^p\|$  от  $p\tau = prh$  в случае  $r = 3/2$  для различных значений  $h$ . В этом случае спектр каждого оператора  $R_h$  состоит из двух точек  $\lambda = 0$  и  $\lambda = -1/2$ , лежащих внутри единичного круга. Этим предопределяется поведение графика величины  $\|R_h^p\|$  при больших значениях  $p\tau$ . Величина  $\|R_h^p\|$  стремится к нулю при  $p\tau \rightarrow \infty$ , т. е. ось абсцисс является асимптотой (в подробных курсах алгебры доказывается, что норма степеней матрицы стремится к нулю при росте показателя степени, если все собственные значения матрицы по модулю меньше единицы).

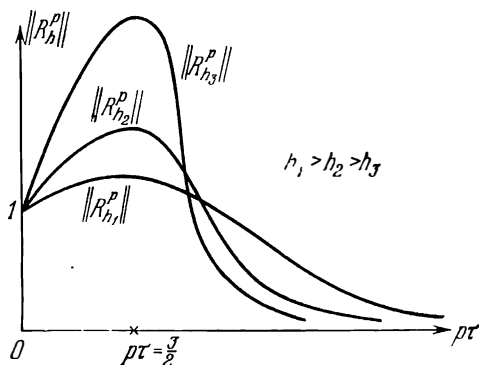


Рис. 53.

То обстоятельство, что спектр семейства операторов  $\{R_h\}$  не целиком лежит в единичном круге, сказывается на поведении величины  $\|R_h^p\|$  при  $h \rightarrow 0$  и при не слишком больших значениях  $p\tau$ . Наибольшее значение величины  $\|R_h^p\|$  на отрезке  $0 < p\tau < T$  ( $T$  — произвольная положительная постоянная) быстро растет при уменьшении  $h$ . Но это и означает неустойчивость на отрезке  $0 < t < T$ , в то время как поведение  $\|R_h^p\|$  при  $p\tau \rightarrow \infty$ , связанное со спектром каждого отдельного оператора  $R_h$ , совершенно несущественно при исследовании устойчивости.

**5. Близость необходимого признака устойчивости к достаточному.** Справедлива следующая

**Теорема 2.** Пусть оператор  $R_h$  определен на конечномерном при каждом  $h$  нормированном пространстве  $U'_h$  и равномерно по  $h$  ограничен некоторой постоянной  $c$ :

$$\|R_h\| < c. \quad (12)$$

Пусть, далее, спектр семейства операторов  $\{R_h\}$  целиком лежит в замкнутом единичном круге  $|\lambda| \leq 1$ .

Тогда при любом  $\varepsilon > 0$  нормы степеней операторов  $R_h$  удовлетворяют оценке

$$\|R_h^p\| \leq A(\varepsilon)(1 + \varepsilon)^p, \quad (13)$$

где  $A \neq A(\varepsilon)$  зависит только от  $\varepsilon$ , но не от  $h$ .

Факт, устанавливаемый этой теоремой, означает, что расположение спектра семейства операторов  $\{R_h\}$  в единичном круге

не только необходимо для устойчивости, но и гарантирует от «грубой» неустойчивости. При выполнении условий теоремы величина

$$\max_{1 \leq \rho \leq [T/\tau]} \|R_h^\rho\|$$

при  $h \rightarrow 0$  остается ограниченной либо растет медленнее степени  $\rho^{[T/\tau]}$  с любым основанием  $\rho = 1 + \varepsilon$ , превосходящим единицу.

**Доказательство.** Предварительно покажем, что если спектр семейства операторов  $\{R_h\}$  лежит в круге  $|\lambda| \leq \rho$ , то для любого  $\lambda$ , удовлетворяющего неравенству  $|\lambda| \geq \rho + \varepsilon$ ,  $\varepsilon > 0$ , существуют числа  $A = A(\varepsilon)$  и  $h_0 > 0$  такие, что при любом  $h < h_0$  и любом  $u \in U_h$ ,  $u \neq 0$ , выполнено неравенство

$$\|R_h u - \lambda u\| > \frac{\rho + \varepsilon}{A(\varepsilon)} \|u\|. \quad (14)$$

Допустим противное. Тогда найдутся  $\varepsilon > 0$ ; последовательности чисел  $h_k > 0$ ,  $h_k \rightarrow 0$ ; комплексных чисел  $\lambda_k$ ,  $|\lambda_k| > \rho + \varepsilon$ ; векторов  $u_{h_k} \in U_{h_k}$  такие, что

$$\|R_{h_k} u_{h_k} - \lambda_k u_{h_k}\| < \frac{\rho + \varepsilon}{k} \|u_{h_k}\|. \quad (15)$$

При достаточно больших значениях  $k$ , при которых  $\frac{\rho + \varepsilon}{k} < 1$ , числа  $\lambda_k$  в силу (12) не могут лежать вне круга  $|\lambda| \leq \rho + 1$ , так как вне этого круга

$$\|R_{h_k} u - \lambda u\| \geq (|\lambda| - \|R_{h_k}\|) \|u\| \geq \|u\|.$$

Таким образом, последовательность  $\lambda_k$  ограничена, а следовательно, имеет предельную точку  $\tilde{\lambda}$ ,  $|\tilde{\lambda}| \geq \rho + \varepsilon$ . Легко видеть, что в силу (15) точка  $\tilde{\lambda}$  принадлежит спектру семейства операторов  $\{R_h\}$ , вопреки предположению, что спектр лежит в круге  $|\lambda| \leq \rho$ .

Пусть теперь  $R$  — линейный оператор, переводящий некоторое конечномерное нормированное пространство  $U$  в себя. Пусть для любого комплексного  $\lambda$ ,  $|\lambda| \geq r > 0$ , и любого  $u \in U$  при некотором  $a = \text{const} > 0$  справедливо неравенство

$$\|Ru - \lambda u\| \geq a \|u\|. \quad (16)$$

Тогда

$$\|R^p\| \leq \frac{r^{p+1}}{a}, \quad p = 1, 2, \dots \quad (17)$$

Неравенство (17) вытекает из следующего известного равенства:

$$R^p = -\frac{1}{2\pi i} \oint_{|\lambda|=r} \lambda^p (R - \lambda E)^{-1} d\lambda, \quad (18)$$

и условия (16), в силу которого  $\|(R - \lambda E)^{-1}\| \leq \frac{1}{a}$ . Для доказательства неравенства (13) положим  $a = \frac{\rho + \varepsilon}{A(\varepsilon)}$ ,  $r = \rho + \varepsilon$ ,  $\rho = 1$ ,  $R = R_h$ . Тогда (17) совпадает с (13).

В заключение напомним доказательство равенства (18). Положим

$$u^{p+1} = Ru^p, \quad U(\lambda) = \sum_{p=0}^{\infty} \frac{u^p}{\lambda^p}.$$

Умножим обе части равенства  $u^{p+1} = Ru^p$  на  $\lambda^{-p}$  и просуммируем по  $p$  от  $p=0$  до  $p=\infty$ . Получим

$$\lambda U(\lambda) - \lambda u^0 = RU(\lambda),$$

или

$$(R - \lambda E)U(\lambda) = -\lambda u^0, \quad U(\lambda) = -\lambda (R - \lambda E)^{-1} u^0.$$

Из определения  $U(\lambda)$  видно, что  $u^p$  является вычетом вектор-функции  $\lambda^{p-1}U(\lambda)$ :

$$u^p = \frac{1}{2\pi i} \oint_{|\lambda|=r} \lambda^{p-1} U(\lambda) d\lambda = -\frac{1}{2\pi i} \oint_{|\lambda|=r} \lambda^p (R - \lambda E)^{-1} u^0 d\lambda.$$

Но  $u^p = R^p u^0$ , так что последнее равенство равносильно операторному равенству (18).

В этом параграфе мы сформулировали спектральную постановку задачи об устойчивости эволюционных разностных схем, имеющую смысл для любых эволюционных разностных схем, приводимых к виду

$$\left. \begin{aligned} u^{p+1} &= R_h u^p + \tau \rho^p, \\ u^0 &\text{ задано} \end{aligned} \right\}$$

так, чтобы выполнение условия

$$\|R_h^p\| < K, \quad p = 1, 2, \dots, [T/\tau],$$

было равносильно устойчивости. Это могут быть двуслойные, многослойные схемы, схема расщепления и так далее для задач на отрезке, в многомерных или составных областях.

Эта спектральная постановка задачи требует выяснить, лежит ли спектр семейства операторов  $\{R_h\}$  в единичном круге  $|\lambda| \leq 1$ .

#### § 45. Алгоритм вычисления спектра семейства разностных операторов над сеточными функциями на отрезке

В этом параграфе мы опишем алгоритм вычисления спектра семейства разностных операторов  $\{R_h\}$  над пространствами сеточных функций (или вектор-функций) на отрезке. За норму функции (или вектор-функции) примем максимум абсолютных

величин значений, принимаемых функцией (или компонентами вектор-функции).

**1. Характерный пример.** Семейство операторов  $\{R_h\}$ ,  $v = R_h u$ , определим равенствами

$$\left. \begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, \quad m=0, 1, \dots, M-1, \\ v_M &= 0. \end{aligned} \right\} \quad (1)$$

Оператор  $R_h$ , определенный равенствами (1), возникает при естественном приведении разностной краевой задачи

$$\left. \begin{aligned} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} &= \varphi(x_m, t_p), \\ p &= 0, 1, \dots, [T/\tau] - 1, \\ u_M^{p+1} = 0, u_m^0 &= \psi(x_m), \quad m=0, 1, \dots, M-1, \end{aligned} \right\} \quad (2)$$

к виду

$$u^{p+1} = R_h u^p + \tau \rho^p, \quad u^0 \text{ задано.}$$

Соотношения (2) являются разностным аналогом дифференциальной краевой задачи

$$\begin{aligned} u_t - u_x &= \varphi(x, t), \quad 0 \leq x \leq 1, \quad 0 \leq t \leq T, \\ u(x, 0) &= \psi(x), \quad u(1, t) = 0. \end{aligned}$$

Мы уже рассматривали разностную схему (2) в п. 2 § 26 в качестве примера, иллюстрирующего применение признака Бабенко — Гельфанда. Напомним, что согласно этому признаку исследование исходной задачи на отрезке следует разбить на исследование трех вспомогательных задач: задачи без боковых границ, задачи с одной только левой границей и задачи с одной только правой границей, для каждой из которых надо найти все собственные значения операторов перехода от  $u^p$  к  $u^{p+1}$ .

Оказывается, что алгоритм вычисления спектра семейства операторов  $\{R_h\}$  совпадает с процедурой Бабенко — Гельфанда.

Чтобы описать алгоритм вычисления спектра семейства операторов  $\{R_h\}$ , наряду с оператором  $R_h$ , заданным равенствами (1), рассмотрим три вспомогательных оператора:  $\overleftrightarrow{R}$ ,  $\overrightarrow{R}$  и  $\overleftarrow{R}$ . Оператор  $\overleftrightarrow{R}$ ,  $v = \overleftrightarrow{R}u$ , задается на линейном пространстве ограниченных функций  $u = \{\dots, u_{-1}, u_0, u_1, \dots\}$ , определенных на всей сеточной прямой  $-\infty < mh < \infty$ , по формуле

$$v_m = (1-r)u_m + ru_{m+1}, \quad m=0, \pm 1, \dots \quad (3)$$

Эта формула получается из равенств (1) при удалении левой границы в  $-\infty$ , а правой в  $+\infty$ , что отражено стрелкой с двумя концами в обозначении оператора:  $\overleftrightarrow{R}$ . Оператор  $\overrightarrow{R}$ ,

$v = \vec{R}u$ , задается на линейном пространстве сеточных функций  $u = (u_0, u_1, \dots, u_m, \dots)$ , определенных на сеточной полупрямой  $x_m = mh, m = 0, 1, 2, \dots$ , и стремящихся к нулю при  $m \rightarrow +\infty$ . Он задается формулой

$$v_m = (1 - r)u_m + ru_{m+1}, \quad m = 0, 1, \dots \tag{4}$$

Эта формула получается из формул (1) при удалении правой границы в  $+\infty$ , что отражено мнемоническим значком  $\rightarrow$  в обозначении оператора:  $\vec{R}$ .

Наконец, оператор  $\overleftarrow{R}$ ,  $v = \overleftarrow{R}u$ , над функциями

$$u = (\dots, u_m, \dots, u_{M-1}, u_M), \quad u_m \rightarrow 0 \text{ при } m \rightarrow -\infty,$$

определенными на сеточной полупрямой  $x_m = mh, m = \dots, -2, -1, 0, 1, \dots, M$ , зададим формулами

$$\left. \begin{aligned} v_m &= (1 - r)u_m + ru_{m+1}, \quad m = \dots, -1, 0, 1, \dots, M - 1, \\ v_M &= 0. \end{aligned} \right\} \tag{5}$$

Эти формулы получились из формул (1) при удалении левой границы в  $-\infty$ , что также отражено в обозначении оператора:  $\overleftarrow{R}$ .

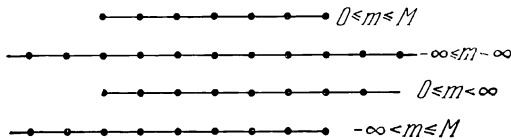


Рис. 54.

Мы видим, что операторы  $\overleftrightarrow{R}$ ,  $\vec{R}$  и  $\overleftarrow{R}$  от  $h$  не зависят. Области определения функций  $u = \{u_m\}$  для операторов (1), (3), (4) и (5) показаны на рис. 54. Будет показано, что совокупность собственных значений всех трех операторов и составляет спектр семейства разностных операторов  $\{R_h\}$ .

Собственные значения операторов  $\overleftrightarrow{R}$ ,  $\vec{R}$  и  $\overleftarrow{R}$  мы уже вычисляли в § 26, однако воспроизведем здесь их вычисление, так как, прежде чем переходить к доказательству сформулированного утверждения, надо отчетливо представить себе структуру собственных функций операторов  $\overleftrightarrow{R}$ ,  $\vec{R}$  и  $\overleftarrow{R}$ .

Прежде всего выясним, каково множество точек  $\lambda$  на комплексной плоскости, для которых уравнение

$$\overleftrightarrow{R}u - \lambda u = 0$$

имеет ограниченное решение  $u = \{u_m\}$ ,  $m = 0, \pm 1, \dots$ . Эти числа  $\lambda$  и есть собственные значения оператора  $\overleftarrow{R}$ . В нашем примере уравнение  $\overleftrightarrow{R}u - \lambda u = 0$  имеет вид

$$(1 - r - \lambda)u_m + ru_{m+1} = 0, \quad m = 0, \pm 1, \dots$$

Всякое решение этого обыкновенного разностного уравнения первого порядка, как вытекает из § 1, может лишь постоянным множителем отличаться от сеточной функции  $u_m = q^m$ ,  $m = 0, \pm 1, \dots$ , где  $q$  — корень характеристического уравнения  $(1 - r - \lambda) + rq = 0$ . Связь между числами  $\lambda$  и  $q$  можно записать также в форме

$$\lambda = 1 - r + rq.$$

Решение  $u_m = q^m$  ограничено при  $m \rightarrow +\infty$  и при  $m \rightarrow -\infty$  только в том случае, если  $|q| = 1$ ,  $q = e^{i\alpha}$ ,  $0 \leq \alpha \leq 2\pi$ . Поэтому множество тех значений  $\lambda$ , при которых решение  $u_m = q^m$  ограничено, получается по формуле

$$\lambda = 1 - r + rq = 1 - r + re^{i\alpha},$$

когда  $q = e^{i\alpha}$  пробегает единичную окружность  $|q| = 1$  на комплексной плоскости. Точка  $\lambda$  пробегает при этом окружность  $\overleftarrow{\Lambda}$  радиуса  $r$  с центром в точке  $1 - r$  (рис. 26, а, стр. 289).

Вычислим собственные значения оператора  $\overrightarrow{R}$ , т. е. те  $\lambda$ , при которых уравнение

$$\overrightarrow{R}u - \lambda u = 0$$

имеет решение  $u = (u_0, u_1, \dots, u_m, \dots)$ , стремящееся к нулю при  $m \rightarrow +\infty$ .

Уравнение  $\overrightarrow{R}u - \lambda u = 0$  в развернутом виде можно записать так:

$$(1 - r - \lambda)u_m + ru_{m+1} = 0, \quad m = 0, 1, \dots$$

Его решение  $u_m = q^m$ ,  $m = 0, 1, \dots$ , стремится к нулю при  $m \rightarrow +\infty$ , если  $|q| < 1$ . Соответствующие собственные значения  $\lambda = 1 - r + rq$  заполняют при этом внутренность круга  $\overleftarrow{\Lambda}$  радиуса  $r$  с центром в точке  $1 - r$  (рис. 26, б):

Алгоритм вычисления собственных чисел оператора  $\overrightarrow{R}$  аналогичен алгоритму вычисления собственных чисел оператора  $\overleftarrow{R}$ .

Уравнение  $\overleftarrow{R}u - \lambda u = 0$  запишем развернуто:

$$\left. \begin{aligned} (1 - r - \lambda)u_m + ru_{m+1} &= 0, \quad m = \dots, -1, 0, 1, \dots, M-1, \\ -\lambda u_M &= 0. \end{aligned} \right\} \quad (6)$$

Всякая сеточная функция  $u = \{u_m\}$ ,  $m = M, M-1, \dots$ , удовлетворяющая первому из этих соотношений, с точностью до постоянного множителя по-прежнему имеет вид  $u_m = q^m$ , причем  $\lambda$  и  $q$  по-прежнему связаны равенством  $\lambda = 1 - r + rq$ . Решение  $u_m = q^m$ ,  $m = M, M-1, \dots$ , стремится к нулю при  $m \rightarrow -\infty$ , если  $|q| > 1$ . Второе соотношение (6), т. е. равенство  $-\lambda u_M = 0$ , накладывает на решение  $u_m = q^m$  дополнительное требование  $-\lambda u_M = -\lambda q^M = 0$  или  $\lambda = 0$ . Если точка  $\lambda = 0$  лежит вне круга радиуса  $r$  с центром в точке  $1 - r$ , изображенного на рис. 26, в, т. е. если  $r < 1/2$ , то ей соответствует некоторое значение  $q$ ,  $|q| > 1$ . Множество  $\overset{\leftarrow}{\Lambda}$  тех  $\lambda$ , при которых уравнение  $\overset{\leftarrow}{R}u - \lambda u = 0$  имеет стремящееся к нулю при  $m \rightarrow -\infty$  решение, состоит из одной этой точки  $\lambda = 0$ . В случае  $r \geq 1/2$ , как следует из проделанного анализа, уравнение  $\overset{\leftarrow}{R}u - \lambda u = 0$  не имеет стремящихся к нулю при  $m \rightarrow -\infty$  решений ни при каком комплексном (или вещественном)  $\lambda$ .

Объединение собственных значений операторов  $\overset{\leftrightarrow}{R}$ ,  $\overset{\rightarrow}{R}$  и  $\overset{\leftarrow}{R}$  изображено для случая  $r < 1/2$  на рис. 27, а, а для случая  $r > 1/2$  на рис. 27, б и 27, в.

Докажем теперь, что спектр семейства операторов  $\{R_h\}$  совпадает с объединением  $\Lambda$  множеств  $\overset{\leftrightarrow}{\Lambda}$ ,  $\overset{\rightarrow}{\Lambda}$  и  $\overset{\leftarrow}{\Lambda}$  собственных значений вспомогательных операторов  $\overset{\leftrightarrow}{R}$ ,  $\overset{\rightarrow}{R}$ ,  $\overset{\leftarrow}{R}$ .

Надо показать, что каждая точка множества  $\Lambda$  принадлежит спектру семейства разностных операторов  $\{R_h\}$  и что других точек спектр не содержит.

Сначала покажем, что всякая точка  $\lambda_0 \in \Lambda$  принадлежит спектру семейства разностных операторов. Для этого достаточно установить, что, каково бы ни было  $\varepsilon > 0$ , неравенство

$$\|R_h u - \lambda_0 u\| < \varepsilon \|u\| \quad (7)$$

имеет решение  $u$  при всех достаточно малых положительных значениях  $h$ . Решение  $u = (u_0, u_1, \dots, u_M)$  можно назвать «почти собственным вектором» оператора  $R_h$ , поскольку решение уравнения  $R_h u - \lambda u = 0$  в алгебре принято называть собственным вектором.

Построения, с помощью которых проводится доказательство, зависят от того, какому из трех множеств  $\overset{\leftrightarrow}{\Lambda}$ ,  $\overset{\rightarrow}{\Lambda}$  или  $\overset{\leftarrow}{\Lambda}$  принадлежит точка  $\lambda_0$ . Начнем со случая  $\lambda_0 \in \overset{\leftrightarrow}{\Lambda}$ . Покажем, что при любом  $\varepsilon > 0$  и всех достаточно малых  $h$  неравенство (7) имеет решение  $u$ .

Переходим к построению функции  $u = (u_0, u_1, \dots, u_M)$ . По определению множества  $\overset{\leftrightarrow}{\Lambda}$  существует  $q_0$ ,  $|q_0| = 1$ , такое, что

$\lambda_0 = (1 - r) + rq_0$ , а уравнение  $(1 - r - \lambda_0)v_m + rv_{m+1} = 0$ ,  $m = 0, \pm 1, \dots$ , имеет ограниченное решение  $v_m = q_0^m$ ,  $m = 0, \pm 1, \dots$ . Будем рассматривать это решение только при  $m = 0, 1, \dots, M$ , сохраняя обозначение  $v$ . Вектор

$$v = (v_0, v_1, \dots, v_M) = (1, q_0, \dots, q_0^M),$$

очевидно, удовлетворял бы уравнению  $R_h v - \lambda_0 v = 0$ , которое в развернутом виде записывается соотношениями

$$\left. \begin{aligned} (1 - r - \lambda_0)v_m + rv_{m+1} &= 0, \quad m = 0, 1, \dots, M-1, \\ -\lambda_0 v_M &= 0, \end{aligned} \right\}$$

если бы не нарушалось последнее из этих соотношений. Соотношение  $-\lambda_0 v_M = 0$  можно считать граничным условием для решения обыкновенного разностного уравнения

$$\begin{aligned} (1 - r - \lambda_0)u_m + ru_{m+1} &= 0, \\ m &= 0, 1, \dots, M-1. \end{aligned}$$

Чтобы удовлетворить этому граничному условию, которое задано при  $m = M$ , т. е. на правом конце отрезка  $0 \leq x \leq 1$ , «подправим» вектор  $v = (1, q_0, \dots, \dots, q_0^M)$ , помножив каждую из его компонент  $v_m$  на множитель  $(M - m)h$ . Полученный вектор обозначим  $u = (u_0, u_1, \dots, u_M)$ ,  $u_m = (M - m)hq_0^m$ .

На рис. 55 приведены графики функций  $v = \{v_m\}$  и  $u = \{u_m\}$  в случае  $q_0 = -1$ . Норма вектора  $u$  равна единице:

$$\|u\| = \max_m |u_m| = \max_m |(M - m)hq_0^m| = Mh = 1.$$

Оценим норму вектора  $\omega = (\omega_0, \omega_1, \dots, \omega_M)$ , определенного равенством  $\omega \equiv R_h u - \lambda_0 u$ . Для координат вектора  $\omega$  получаем следующие выражения:

$$\begin{aligned} |\omega_m| &= |(1 - r - \lambda_0)(M - m)hq_0^m + r(M - m - 1)hq_0^{m+1}| = \\ &= |[ (1 - r - \lambda_0) + rq_0 ](M - m)hq_0^m - rhq_0^{m+1}| = \\ &= |0 \cdot (M - m)hq_0^m - rhq_0^{m+1}| = rh, \quad m = 0, 1, \dots, M-1, \\ |\omega_M| &= 0 - \lambda_0 \cdot 0 = 0. \end{aligned}$$

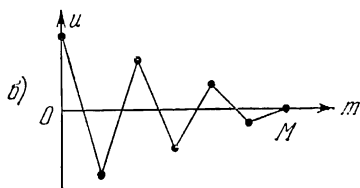
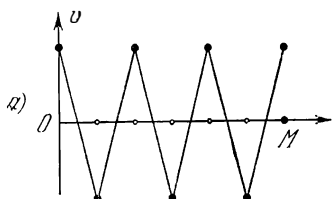


Рис. 55.



Отсюда видно, что  $\|w\| = rh$ , и при  $h < \varepsilon/r$  выполнено неравенство  $\|w\| = \|R_h u - \lambda_0 u\| < \varepsilon \|u\|$ . Этим и завершается доказательство того, что точка  $\lambda_0 \in \overset{\leftrightarrow}{\Lambda}$  принадлежит спектру семейства разностных операторов  $\{R_h\}$ .

Покажем, что если точка  $\lambda_0$  принадлежит одному из множеств  $\overset{\rightarrow}{\Lambda}$  или  $\overset{\leftarrow}{\Lambda}$ , то она является точкой спектра семейства операторов  $\{R_h\}$ . Пусть для определенности  $\lambda_0 \in \overset{\rightarrow}{\Lambda}$ . Тогда по определению множества  $\overset{\rightarrow}{\Lambda}$  уравнение  $\overset{\rightarrow}{R}v - \lambda_0 v = 0$ , которое в развернутом виде записывается равенствами

$$(1 - r - \lambda_0)v_m + rv_{m+1} = 0, \quad m = 0, 1, 2, \dots,$$

имеет решение  $v_m = q_0^m, |q_0| < 1, m = 0, 1, \dots$

Будем рассматривать это решение только при  $m = 0, 1, \dots, M$ , положив

$$u = (u_0, u_1, \dots, u_M) = (1, q_0, \dots, q_0^M).$$

Вычислим для этой сеточной функции  $u$ , график которой в случае  $q = 1/2$  изображен на рис. 56, норму вектора  $w \equiv R_h u - \lambda_0 u$ . Из равенств

$$|w_m| = |(1 - r - \lambda_0)q_0^m + rq_0^{m+1}| = 0, \quad m = 0, 1, \dots, M - 1,$$

$$|w_M| = |q_0|^M$$

следует, что  $\|w\| = |q_0|^M = |q_0|^{1/h}$ . Если  $h$  настолько мало, что  $|q_0|^{1/h} < \varepsilon$ , то  $\|w\| = \|R_h u - \lambda_0 u\| < \varepsilon = \varepsilon \|u\|$ , поскольку  $\|u\| = 1$ .

Итак, доказано, что в нашем примере все точки множеств  $\overset{\leftrightarrow}{\Lambda}, \overset{\rightarrow}{\Lambda}$  и  $\overset{\leftarrow}{\Lambda}$  принадлежат спектру семейства разностных операторов.

Покажем теперь, что всякая точка  $\lambda_0$ , не принадлежащая множествам  $\overset{\leftrightarrow}{\Lambda}, \overset{\rightarrow}{\Lambda}$  и  $\overset{\leftarrow}{\Lambda}$ , не принадлежит спектру семейства  $\{R_h\}$ .

Именно покажем, что существует число  $A > 0$ , не зависящее от  $h$  и такое, что для любой функции  $u = (u_0, u_1, \dots, u_M)$  выполнено неравенство

$$\|R_h u - \lambda_0 u\| \geq A \|u\|. \tag{8}$$

Тогда при  $\varepsilon < A$  неравенство  $\|R_h u - \lambda_0 u\| < \varepsilon \|u\|$  не имеет решения и точка  $\lambda_0$  не принадлежит спектру. Обозначим  $f \equiv \overset{\rightarrow}{R}u - \lambda_0 u$ , тогда неравенство (8) запишется так:

$$\|f\| \geq A \|u\|. \tag{9}$$

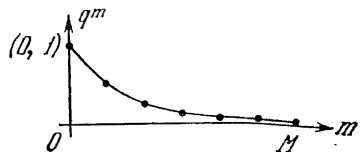


Рис. 56.

Эту оценку мы и будем обосновывать. Равенство  $R_h u - \lambda_0 u = f$  запишем в развернутом виде:

$$\left. \begin{aligned} (1 - r - \lambda_0) u_m + r u_{m+1} &= f_m, \quad m = 0, 1, \dots, M-1, \\ -\lambda_0 u_M &= f_M. \end{aligned} \right\} \quad (10)$$

Будем рассматривать эти соотношения как уравнение относительно  $u$ , а  $f$  будем считать заданной правой частью. Запишем решение  $u = \{u_m\}$  в виде суммы, положив

$$u_m = \alpha_m + \beta_m, \quad m = 0, 1, \dots, M, \quad (11)$$

где  $\alpha_m$  — компоненты ограниченного решения  $\alpha = \{\alpha_m\}$  следующего уравнения:

$$\begin{aligned} (1 - r - \lambda_0) \alpha_m + r \alpha_{m+1} &= F_m = \\ &= \begin{cases} 0, & \text{если } m < 0, \\ f_m, & \text{если } m = 0, 1, \dots, M-1, \\ 0, & \text{если } m \geq M. \end{cases} \end{aligned} \quad (12)$$

Тогда в силу линейности вектор  $\beta = \{\beta_m\}$ , компоненты которого входят в равенство (11), есть решение уравнения

$$\left. \begin{aligned} (1 - r - \lambda_0) \beta_m + r \beta_{m+1} &= 0, \quad m = 0, 1, \dots, M-1, \\ -\lambda_0 \beta_M &= f_M + \lambda_0 \alpha_M. \end{aligned} \right\} \quad (13)$$

Для доказательства оценки (9), которую при сделанном выборе нормы можно переписать в форме  $|u_m| < \frac{1}{A} \max_m |f_m|$ , в силу соотношения  $u_m = \alpha_m + \beta_m$  достаточно установить оценки вида

$$|\alpha_m| \leq A_1 \max |f_m|, \quad (14)$$

$$|\beta_m| \leq A_2 \max |f_m|, \quad (15)$$

где  $A_1$  и  $A_2$  — некоторые постоянные. Начнем с оценки (14). Заметим, что уравнение (12) есть уравнение первого порядка вида

$$a \alpha_m + b \alpha_{m+1} = F_m, \quad m = 0, \pm 1, \dots,$$

где  $a = 1 - r - \lambda_0$ ,  $b = r$ . Уравнение такого вида было рассмотрено в § 2, где получена оценка

$$|\alpha_m| \leq \frac{\max_m |F_m|}{|a| - |b|} = \frac{\max_m |f_m|}{|a| - |b|}. \quad (16)$$

В рассматриваемом примере  $|a| - |b| > \delta_0/2$ , где  $\delta_0$  — расстояние от точки  $\lambda_0$  до множества  $\overleftarrow{\Lambda} + \overrightarrow{\Lambda} + \overleftarrow{\Lambda}$ . Из (16) поэтому

вытекает доказываемое неравенство (14). Оценка (15) вытекает из записи решения уравнения (13) в следующем виде:

$$\beta_m = - \frac{f_M + \lambda_0 \alpha_M}{\lambda_0} q_0^{m-M}, \tag{17}$$

где  $q_0$  определяется соотношением  $(1 - r - \lambda_0) + r q_0 = 0$ . По предположению точка  $\lambda_0$  не принадлежит множеству  $\Lambda$  и поэтому лежит вне круга с центром в точке  $1 - r$  и радиусом  $r$ . А в этом случае  $|q_0| > 1$ . Далее,  $|\lambda_0| = \delta_1 > 0$ , так как если бы было  $\lambda_0 = 0$ , то  $\lambda_0$  принадлежало бы множеству  $\overleftrightarrow{\Lambda} + \overrightarrow{\Lambda} + \overleftarrow{\Lambda}$ . Итак, используя равенство (17) и учитывая уже доказанную оценку (14), получим неравенство (15):

$$|\beta_m| = \left| \frac{f_M + \lambda_0 \alpha_M}{\lambda_0} \right| \cdot |q_0^{m-M}| \leq \frac{|f_M|}{|\lambda_0|} + |\alpha_m| \leq \leq \frac{\max |f_m|}{\delta_1} + A_1 \max |f_m| = A_2 \max |f_m|.$$

Итак, доказано, что спектр семейства операторов  $\{R_h\}$ , определенного формулой (1), совпадает с объединением множеств  $\overleftrightarrow{\Lambda}$ ,  $\overrightarrow{\Lambda}$  и  $\overleftarrow{\Lambda}$  на комплексной плоскости.

**2. Алгоритм вычисления спектра в общем случае.**

*Теорема. Пусть оператор  $R_h, b = R_h a, a, b \in U_h$ , задан равенством  $B_h b = A_h a$ , где  $A_h$  и  $B_h$  — некоторые линейные операторы, определенные на конечномерном линейном нормированном пространстве  $U_h$  со значениями из некоторого линейного нормированного пространства  $F_h$ . Пусть, далее, операторы  $A_h$  и  $B_h$  равномерно по  $h$  ограничены, а оператор  $B_h$  имеет равномерно ограниченный обратный  $B_h^{-1}: \|A_h\|, \|B_h\|, \|B_h^{-1}\| < c$ .*

*В таком случае спектр семейства операторов  $\{R_h\}$  не содержит тех и только тех  $\lambda$  на комплексной плоскости, при которых оператор  $A_h - \lambda B_h$  имеет при всех достаточно малых  $h$  равномерно по  $h$  ограниченный обратный оператор.*

Доказательство очевидно, и мы его излагать не будем.

Пусть теперь оператор  $R_h, v = R_h u$ , задан разностными соотношениями

$$\left. \begin{aligned} \sum_{k=-k_0}^{k_0} B_{km} v_{m+k} &= \sum_{k=-k_0}^{k_0} A_{km} u_{k+m}, \quad k_0 \leq m \leq M - k_0, \\ \sum_{i=0}^{2k_0} b_i v_i &= \sum_{i=0}^{2k_0} a_i u_i, \quad \sum_{i=0}^{2k_0} \beta_i v_{N-i} = \sum_{i=0}^{2k_0} \alpha_i u_{N-i}, \end{aligned} \right\} \tag{18}$$

причем задача

$$\left. \begin{aligned} \sum_{k=-k_0}^{k_0} B_{km} v_{m+k} &= \varphi_m, \quad k_0 \leq m \leq M - k_0, \\ \sum_{i=0}^{2k_0} b_i v_i &= \psi_1, \quad \sum_{i=0}^{2k_0} \beta_i v_{M-i} = \psi_2 \end{aligned} \right\} \tag{19}$$

хорошо обусловлена.

Предполагается, что

$$A_{km} = A_k \left( \frac{m}{M} \right), \quad B_{km} = B_k \left( \frac{m}{B} \right),$$

где  $A_k(x)$  и  $B_k(x)$  — квадратные матрицы, определенные на отрезке  $0 \leq x \leq 1$ , удовлетворяющие на этом отрезке условиям гладкости (14) § 4;  $a_i, b_i, \alpha_i, \beta_i$  — прямоугольные числовые матрицы, не зависящие от  $M$ . В таком случае применима теорема и спектр семейства операторов  $\{R_h\}$  состоит из всех тех  $\lambda$ , для которых разностная краевая задача

$$\left. \begin{aligned} \sum_{k=-k_0}^{k_n} (\lambda B_{km} - A_{km}) u_{m+k} &= \varphi_m, \quad k_0 \leq m \leq M - k_0, \\ \sum_{i=0}^{2k_0} (\lambda b_i - a_i) u_i &= \psi_1, \quad \sum_{i=0}^{2k_n} (\lambda \beta_i - \alpha_i) u_{M-i} = \psi_2 \end{aligned} \right\} \quad (20)$$

не является хорошо обусловленной. Для выяснения того, является ли задача (20) хорошо обусловленной, при каждом  $\lambda$  можно воспользоваться критерием п. 7 § 4.

### ЗАДАЧИ

1. Доказать, что для семейства разностных операторов  $\{R_h\}$ ,  $v = R_h u$ , заданного равенствами

$$\left. \begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, \quad m = 0, 1, \dots, M-1, \\ v_M &= 0, \quad Mh = 1, \end{aligned} \right\}$$

и рассмотренного в этом параграфе, спектр не изменится, если норму определить не по формуле  $\|u\| = \max_m |u_m|$ , а по формуле  $\|u\| = \left( h \sum_m |u_m|^2 \right)^{1/2}$ .

2. Доказать, что спектр семейства разностных операторов  $\{R_h\}$ ,  $v = R_h u$ , заданного равенствами

$$\left. \begin{aligned} v_m &= (1-r+\gamma h)u_m + ru_{m+1}, \quad m = 0, 1, \dots, M-1, \\ v_M &= 0, \quad Mh = 1, \end{aligned} \right\}$$

не зависит от значения числа  $\gamma$  и совпадает со спектром, построенным в этом параграфе для случая  $\gamma = 0$ .

3. Вычислить спектр семейства операторов  $\{R_h\}$ ,  $v = R_h u$ , заданного равенствами

$$\left. \begin{aligned} v_m &= (1-r)u_m + r(u_{m-1} + u_{m+1}), \quad m = 1, 2, \dots, M-1, \\ au_0 + bu_1 &= 0, \quad u_M = 0, \quad Mh = 1, \quad r = \text{const}, \end{aligned} \right\}$$

где  $a$  и  $b$  — заданные числа. Рассмотреть случаи  $|a| > |b|$  и  $|a| < |b|$ .

### § 46. Ядра спектров семейств операторов

Пусть  $R_h$  отображает линейное нормированное пространство  $U'_h$  некоторой размерности  $N$ ,  $N = N(h)$ , в себя. Будем писать вместо  $R_h$  и  $U'_h$  соответственно  $R_N$  и  $U_N$ , чтобы в обозначениях была явно указана размерность. Предполагается, что  $N \rightarrow \infty$  при  $h \rightarrow 0$ ,

Здесь мы обсудим вопрос о том, насколько спектр семейства операторов  $\{R_N\}$  зависит от выбора последовательности норм  $\|\cdot\|_N$  в пространствах  $U_N$  и, тем самым, насколько инвариантен спектральный признак ограниченности норм степеней оператора  $R_N$  (теорема 1 из § 44) относительно выбора норм.

Относительно семейства операторов  $\{R_N\}$  будем предполагать, что собственные числа всех операторов  $R_N$  ограничены в совокупности, т. е. лежат в некотором круге

$$|\lambda| \leq c = \text{const.} \quad (1)$$

Очевидно, что для выполнения условия (1) достаточно, чтобы существовала хотя бы одна последовательность норм  $\|\cdot\|_N$  такая, чтобы выполнялись неравенства  $\|R_N\|_N < c' = \text{const.}$  Отсюда видно, что ограничение (1) естественно: оно выполняется для семейств операторов  $\{R\}_N$  перехода со слоя на слой, возникающих при рассмотрении эволюционных разностных краевых задач. Переходим к определению понятия ядра спектра, с помощью которого и будут сформулированы результаты этого параграфа.

Пусть заданы: какая-либо последовательность норм  $\|\cdot\|_N$ , число  $a \in [0, 1]$  и целое  $k \geq 0$ . Обозначим через  $\Lambda(a, k, N)$  множество точек  $\lambda$ , для которых неравенство  $\|R_N u - \lambda u\| < a^N N^{-k} \|u\|$  имеет решение  $u \in U_N$ . Значок  $N$  при написании норм мы опускаем.

**О п р е д е л е н и е.** Ядром показателя  $a \in [0, 1]$  семейства операторов  $\{R_N\}$  назовем следующее множество  $\Lambda(a)$  на комплексной плоскости:

$$\Lambda(a) \equiv \bigcap_{k \geq 0} \bigcap_{s > 0} \left( \overline{\bigcup_{N > s} \Lambda(a, k, N)} \right).$$

Здесь

$$\overline{\bigcup_{N > s} \Lambda(a, k, N)} \equiv \Lambda_s(a, k)$$

есть теоретико-множественное замыкание объединения множеств  $\Lambda(a, k, N)$  при всех  $N > s$ ;  $\bigcap_{s > 0} \Lambda_s(a, k) \equiv \Lambda(a, k)$  — пересечение всех множеств

$\Lambda_s(a, k)$ ;  $\bigcap_{k \geq 0} \Lambda(a, k) = \Lambda(a)$  — пересечение всех множеств  $\Lambda(a, k)$ .

**Т е о р е м а 1.** Ядро  $\Lambda(a)$ ,  $a \in [0, 1]$ , целиком содержится в спектре семейства операторов  $\{R_N\}$  и замкнуто.

**Доказательство.** Покажем, что если точка  $\lambda_0$  не принадлежит спектру семейства операторов  $\{R_N\}$ , то она не принадлежит и ядру. Действительно, найдутся  $\varepsilon > 0$  и  $N_0$  такие, что при всех  $N > N_0$  для любого  $u \in U_N$  выполнено неравенство  $\|R_N u - \lambda_0 u\| \geq \varepsilon \|u\|$ . Но тогда для всех  $\lambda$  из круга  $|\lambda - \lambda_0| < \varepsilon/2$  выполнено также неравенство  $\|R_N u - \lambda u\| \geq \frac{\varepsilon}{2} \|u\|$ . Ввиду этого при  $N > N_0$  ни одно множество  $\Lambda(a, k, N)$  не содержит точек в круге  $|\lambda - \lambda_0| < \varepsilon/2$ . Отсюда следует справедливость первого утверждения теоремы. Для доказательства замкнутости ядра  $\Lambda(a)$  заметим, что  $\Lambda_s(a, k)$  замкнуты по построению, а множества  $\Lambda(a, k)$  и  $\Lambda(a)$  — как пересечения замкнутых множеств.

**П р и м е р.** Вычислим ядро  $\Lambda(a)$ ,  $a \in [0, 1]$ , для семейства операторов  $\{R_N\}$ , если оператор  $R_{N+1}$ ,  $v = R_{N+1} u$ , задан равенствами

$$\left. \begin{aligned} v_n &= (1-r)u_n + ru_{n+1}, & n &= 0, 1, \dots, N-1, \\ v_N &= 0, \end{aligned} \right\} \quad (2)$$

а норма — равенством  $\|u\| = \|(u_0, u_1, \dots, u_N)\| = \max |u_n|$ .

Покажем, что  $\Lambda(a)$  состоит из точки  $\lambda = 0$  и из замкнутого круга радиуса  $ar$  с центром в точке  $1 - r$ :

$$|\lambda - (1 - r)| \leq ar. \quad (3)$$

Действительно,  $\lambda = 0$ , как мы видели в п. 1 § 42, является собственным числом для всех операторов  $R_N$ , а потому принадлежит всем множествам  $\Lambda(a, k, N)$  и, следовательно, ядру. Далее, для любого  $\lambda_0$ , лежащего строго внутри круга (3), при некоторых вещественном  $\alpha \geq 0$  и  $b > 1$  имеет место представление

$$\lambda_0 = 1 - r + \frac{ar}{b} e^{i\alpha}.$$

Неравенство  $\|R_N u - \lambda_0 u\| < a^N N^{-k} \|u\|$  при любом фиксированном  $k$  и всех достаточно больших  $N$  имеет решение

$$u_n = \begin{cases} (a/b)^n e^{i\alpha n}, & n = 0, 1, \dots, N-1, \\ 0, & n = N. \end{cases}$$

Следовательно, при всех достаточно больших  $N$  множества  $\Lambda(a, k, N)$  содержат точку  $\lambda_0$ , а значит, ее содержит и  $\Lambda(a)$ . Итак, внутренние точки круга (3) принадлежат ядру  $\Lambda(a)$ , а ввиду замкнутости ядра ему принадлежит и граница круга (3).

Если точка  $\lambda_0 \neq 0$  не принадлежит кругу (3), т. е.

$$\lambda_0 = 1 - r + \frac{ar}{b} e^{i\alpha}, \quad \alpha \geq 0, \quad b = 1 - 2\delta, \quad \delta > 0,$$

то, выписав функцию Грина разностного уравнения первого порядка (§ 2), можно установить, что при любом  $\lambda$  из круга  $|\lambda - \lambda_0| < \min \left[ |\lambda_0|, \frac{ar}{1-\delta} \right]$  при всех достаточно больших  $N$  и всех  $u \in U_N$  выполнено неравенство  $\|R_N u - \lambda u\| > a^N \|u\|$ . Отсюда следует, что точки этого круга не принадлежат  $\Lambda(a, k, N)$ , если  $N$  достаточно велико, а следовательно, и не принадлежат ни замыканию их объединений  $\Lambda_s(a, k)$ , ни ядру  $\Lambda(a)$ .

Заметим, что ядро  $\Lambda(0)$  показателя  $a = 0$  в рассмотренном примере состоит из двух точек  $\lambda = 0$  и  $\lambda = 1 - r$ , а ядро  $\Lambda(1)$  совпадает со всем спектром семейства операторов  $\{R_N\}$ , который был вычислен в § 45.

На этом закончим рассмотрение примера и вернемся к общим построениям.

**Определение.** Ядро  $\Lambda(0)$  назовем *абсолютным ядром*.

**Теорема 2.** *Абсолютное ядро семейства операторов  $\{R_N\}$  не зависит от выбора последовательности норм  $\|\cdot\|_N$ .*

Доказательство следует из того факта, что при  $a = 0$  множество  $\Lambda(a, k, N)$  совпадает при каждом  $N$  с множеством собственных значений оператора  $R_N$ , которое не зависит от нормы в пространстве  $U_N$ .

**Теорема 3.** *При условии (1) последовательность норм  $\|\cdot\|_N$  всегда можно выбрать так, чтобы спектр семейства операторов  $\{R_N\}$  совпал со своим абсолютным ядром.*

Доказательство. Укажем конструкцию норм, существование которых утверждается в теореме. Выберем базис в пространстве  $U_N$  так, чтобы матрица преобразования  $R_N$  в этом базисе была жордановой и модули всех внедиагональных членов были меньше чем  $1/N$ . Введем скалярное умножение и порожденную им норму, объявив этот базис ортонормальным. Если  $\lambda_0$  — произвольная точка, не принадлежащая  $\Lambda(0)$ , и  $\varepsilon > 0$  — расстояние от этой

точки до замкнутого в силу теоремы 1 множества  $\Lambda(0)$ , то можно проверить, что  $\|Ru - \lambda_0 u\| \geq \frac{\varepsilon}{4} \|u\|$  при всех  $N > 8/\varepsilon$  и всех  $u \in U_N$ , так что  $\lambda_0$  не принадлежит спектру семейства операторов  $\{R_N\}$ .

Итак, если спектр семейства операторов  $\{R_N\}$  не совпадает со своим ядром  $\Lambda(0)$  показателя  $a = 0$  при заданном выборе норм, как это имеет место в рассмотренном выше примере (2) при норме  $\|u\| = \max |u_n|$ , то за счет выбора другой последовательности норм можно получить в качестве спектра более узкое множество  $\Lambda(0)$ .

Однако в теории разностных схем используются не вполне произвольные нормы.

Обозначим через  $\|\cdot\|_{cN}$  норму, равную максимуму абсолютных величин всех компонент, образующих сеточную функцию (или вектор-функцию) из  $U_N$ . Выделим класс последовательностей норм  $\|\cdot\|_N$ , для которых существует натуральное  $s$ , зависящее от последовательности, и такое, что при всех достаточно больших  $N$  справедливо неравенство

$$\sup_{\|u\|_c=1} \|u\|_N \leq N^s \inf_{\|u\|_c=1} \|u\|_N. \quad (4)$$

Очевидно, что сама норма  $\|\cdot\|_{cN}$  и все встречавшиеся нам в связи с разностными уравнениями нормы при возрастающем  $N$  образуют последовательности из указанного класса (4).

**Теорема 4.** *Ядро  $\Lambda(a)$  показателя  $a \in [0, 1]$  не зависит от выбора последовательности норм из числа удовлетворяющих требованию (4).*

Доказательство непосредственно следует из определений.

Рассмотрим теперь семейство операторов  $\{R_h\}$ , определенных равенствами (18) и (19) из § 45, предположив дополнительно, что матричные коэффициенты  $A_k$  и  $B_k$  постоянны:  $A_k(x) \equiv A_k(0)$ ,  $B_k(x) \equiv B_k(0)$ . Для этого семейства операторов справедлива следующая важная

**Теорема 5** (А. В. Соколов). *Если в пространствах  $U'_h = U_N$ , где действуют операторы  $R_h = R_N$ , введены нормы  $\|\cdot\|_{cN}$ , то ядро  $\Lambda(1)$  показателя  $a = 1$  спектра семейства операторов  $\{R_N\}$  совпадает со всем спектром этого семейства операторов.*

Из этой теоремы и теоремы 4 следует, что при любом выборе последовательности норм из класса норм, удовлетворяющих условию (4), спектр семейства операторов  $\{R_N\}$  содержит спектр семейства операторов  $\{R_N\}$ , полученного при использовании норм  $\|\cdot\|_{cN}$ , способ вычисления которого описан в п. 2 § 45. Поэтому если не выполнено спектральное условие ограниченности норм степеней операторов  $R_N$  (теорема 1 из § 44) при выборе норм  $\|\cdot\|_{cN}$ , то оно не выполнено и при любом другом выборе последовательности норм из числа удовлетворяющих условию (4).

Доказательство теоремы А. В. Соколова требует сложного исследования, и мы его не излагаем.

## § 47. Об устойчивости итерационных алгоритмов решения несамосопряженных разностных уравнений

Решение стационарных задач установлением можно понимать как некоторый итерационный процесс, а результаты, полученные на очередном временном слое, — как соответствующее приближение. В § 35 была рассмотрена в качестве примера разностная задача Дирихле для уравнения Пуассона.

В случае нулевых условий на границе — это самосопряженная разностная задача. В соответствии с этим в процессе установления можно было разлагать ошибку по полной ортогональной системе собственных функций. По расположению собственных чисел можно было судить одновременно и о скорости убывания погрешности, и о влиянии ошибок округления, вносимых на промежуточных слоях.

Оказывается, при решении несамосопряженных разностных уравнений установлением дело обстоит, вообще говоря, не так. Может возникнуть, несмотря на сходимость итерационного процесса, неустойчивость из-за большой чувствительности к ошибкам округления.

Это явление мы здесь точно определим и обсудим. При этом окажется полезным понятие спектра и ядра спектра семейства разностных операторов.

Пусть

$$u = R_N u + f_N \quad (1)$$

— семейство линейных уравнений («разностное уравнение») относительно неизвестного элемента  $u$  из некоторого  $N$ -мерного линейного нормированного пространства  $U_N$ , зависящее от натурального параметра  $N$ . Мы будем рассматривать итерационный процесс

$$u^{m+1} = R_N u^m + f_N, \quad m = 0, 1, \dots \quad (2)$$

вычисления решения  $u$ . Будем предполагать, что все собственные значения  $\lambda_k = \lambda_k(N)$  оператора  $R_N$  по модулю меньше единицы

$$|\lambda_k| \leq \rho_N < 1, \quad (3)$$

т. е. что известный критерий сходимости процесса (2) выполнен, причем

$$\|u - u^m\| = O(\rho_N^m). \quad (4)$$

Пусть теперь вычисления (2) ведутся приближенно с некоторым числом  $p = q + \alpha$  десятичных знаков, т. е. по формуле

$$\tilde{u}^{m+1} = R_N \tilde{u}^m + f_N + 10^{-p} \|\tilde{u}^m\| \delta_m, \quad (5)$$

где  $\delta_m \in U_N$ ,  $\|\delta_m\| \leq 1$  произвольны.

Зададим натуральное  $q$  и будем требовать, чтобы при произвольных  $\delta_k$ ,  $\|\delta_k\| \leq 1$ ,  $k = 0, 1, \dots$ , выполнялось неравенство

$$\overline{\lim}_{m \rightarrow \infty} \|u - \tilde{u}^m\| \leq 10^{-q} \|u\|. \quad (6)$$

Неравенство (6) обеспечивает возможность вычислить решение  $u$  по формулам (5) с ошибкой, не превосходящей единицы  $q$ -го десятичного знака (в смысле нормы в  $U_N$ ).

*Лемма.* Для выполнения требования (6) необходимо, чтобы число  $\alpha$  «запасных знаков» в формуле (5) удовлетворяло неравенству

$$(1 - 10^{-q}) \varphi \leq 10^\alpha,$$

и достаточно, чтобы  $\alpha$  удовлетворяло неравенству

$$(1 + 10^{-q}) \varphi \leq 10^\alpha,$$

где

$$\varphi = \lim_{m \rightarrow \infty} \max_{\|\delta_k\|=1} \left\| \sum_{k=0}^m R_N^{m-k} \delta_k \right\|.$$

Доказательство предоставляем читателю.



Отметим, что существование  $\varphi = \varphi(N)$  следует из условия (3). В дальнейшем будем понимать под  $\alpha = \alpha(N)$  наименьшее целое, обеспечивающее выполнение требования (6). Из леммы видно, что такое число существует, отрицательно, от  $q$  зависит несущественно либо не зависит вовсе.

О п р е д е л е н и е. Сходящийся итерационный алгоритм (2) будем называть *устойчивым*, если существует не зависящая от  $N$  постоянная  $C$ , при которой выполнено неравенство

$$\alpha(N) < C; \quad (7)$$

сходящийся итерационный алгоритм будем называть *слабо устойчивым*, если существует не зависящая от  $N$  постоянная  $C$ , при которой выполнено неравенство

$$\alpha(N) \leq C \ln N, \quad (8)$$

но устойчивость места не имеет. Наконец, сходящийся итерационный алгоритм будем называть *неустойчивым*, если он не является ни устойчивым, ни слабо устойчивым.

П р и м е р. Запишем уравнение

$$\left. \begin{aligned} -2u_n + u_{n+1} - f_n &= 0, & n = 0, 1, \dots, N-1, \\ u_N &= 0 \end{aligned} \right\} \quad (9)$$

в форме

$$\left. \begin{aligned} u_n &= (1-2r)u_n + ru_{n+1} + rf_n, & n = 0, 1, \dots, N-1, \\ u_N &= 0, \end{aligned} \right\} \quad (10)$$

где  $r$  — параметр. Итерационный алгоритм (2) для уравнения (10) примет вид

$$u_n^{m+1} = (1-2r)u_n^m + ru_{n+1}^m + rf_n^m, \quad n = 0, 1, \dots, N-1, \quad (11)$$

так что оператор  $R_N$ ,  $v = R_N u$  запишется формулами

$$\begin{aligned} v_n &= (1-2r)u_n + ru_{n+1}, & n = 0, 1, \dots, N-1, \\ v_N &= 0. \end{aligned}$$

Оператор  $R_N$  имеет, как легко видеть, только два собственных значения  $\lambda_1(N) = 1-2r$  и  $\lambda_2(N) = 0$ .

Неравенство (3) выполнено и итерационный алгоритм (11) сходится при  $r < 1$ . Примем за норму  $\|u\| = \max_n |u_n|$ . Покажем, что при  $r < \frac{2}{3}$  он устойчив, а при  $r > \frac{2}{3}$  неустойчив. Действительно, если  $r < \frac{2}{3}$ , то

$$\max_n |v_n| \leq \max(|1-3r|, |1-r|) \max_n |u_n|,$$

так что  $\|R_N\| \leq \max(|1-3r|, |1-r|) = \rho < 1$ . Поэтому  $\varphi(N) \leq \frac{1}{1-\rho}$ ,

и в силу леммы имеет место оценка (7) с постоянной  $C = -2 \ln(1-\rho)$ .

Пусть теперь  $r > \frac{2}{3}$ . Положим в (11)  $f_n = 0$ ,  $u_n^0 = (-1)^n$ . Легко видеть, что

в таком случае  $u_n^m = (1-3r)^m (-1)^n$ ,  $n = 0, 1, \dots, N-m$ . Отсюда следует, что

$\|R_N^m\| \geq \rho^m$ ,  $m = 1, 2, \dots, N-1$ , где  $\rho = |1-3r| > 1$ . Поэтому  $\varphi(N) > \rho^N$ ,

а в силу леммы  $\alpha \approx N \lg \rho$ , что доказывает неустойчивость. Можно показать, что при  $r = \frac{2}{3}$  итерационный алгоритм (11) слабо устойчив.

Таким образом, спектральный критерий сходимости (3) итерационного алгоритма не определяет его устойчивости. Спектральный критерий и признаки устойчивости формулируются не в терминах расположения спектров каждого из операторов  $R$ , а в терминах расположения спектра и ядер спектра семейства операторов  $\{R_N\}$ . Именно, в предположении, что семейство операторов  $\{R_N\}$  равномерно ограничено,  $\|R_N\| < C$ , легко проверить следующие утверждения.

*Лемма. Для того чтобы при всех достаточно больших значениях  $N$  итерационный алгоритм (2) был сходящимся, достаточно, чтобы радиус  $\rho$  какого-нибудь ядра спектра семейства операторов  $\{R_N\}$  был строго меньше единицы.*

*Критерий устойчивости. Для устойчивости итерационного алгоритма (2) необходимо и достаточно, чтобы спектр семейства операторов  $\{R_N\}$  лежал строго внутри единичного круга.*

*Теорема. Для того чтобы итерационный алгоритм (2) сходил к и был устойчив или слабо устойчив, достаточно, чтобы радиус  $\rho$  ядра  $\Lambda(1)$  спектра семейства операторов  $\{R_N\}$  был строго меньше единицы; для неустойчивости сходящегося итерационного алгоритма (2) достаточно, чтобы радиус  $\rho$  этого ядра спектра семейства операторов  $\{R_N\}$  был строго больше единицы.*

В § 46 показано, что ядро  $\Lambda(1)$  спектра семейства операторов  $\{R_N\}$  не зависит от выбора норм из естественного для разностных уравнений класса (4) § 46. Отсюда следует, в частности, что если операторы  $R_N$  являются равномерно по  $N$  сжимающими,  $\|R_N\| \leq \rho < 1$ , так что спектр, а значит, и ядро  $\Lambda(1)$  спектра семейства операторов  $\{R_N\}$  лежат в круге  $|\lambda| \leq \rho < 1$ , то итерационный алгоритм (2) устойчив и остается устойчивым (сильно или слабо) и в любой другой норме (4) § 46, в которой операторы  $R_N$  могут перестать быть сжимающими.

В рассмотренном выше примере спектр семейства операторов  $\{R_N\}$  состоит из круга  $|\lambda - (1 - 2r)| \leq r$  и точки  $\lambda = 0$ , причем совпадает со своим ядром  $\Lambda(1)$ . Утверждение об устойчивости алгоритма (11) при  $r < \frac{2}{3}$  и неустойчивости при  $r > \frac{2}{3}$  можно сделать поэтому и с помощью спектральных признаков.

Для вычисления решения (несамоспряженного) уравнения вида

$$A_N u + f_N = 0 \quad (12)$$

можно пытаться строить итерационный алгоритм в форме

$$B_N u^{m+1} = B_N u^m + (A_N u^m + f_N). \quad (13)$$

При этом оператор  $B_N$  надо подобрать так, чтобы его было легко численно обратить и чтобы спектр семейства операторов  $\{B_N^{-1} A_N\}$  имел возможно меньший радиус  $\rho$ ,  $\rho < 1$ . В силу оценки  $\|R_N^m\| \leq C(\epsilon) \cdot (\rho + \epsilon)^m$ , где  $\epsilon > 0$  произвольно, а  $C(\epsilon)$  не зависит от  $N$ , это обеспечит быструю сходимость, а в силу критерия устойчивости, сформулированного выше, — устойчивость итерационного алгоритма (13).

## ДОПОЛНЕНИЕ

### МЕТОД ВНУТРЕННИХ ГРАНИЧНЫХ УСЛОВИЙ

В теории краевых задач для аналитических функций, т. е. для решений системы уравнений Коши — Римана, а также для решений более общих систем уравнений с частными производными, применяется метод сингулярных интегральных уравнений. Он состоит в сведении краевых задач к некоторым интегральным уравнениям на границе рассматриваемой области. При этом в дополнение к заданным граничным условиям используются следствия самой системы дифференциальных уравнений — соотношения, которым должны удовлетворять функции (и их нормальные производные) на границе области, чтобы их можно было доопределить внутри области до некоторого решения соответствующей системы. В случае аналитических функций — это классическое условие Сохоцкого — Племеля, которое возникает при переходе в интегральной формуле Коши

$$\varphi(z) = \frac{\Gamma}{2\pi i} \int_{\gamma} \frac{f(\xi)}{\xi - z} d\xi$$

к пределу при стремлении  $z$  к границе  $\gamma$ . В случае дифференциальных уравнений второго порядка соответствующее условие возникает из формулы Грина, выражающей решение в каждой точке области через значения этого решения и его нормальной производной на границе. Чтобы получить это условие, также надо перейти к пределу при стремлении точки изнутри области к ее границе, воспользовавшись свойствами потенциалов простого и двойного слоев.

Метод внутренних граничных условий по идее аналогичен описанному методу редукции краевых задач для уравнений с частными производными к интегральным уравнениям на границе. Роль дополнительных граничных условий, аналогичных условию Сохоцкого — Племеля, играют внутренние граничные условия, возникающие из разностного аналога интегральной формулы Коши (или разностного аналога формулы Грина).

**1. Класс систем разностных уравнений.** Рассматриваются краевые задачи для общих систем разностных уравнений с постоянными коэффициентами, которые в векторной записи имеют вид

$$Lu \equiv \sum_{k \in K} A_k u_{n+k} - f_n, \quad (1)$$

где  $n = (n_1, n_2, \dots, n_s)$ ,  $k = (k_1, k_2, \dots, k_s)$  — мультииндексы,  $A_k$  — квадратные матрицы,  $f_n$  — заданная и  $u_n$  — искомая вектор-функции,  $K$  — конечное множество (шаблои). Будем предполагать, что система (1) удовлетворяет следующему алгебраическому условию: характеристическая матрица

$$A(\xi) \equiv \sum_{k \in K} A_k \xi^k, \quad (2)$$

где  $\xi^k \equiv \xi_1^{k_1} \dots \xi_s^{k_s}$  и  $\xi_1, \dots, \xi_s$  — комплексные параметры, не является тождественно по  $\xi$  вырожденной:

$$\det A(\xi) \neq 0. \quad (3)$$

Это ограничение естественно: можно показать, что в случае  $\det A(\xi) \equiv 0$  уравнение (1) имеет решение не при всякой финитной (по  $n$ ) правой части  $f_n$ .

**2. Фундаментальное решение.** Матричную функцию  $G_n$  назовем *фундаментальным решением системы* (1), если она одновременно удовлетворяет следующим двум уравнениям:

$$\sum_{k \in K} A_k G_{n-k} = \delta_n^0 E, \quad (4)$$

$$\sum_{k \in K} G_{n-k} A_k = \delta_n^0 E. \quad (4')$$

*Лемма.* Пусть  $Q(\xi_1, \dots, \xi_t)$  есть произвольный многочлен от произвольного числа  $t$  комплексных аргументов, не обращающийся тождественно в нуль. Тогда можно выбрать радиусы  $r_j$  окружностей  $|\xi_j| = r_j$  так, чтобы выполнялось неравенство  $Q(\xi_1, \dots, \xi_t) \neq 0$ , если  $|\xi_1| = r_1, \dots, |\xi_t| = r_t$ .

Доказательство проведем индукцией по числу аргументов  $t$ . При  $t = 1$  число корней  $Q(\xi_1) = 0$  конечно и утверждение очевидно. Считая, что утверждение доказано для  $t = p$ , установим его в случае  $t = p + 1$ . Многочлен  $Q(\xi_1, \dots, \xi_{p+1})$  расположим по степеням  $\xi_{p+1}$ :

$$Q(\xi_1, \dots, \xi_{p+1}) = Q_0(\xi_1, \dots, \xi_p) \xi_{p+1}^M + \dots + Q_M(\xi_1, \dots, \xi_p),$$

где  $M$  — некоторое натуральное число и  $Q_0(\xi_1, \dots, \xi_p)$  не обращается тождественно в нуль. Выберем  $r_1, \dots, r_p$  так, чтобы  $Q_0(\xi_1, \dots, \xi_p) \neq 0$  при  $|\xi_1| = r_1, \dots, |\xi_p| = r_p$ . Это возможно по предположению индукции. Выбирая теперь  $r_{p+1}$  достаточно большим, можно добиться, чтобы при  $|\xi_j| = r_j$ ,  $j = 1, \dots, p + 1$ , выполнялось неравенство  $Q(\xi_1, \dots, \xi_{p+1}) \neq 0$ .

**Теорема 1.** Матрица  $G_n$ , определяемая равенством

$$G_n = \frac{1}{(2\pi i)^s} \oint_{|\xi_1|=r_1} \dots \oint_{|\xi_p|=r_p} \dots \oint_{\xi_1^{n_1+1} \dots \xi_s^{n_s+1}} \frac{A^{-1}(\xi)}{\xi_1^{n_1+1} \dots \xi_s^{n_s+1}} d\xi_1 \dots d\xi_s, \quad (5)$$

является фундаментальным решением.

Здесь  $r_j$  в соответствии с леммой выбраны так, чтобы  $\det A(\xi) \neq 0$ , если  $|\xi_j| = r_j$ .

Доказательство получается непосредственной проверкой. Учитывая свойства вычетов, получаем

$$\sum_{k \in K} A_k G_{n-k} = \frac{1}{(2\pi i)^s} \int \dots \int \frac{A(\xi) A^{-1}(\xi)}{\xi_1^{n_1+1} \dots \xi_s^{n_s+1}} d\xi_1 \dots d\xi_s = \delta_n^0 E,$$

$$\sum_{k \in K} G_{n-k} A_k = \frac{1}{(2\pi i)^s} \int \dots \int \frac{A^{-1}(\xi) A(\xi)}{\xi_1^{n_1+1} \dots \xi_s^{n_s+1}} d\xi_1 \dots d\xi_s = \delta_n^0 E.$$

**3. Граница сеточной области.** Рассмотрим уравнение (1) на некотором ограниченном множестве

$$Lu \equiv \sum_{k \in K} A_k u_{n+k} = f_n, \quad n \in D_0, \quad (6)$$

где  $D_0$  — произвольная сеточная область определения правой части  $f_n$ . Тогда область определения решения  $u_n$  есть множество  $D$ , которое пробегает точка

$n+k$ , если  $n$  и  $k$  пробегает независимо  $D_0$  и  $K$  соответственно. Сопоставим каждому  $r \in D$  подмножество  $K_r$  множества  $K$ , состоящее из всех тех  $k \in K$ , для которых  $r-k \in D_0$ . Границей  $\Gamma$  назовем совокупность всех тех точек  $r \in D$ , для которых  $K_r$  непусто. Например, для простейшего разностного аналога уравнения Пуассона

$$\begin{aligned} Lu \equiv u_{n_1-1, n_2} + u_{n_1, n_2+1} + u_{n_1+1, n_2} + \\ + u_{n_1, n_2-1} - 4u_{n_1 n_2} = h^2 F_{n_1 n_2}, \\ |n_1| < N, \quad |n_2| < N; \quad Nh = 1, \end{aligned}$$

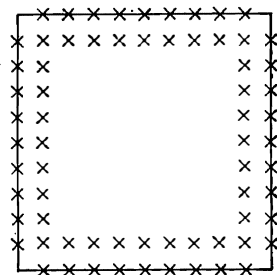


Рис. 57.

множество  $D_0$  состоит из тех точек  $(n_1 h, n_2 h)$ , которые попали внутрь квадрата  $|x_1| \leq 1$ ,  $|x_2| \leq 1$ . Множество  $K$  — из пяти векторов  $(1, 0)$ ,  $(0, 1)$ ,  $(-1, 0)$ ,  $(0, -1)$ ,  $(0, 0)$ . Множество  $D$  — совокупность всех целочисленных точек квадрата  $|n_1| \leq N$ ,  $|n_2| \leq N$ , кроме четырех угловых  $|n_1| = |n_2| = N$ . Граница  $\Gamma$  состоит из двух слоев точек, отмеченных на рис. 57 крестиками.

#### 4. Разностные аналоги интегральных формул Коши и типа Коши.

Л е м м а. Пусть  $B_n$  — произвольная матрица-функция, для которой имеет смысл умножение справа на квадратную матрицу порядка  $t$ , определенная на всей целочисленной сетке. Справедливо следующее тождество:

$$\begin{aligned} \sum_{n \in D_0} B_{-n} \sum_{k \in K} A_k u_{n+k} \equiv \\ \equiv \sum_{n \in D} \left( \sum_{k \in K} B_{-n+k} A_k \right) u_n - \sum_{r \in \Gamma} \left( \sum_{k \in K_r} B_{-r+k} A_k \right) u_r. \quad (7) \end{aligned}$$

Доказательство. Вектор-функцию  $u_n$ ,  $n \in D$ , можно записать в виде

$$u_n = \sum_{t \in D} \delta_n^t u_t.$$

Левая и правая части тождества (7) линейно зависят от  $u$ . Поэтому для доказательства достаточно проверить справедливость тождества (7) для вектор-функции

$$v_n \equiv v_n^t \equiv \delta_n^t u_t = \begin{cases} 0, & \text{если } n \neq t, \\ u_t, & \text{если } n = t, \end{cases}$$

при каждом фиксированном  $t \in D$ :

$$\begin{aligned} \sum_{n \in D_0} B_{-n} \sum_{k \in K} A_k v_{n+k} &\equiv \sum_{k \in K} \sum_{n \in D_0} B_{-n} A_k v_{n+k} \equiv \\ &\equiv \sum_{k \in K} \delta_{D_0}^{t-k} B_{-t+k} A_k v_t \equiv \sum_{k \in K} B_{-t+k} A_k v_t - \sum_{k \in K_t} B_{-t+k} A_k v_t \equiv \\ &\equiv \sum_{n \in D} \left( \sum_{k \in K} B_{-n+k} A_k \right) v_n - \sum_{r \in \Gamma} \left( \sum_{k \in K_r} B_{-r+k} A_k \right) v_r, \end{aligned}$$

где  $\delta_D^{t-k} = \begin{cases} 1, & \text{если } t-k \in D_0, \\ 0, & \text{если } t-k \notin D_0. \end{cases}$

**Теорема 2.** Пусть  $\{u_n\}$ ,  $n \in D$ , — произвольное решение уравнения (6), а  $G_n$  — произвольное фундаментальное решение. Тогда справедлива формула

$$\sum_{r \in \Gamma} \left( \sum_{k \in K_r} G_{n-r+k} A_k \right) u_r + \sum_{m \in D_0} G_{n-m} f_m = \begin{cases} u_n, & \text{если } n \in D, \\ 0, & \text{если } n \in \bar{D}. \end{cases} \quad (8)$$

**Доказательство.** Помножим обе части равенства (6) слева на матрицу  $G_{t-n}$  и просуммируем по всем  $n \in D_0$ . Воспользовавшись тождеством (7), а затем равенством (4'), получим формулу (8).

**Следствие.** Каждое решение  $\{u_n\}$  уравнения (6) полностью определяется своими значениями на  $\Gamma$  и восстанавливается по этим значениям по формуле (8).

**Теорема 3.** Пусть  $\{v_r\}$  — произвольная вектор-функция размерности  $m$ , определенной на  $\Gamma$ , и пусть  $G_n$  — произвольное фундаментальное решение. Тогда формула

$$u_n = \sum_{r \in \Gamma} \left( \sum_{k \in K_r} G_{n-r+k} A_k \right) v_r + \sum_{m \in D_0} G_{n-m} f_m, \quad n \in D, \quad (9)$$

задает некоторое решение уравнения (6).

**Доказательство.** Применим оператор  $L$  к вектор-функции  $\{u_n\}$ , определенной формулой (9):

$$Lu_n = \sum_{r \in \Gamma} \left[ \sum_{k \in K_r} (L G_{n-r+k}) A_k \right] v_r + \sum_{m \in D_0} (L G_{n-m}) f_m, \quad n \in D_0. \quad (10)$$

Вычислим правую часть. В силу (4) имеем

$$L G_{n-r+k} = \begin{cases} E, & \text{если } n = r - k, \\ 0, & \text{если } n \neq r - k. \end{cases}$$

Но в силу определения множества  $K_r$  точка  $n = r - k$  не принадлежит  $D_0$ , так что первое слагаемое в правой части формулы (10) есть нулевой вектор. Второе слагаемое есть, очевидно,  $f_n$ , так что  $Lu_n = f_n$ , и теорема доказана.

Равенство (8) аналогично интегральной формуле Коши для аналитических функций  $\varphi(z)$  в ограниченной области  $d$  с границей  $\gamma$ :

$$\frac{1}{2\pi i} \oint_{\zeta \in \gamma} \frac{\varphi(\zeta)}{\zeta - z} d\zeta = \begin{cases} \varphi(z), & \text{если } z \in d, \\ 0, & \text{если } z \in \bar{d} \cup \gamma. \end{cases} \quad (11)$$

При этом роль аналитических функций, границы области и ядра Коши  $\frac{1}{2\pi i} \frac{1}{\zeta - z}$  играют соответственно решения  $\{u_n\}$  задачи (6), граница  $\Gamma$  точной области  $D$  и выражение  $\left( \sum_{k \in K_r} G_{n-r+k} A_k \right)$ , учитывающее через

множество  $K_r$ , по которому ведется суммирование, структуру границы вблизи точки  $r \in \Gamma$ .

Формулу (9) в таком случае естественно сравнить с интегральной формулой типа Коши. Формула (8) аналогична также формуле Грина для уравнения Лапласа.

Подчеркнем, однако, следующее существенное различие между формулами (11) и (8): интегральная формула Коши справедлива только строго внутри области  $d$ , а разностная формула (8) — всюду на  $D$ , включая точки границы  $\Gamma$ .

Аналогичное различие имеется также между формулой (9) и формулой Грина.

### 5. Внутренние граничные условия.

**Теорема 4.** Пусть  $G_n$  — какое-нибудь фундаментальное решение уравнения (1). Для того чтобы заданную на  $\Gamma$  вектор-функцию  $\{u_r\}$ ,  $r \in \Gamma$ , можно было доопределить всюду в ограниченной сеточной области  $D$  до некоторого решения уравнения (6), необходимо и достаточно, чтобы при всех  $n \in \Gamma$  выполнялись равенства

$$\sum_{r \in \Gamma} \left( \sum_{k \in K_r} G_{n-r+k} \right) u_r + \sum_{m \in D_0} G_{n-m} f_m = u_n, \quad n \in \Gamma. \quad (12)$$

**Доказательство.** Если  $\{u_r\}$ ,  $r \in \Gamma$ , можно доопределить всюду на  $D$  до некоторого решения  $\{u_n\}$ ,  $n \in D$ , то, применив к этому решению формулу (8), а затем рассматривая полученное равенство только при  $n \in \Gamma$ , убедимся в выполнении (12). Обратно, если  $\{u_r\}$ ,  $r \in \Gamma$ , удовлетворяет (12), то примем  $v_r \equiv u_r$  и построим некоторое решение  $\{u_n\}$ ,  $n \in D$ , по формуле (9). В силу (12) граничные значения этого решения  $\{u_r\}$ ,  $r \in \Gamma$ , совпадут с заданными.

Доказанная теорема 4 дает основание назвать равенства (12) *внутренними граничными условиями*: эти условия не задаются извне, а являются следствиями самого разностного уравнения.

Если формулы (8) и (9) трактовать как аналоги интегральных формул Коши и типа Коши, то внутренние граничные условия аналогичны классическим условиям Сохоцкого — Племелья, при которых заданную на границе  $\gamma$  области  $d$  на комплексной плоскости функцию  $\varphi(z)$  можно доопределить всюду в области  $d$  до некоторой аналитической функции.

Формулу (8) можно понимать как адекватную системе (6) разностную формулу Грина, которая неявно учитывает «скачки потенциалов» на границе  $\Gamma$  и приводит к внутренним граничным условиям (12).

**6. Оператор граничного проектирования.** Возможна отличная от (12) запись внутренних граничных условий. Будем обозначать через  $U_\Gamma$  линейное пространство всех сеточных вектор-функций  $u_\Gamma = \{u_r\}$ ,  $r \in \Gamma$ , а через  $U'_\Gamma$  — подпространство тех из них, которые можно доопределить всюду в  $D$  до решений  $\{u_n\}$ ,  $n \in D$ , однородного уравнения, соответствующего уравнению (6).

Определим линейное отображение  $P$ ,  $u_\Gamma = P v_\Gamma$ , пространства  $U_\Gamma$  в себя следующей формулой:

$$u_n = \sum_{r \in \Gamma} \left( \sum_{k \in K_r} G_{n-r+k} A_k \right) v_r, \quad n \in \Gamma. \quad (13)$$

**Теорема 5.** Оператор  $P$  есть оператор проектирования  $U_\Gamma$  на  $U'_\Gamma$ .

**Доказательство.** Действительно, при любом  $v_\Gamma \in U_\Gamma$  в силу теоремы 3 элемент  $u_\Gamma = P v_\Gamma$  принадлежит  $U'_\Gamma$ . Если  $v_\Gamma \in U'_\Gamma$ , то в силу теоремы 2 получим  $P u_\Gamma = u_\Gamma$ . Теорема доказана.

Оператор  $P$ , определенный формулой (13), будем называть *оператором граничного проектирования*. С его помощью внутренние граничные условия (12) в случае  $f_n \equiv 0$  записываются в форме

$$u_\Gamma - P u_\Gamma = 0. \quad (14)$$

Подчеркнем, что оператор граничного проектирования зависит от выбора фундаментального решения  $G_n$ .

**7. Общая краевая задача.** В силу следствия из теоремы 2 каждое решение уравнения (6) восстанавливается по его значениям на границе  $\Gamma$ . Это

дает основание определить общую линейную краевую задачу для уравнения (6) как краевую задачу вида

$$\left. \begin{aligned} Lu &\equiv \sum_{k \in K} A_k u_{n+k} = f_n, \quad n \in D_0, \\ lu_\Gamma &= \varphi, \quad \varphi \in \Phi, \end{aligned} \right\} \quad (15)$$

где  $l$  — какой-нибудь линейный оператор, отображающий пространство  $U_\Gamma$  на некоторое линейное пространство  $\Phi$ .

Естественные разностные схемы, аппроксимирующие первую, вторую или третью краевые задачи для уравнения Пуассона, например, легко записать в виде (15).

Название «общая краевая задача» несколько условно: могут встретиться разностные краевые задачи, имеющие иной, чем (15), вид. Например, таковы естественные разностные схемы для дифференциальных краевых задач, в которых порядок дифференциального уравнения ниже порядка дифференциальных краевых условий.

**8. Основная идея метода внутренних граничных условий.** Пусть для простоты  $f_n \equiv 0$ . Между разностной краевой задачей

$$Lu = 0, \quad lu_\Gamma = \varphi \quad (16)$$

и между задачами

$$u_\Gamma - Pu_\Gamma = 0, \quad lu_\Gamma = \varphi \quad (17)$$

существует тесная связь. Именно, граничные значения  $u_\Gamma = \{u_r\}$ ,  $r \in \Gamma$ , каждого решения  $\{u_n\}$ ,  $n \in D$ , задачи (16) удовлетворяют в силу теоремы 2 уравнениям (17). Обратно, каждое решение  $u_\Gamma = \{u_r\}$ ,  $r \in \Gamma$ , задачи (17) в силу теоремы 4 и следствия из теоремы 2 можно единственным образом определить всюду в  $D$  до решения задачи (16). Основная идея метода внутренних граничных условий состоит в переходе от исходной разностной краевой задачи (16) к системе уравнений (17) на границе  $\Gamma$ . Продвижения, которые при этом удаётся получить, основаны на двух обстоятельствах. Первое из них — малое по сравнению с задачей (16) число неизвестных, участвующих в задаче (17). Второе — специальный вид системы (17), в структуру которой органически включен оператор граничного проектирования.

**9. Устойчивость внутренних граничных условий.** Можно опасаться, что внутренние граничные условия  $u_\Gamma - Pu_\Gamma = 0$  «почти вырождены», и поэтому задача (17) плохо обусловлена независимо от вида оператора  $l$ , так что переход от задачи (16) к задаче (17) связан с потерей вычислительной устойчивости.

Будем считать, что пространство  $\Phi$  вложено в пространство  $U_\Gamma$ , введем в пространстве  $U_\Gamma$  (а значит, и в пространстве  $\Phi \subset U_\Gamma$ ) норму  $\|\cdot\|$  и докажем теорему, означающую, что при переходе от задачи (16) к задаче (17) не происходит потери вычислительной устойчивости.

**Теорема 6.** Пусть задача (17) имеет решение  $u_\Gamma$  при любом  $\varphi \in \Phi$ , причем выполнена оценка

$$\|u_\Gamma\| \leq c \|\varphi\|, \quad (18)$$

где  $c$  от  $\varphi$  не зависит. Пусть, далее,  $v_\Gamma$  — произвольный элемент из  $U_\Gamma$ . Обозначим

$$v_\Gamma - Pv_\Gamma = \tilde{\varphi}, \quad lv_\Gamma = \tilde{\varphi}. \quad (19)$$

Тогда справедлива оценка

$$\|v_\Gamma\| \leq c (\|\tilde{\varphi}\| + \|l\| \|\tilde{\varphi}\|) + \|\tilde{\varphi}\| \quad (20)$$



Если рассматривать (19) как уравнения для определения  $v_\Gamma$ , то оценка (20) означает, что чувствительность решения задачи (19) к возмущениям  $\tilde{\psi}$  правой части внутренних граничных условий характеризуется постоянной  $c$  из оценки (18), т. е. чувствительностью решения к возмущениям правой части заданного граничного условия  $l u_\Gamma = \varphi$ .

**Доказательство.** Обозначим

$$z_\Gamma \equiv v_\Gamma - \tilde{\psi} = P v_\Gamma.$$

В силу теоремы 5 имеем  $z_\Gamma = P v_\Gamma \in U'_\Gamma$ . Поэтому

$$z_\Gamma - P z_\Gamma = 0, \quad l z_\Gamma = l(u_\Gamma - \tilde{\psi}) = \tilde{\varphi} - l \tilde{\psi},$$

т. е.  $z_\Gamma$  удовлетворяет системе вида (17) и в силу (18) оценке

$$\|z_\Gamma\| \leq c \|\tilde{\varphi} - l \tilde{\psi}\| \leq c (\|\tilde{\varphi}\| + \|l\| \|\tilde{\psi}\|).$$

Отсюда с учетом тождества  $z_\Gamma = v_\Gamma - \tilde{\psi}$  следует (20).

**10. Дополнительная идея.** Изложим полезную при численном решении краевых задач для уравнений с частными производными идею применительно к следующей задаче.

Пусть функция  $u(x, y)$  определена в некоторой области  $d$  с достаточно гладкой границей  $\gamma$  как решение задачи Дирихле

$$\left. \begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= 0, & (x, y) \in d, \\ u|_\gamma &= a(s), \end{aligned} \right\}$$

и требуется найти производную

$$\left. \frac{\partial u}{\partial n} \right|_\gamma = b(s)$$

в направлении внутренней нормали. Такая задача возникает, если по температуре  $u|_\gamma = a(s)$  на границе  $\gamma$  требуется найти установившийся тепловой поток через границу.

Под  $s$  понимается длина дуги вдоль границы  $\gamma$ , причем будем считать для определенности, что полная длина границы  $\gamma$  есть  $2\pi$ . Функцию

$$\left. \frac{\partial u}{\partial n} \right|_\gamma = b(s)$$

будем искать приближено в виде частичной суммы

$$b(s) = \sum_{j=0}^k (\alpha_j \cos js + \beta_j \sin js)$$

ее ряда Фурье. Для определения коэффициентов  $\alpha_j$  и  $\beta_j$  воспользуемся методом внутренних граничных условий.

Зададим  $h > 0$ , построим сетку

$$(x_{n_1}, y_{n_2}) = (n_1 h, n_2 h)$$

и разностное уравнение

$$u_{n_1+1, n_2} + u_{n_1-1, n_2} + u_{n_1, n_2+1} + u_{n_1, n_2-1} - 4u_{n_1, n_2} = 0.$$

Отнесем к  $D_0 = D_0^h$  все те точки сетки, которые вместе со всеми четырьмя

соседними точками принадлежат  $dU\gamma$ . Тогда определится сеточная область  $D = D^h$ , ее граница  $\Gamma = \Gamma^h$  и внутренние граничные условия  $u_\Gamma - Pu_\Gamma = 0$

Идея состоит в том, чтобы по функции  $u|_\gamma = a(s)$  и функции

$$\left. \frac{\partial u}{\partial n} \right|_\gamma = b(s),$$

записанной в виде ряда с неопределенными коэффициентами, продолжить решение по формуле Тейлора с границы  $\gamma$  в приграничную полосу, где лежит граница сеточной области  $\Gamma^h$ ; затем подобрать неопределенные коэффициенты

$$\alpha_j = \alpha_j^h, \quad \beta_j = \beta_j^h$$

из условия минимизации невязки, возникающей при подстановке продолженной с границы  $\gamma$  в приграничную полосу функции  $u(x, y)$ , во внутренние граничные условия.

**11. Сопоставление метода внутренних граничных условий с методом сингулярных интегральных уравнений.** В начале Дополнения мы указывали на аналогию между методом внутренних граничных условий и методом сингулярных интегральных уравнений, которая не является полной. Здесь мы сопоставим эти методы, уточняя аналогию и выявляя существенные различия.

Для сопоставления сначала опишем идею метода сингулярных интегральных уравнений для дифференциальных краевых задач на примере задачи

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} - \mu u = 0, \quad x = (x_1, x_2) \in d, \quad (21)$$

$$a_0 u_0 + a_1 u_1 = \varphi(x), \quad x = (x_1, x_2) \in \gamma, \quad (22)$$

где  $\mu = \text{const} > 0$ ,  $d$  — ограниченная область,  $\gamma$  — ее граница. Краевое условие (22) связывает решение  $u = u_0(x)$  на границе области и его производную по направлению внутренней нормали  $\partial u / \partial \nu = u_1(x)$ . Коэффициенты  $a_0$  и  $a_1$  — заданные операторы.

Выпишем классическую формулу Грина для уравнения (21):

$$u(x) = \int_{y \in \gamma} \left[ g(x-y) \frac{\partial u}{\partial \nu} - u \frac{\partial g}{\partial \nu} \right] d\sigma_y, \quad (23)$$

где  $g(x)$  — фундаментальное решение уравнения (21), стремящееся к нулю на бесконечности. Устремим  $x$  к границе  $\gamma$ . Воспользовавшись свойствами потенциалов простого и двойного слоев, получим на границе  $\gamma$  соотношение вида

$$u_0 = b_0 u_0 + b_1 u_1, \quad (24)$$

связывающее решение  $u(x)$  и его нормальную производную  $\partial u / \partial \nu = u_1(x)$  на границе области;  $b_0$  и  $b_1$  — некоторые известные интегральные операторы. Переход от задачи (21), (22) к равносильной системе уравнений (24), (24) относительно функций  $u_0(x)$  и  $u_1(x)$ , определенных на границе  $\gamma$ , и составляет сущность метода сингулярных интегральных уравнений.

Для сравнения рассмотрим теперь метод внутренних граничных условий применительно к следующей общей краевой задаче для разностного аналога уравнения (21) в квадратной сеточной области

$$u_{n_1-1, n_2} + u_{n_1, n_2+1} + u_{n_1+1, n_2} + u_{n_1, n_2-1} - (4 + \mu) u_{n_1, n_2} = 0, \quad (25)$$

$$-N < n_1, n_2 < N,$$

$$lu_\Gamma = \varphi. \quad (26)$$

Запишем внутренние граничные условия  $u_\Gamma - Pu_\Gamma = 0$  в удобной для дальнейшего форме.

Легко проверить, что формула (9) в этом случае может быть переписана в форме

$$u_n \equiv \sum_{r \in Q_0} [G_{n-r}(\Delta_\nu u_r) - u_r(\Delta_\nu G_{n-r})] + \sum_{r \in Q_0} u_r \delta_r^n, \quad n \in D, \quad (27)$$

где  $Q_0$  — совокупность точек  $\Gamma$ , лежащих на сторонах квадрата  $|n_1| = N$ ,  $|n_2| = N$ , т. е. на внешнем слое двухслойной границы  $\Gamma$  сеточной квадратной области (рис. 57), а  $\Delta_\nu$  — разностный аналог производной по направлению внутренней нормали.

Заметим, что формула (27) была бы полным аналогом классической формулы Грина (23), если бы в ее правой части отсутствовало «сингулярное слагаемое»  $\sum \delta_r^n u_r$ . Однако в таком случае равенство (27) имело бы место не при всех  $n \in D$ , а лишь при  $n \in D_0$ . Из него нельзя было бы получить тогда внутренние граничные условия  $u_\Gamma - Pu_\Gamma = 0$ . Эти условия получаются из (27), если  $n$  пробегает не всю область  $D$ , а только точки границы  $\Gamma$ , и записываются двумя системами равенств

$$u_n = \sum_{r \in Q_0} [G_{n-r}(\Delta_\nu u_r) - u_r(\Delta_\nu G_{n-r})] + u_n, \quad n \in Q_0, \quad (28)$$

$$u_n = \sum_{r \in Q_0} [G_{n-r}(\Delta_\nu u_r) - u_r(\Delta_\nu G_{n-r})], \quad n \in \Gamma \setminus Q_0, \quad (29)$$

отвечающих, соответственно, точкам  $n \in Q_0$  внешнего и точкам  $n \in \Gamma \setminus Q_0$  внутреннего слоев двойной границы  $\Gamma$ . В качестве  $G_n$  в равенствах (28) и (29) будем использовать ограниченное фундаментальное решение.

Можно показать, что внутренние граничные условия  $u_\Gamma - Pu_\Gamma = 0$ , т. е. система уравнений (28), (29), алгебраически равносильна каждой из отдельно взятых подсистем (28) или (29).

Подсистема (28) аналогична интегральному соотношению (24), так что разностным аналогом задачи (22), (24) является задача (27), (28), по не задача

$$lu_\Gamma = \varphi, \quad u_\Gamma - Pu_\Gamma = 0,$$

записываемая равенствами (27)–(29), которая рассматривается в методе внутренних граничных условий.

Имеется очевидная разница между внутренними граничными условиями  $u_\Gamma - Pu_\Gamma = 0$ , т. е. системой (28), (29), и одной только подсистемой (28). Внутренние граничные условия

$$u_\Gamma - Pu_\Gamma = 0$$

содержат избыточные равенства (29). В этом смысле разностные внутренние граничные условия

$$u_\Gamma - Pu_\Gamma = 0$$

больше похожи не на интегральное соотношение (24), а на условия Сохоцкого — Племяля для аналитических функций. Эти последние представляют собой два вещественных соотношения, связывающих две вещественных функции, но они не независимы, и многообразие удовлетворяющих условиям Сохоцкого — Племяля пар функций зависит от одной произвольной вещественной функции.

Отметим, что внутренние граничные условия

$$u_{\Gamma} - Pu_{\Gamma} = 0$$

выгодно отличаются от равносильной им подсистемы (28) тем, что в их структуру входит оператор граничного проектирования. Благодаря этому обстоятельству задача

$$Lu_{\Gamma} = \varphi, u_{\Gamma} - Pu_{\Gamma} = \psi$$

устойчива в смысле теоремы 6 относительно возмущения правой части  $\psi$ . В общем случае при вычеркивании части уравнений из числа составляющих систему  $u_{\Gamma} - Pu_{\Gamma} = 0$  может получиться подсистема, алгебраически равносильная исходной, но уже не обладающая свойством устойчивости.

Можно показать, что в нашем примере (25), (26) вместо

$$u_{\Gamma} - Pu_{\Gamma} = 0$$

удобнее использовать не подсистему (28), а подсистему (29), которая устойчива и в отличие от подсистемы (28) состоит из независимых уравнений — ее ранг равен числу составляющих ее уравнений.

Итак, в рассматриваемом примере аналогия между методом внутренних граничных условий и методом сингулярных интегральных уравнений, аналогичным условию Сохоцкого — Племеля, не является полной.

Тем более, нет полной аналогии с классическим методом интегральных уравнений, в котором искомой функцией является не само решение исходной задачи (21), (22) на границе, а некоторая вспомогательная плотность потенциала простого или двойного слоя.

В заключение заметим, что мы употребляем выражение «метод сингулярных интегральных уравнений», поскольку условие Сохоцкого — Племеля содержит сингулярный интеграл. В примере этого пункта условие (26) содержит сходящиеся несобственные интегралы.

## БИБЛИОГРАФИЧЕСКИЕ КОММЕНТАРИИ

**К гл. 1, §§ 1, 2.** С общей теорией линейных разностных уравнений можно познакомиться, например, по гл. V книги [7].

**К гл. 2, § 5.** С методом прогонки и его обоснованием для некоторого класса разностных краевых задач авторы впервые познакомились в 1953 году по рукописи статьи И. М. Гельфанда и О. В. Локушевского «Метод прогонки для решения разностных уравнений» (см., например, [10]). Существуют варианты прогонки, предназначенные для вычисления решений разностных краевых задач, не рассмотренных в нашей книге. С результатами и библиографией можно познакомиться по книгам [4], [15], [23] и др.

**К гл. 3.** Идея использовать при обосновании прогонки непосредственно свойство хорошей обусловленности разностной краевой задачи была высказана Н. С. Бахваловым. Некоторые шаги в осуществление этой идеи были сделаны при изложении прогонки в книге [10], а затем В. В. Огневой, ЖВМ и МФ 7, № 4 (1967), которой принадлежит идея рассмотрения урезанных систем. Модифицированное изложение этой работы имеет в книге [8].

Приведенное в § 6 обоснование хорошей обусловленности разностной краевой задачи использует дипломную работу студента Новосибирского университета Багисбаева, которому, в частности, принадлежит пример, показывающий, что условие гладкости коэффициентов нельзя игнорировать.

**К гл. 6, §§ 19, 20.** Подробнее познакомиться с методами численного решения обыкновенных дифференциальных уравнений можно по книге [4] и указанной в ней литературе.

Разностные схемы для некоторых важных классов дифференциальных уравнений с разрывными коэффициентами построены в теории однородных разностных схем А. Н. Тихонова и А. А. Самарского и изложены в одной из глав книги [23].

**К гл. 7, § 21.** Понятие устойчивости разностных схем относительно ошибок округления при задании начальных данных впервые описано Дж. фон Нейманом и Р. Д. Рихтмайером в 1950 году (см. сб. переводов «Механика», в. 1, 1951) в работе, посвященной расчету газодинамических скачков. Первая система определений устойчивости и аппроксимации, при которой сходимость является следствием аппроксимации и устойчивости, была предложена В. С. Рябеньким, ДАН СССР 86, № 6 (1952), в случае разностных аналогов задачи Коши для систем уравнений с частными производными.

Принятая в нашей книге система основных определений и теорема о том, что из аппроксимации и устойчивости следует сходимость, близки к предложенным А. Ф. Филипповым, ДАН СССР 100, № 6 (1955). См. также [22] или [10]. Отличие состоит главным образом в том, что мы используем более универсальное, чем А. Ф. Филиппов, определение аппроксимации.

Существуют другие естественные системы определений основных понятий, при которых аппроксимация и устойчивость обеспечивают сходимость. Среди них наиболее известна система определений П. Д. Лакса, предложенная в 1956 году (см., например, [20]). В теории Лакса рассматриваются разностные схемы для нестационарных задач, причем предполагается, что эти

разностные схемы действуют не в пространстве сеточных функций, а в том же функциональном пространстве, что и дифференциальное уравнение. При этом (дополнительно) предполагении доказывается, что для аппроксимирующей разностной схемы устойчивость и сходимости имеют место одновременно. Эта теорема эквивалентности Лакса является одной из конкретизаций более общей конструкции Л. В. Канторовича, УМН 3, в 6 (1948).

В последние годы А. А. Самарский предложил и развил в соавторстве с А. В. Гулиным теорию устойчивости, применимую к весьма широкому классу разностных схем (см. [23], [24] и § 43 настоящей книги).

С новыми результатами, библиографией и обзорами работ по устойчивости разностных схем можно познакомиться по кингам [10], [15], [20]—[28].

Следует сказать, что в работе 1923 года Р. Куранта, К. Фридрихса и Г. Леви (см. УМН 8 (1940)) и во многих других работах, где метод конечных разностей используется для доказательства существования решений дифференциальных уравнений, устанавливаются неравенства, которые в современной терминологии можно истолковать как устойчивость в тех или иных нормах. Однако понятие устойчивости возникло в связи с использованием разностных схем для приближенного вычисления решений в предположении, что эти решения существуют. Поэтому устойчивость изучается обычно в более слабых нормах, чем это нужно для доказательства существования.

Отметим, что впервые метод конечных разностей был использован для доказательства существования решений уравнений с частными производными в 1924 году Л. А. Люстерником (см. УМН 8 (1940)), который рассматривал уравнение Лапласа.

**К гл. 7, § 22, п. 3.** Излагаемый здесь прием построения разностных схем предложен в работах: Р. L. I. Brian, A. I. Ch. E. J. 7 (1961); J. Douglas, Num. Math. 4 (1962); J. Douglas, Trans. Amer. Soc. 89 (1958); С. К. Годунов, Разностные методы решения уравнений газовой динамики. Новосибирск, 1962 (ротапринт). Двумерный вариант рассмотренной в этом пункте схемы с пересечением Лакса—Вендрова [20], для газодинамических задач предложен Л. А. Чудовым (см. обзорную статью Г. С. Рослякова и Г. Ф. Теленина в сб. «Численные методы в газовой динамике», М., Изд-во МГУ, в. 2, 1963). Идея метода Рунге—Кутты была применена В. В. Русановым (препринт ИПМ АН СССР, 1967) для построения разностной схемы третьего порядка точности в случае газодинамических расчетов.

Л. А. Чудов (статья в сб. «Некоторые применения метода сеток в газовой динамике», в. 1. «Течения в пограничном слое», Изд-во МГУ, 1971) для уравнений параболического типа построил разностную схему типа Рунге—Кутты второго порядка точности, обладающую хорошими сглаживающими свойствами. Схемы с пересечением применяются во многих газодинамических расчетах. См., например, [1]. Существуют и другие методы построения разностных схем (см. [4], [13], [19]—[28]).

**К гл. 8, § 25, п. 5.** Насколько известно авторам, возможность использовать дифференциальные приближения для исследования разностных уравнений впервые заметил в 50-х годах А. И. Жуков (сообщение на семинаре ИПМ), которому принадлежит рассмотренный здесь пример. Теория дифференциальных приближений, в которой изучаются асимптотические и групповые свойства интересных классов разностных уравнений, построена Н. Н. Яненко и Ю. И. Шокиным, Сиб. матем. ж. 10, № 5 (1969); Численные методы мех. сплошной среды, 2, № 2 (1971). К этим же вопросам относятся статьи Н. Н. Кузнецова, ДАН 200, № 5, (1971); ДАН 204, № 2 (1972); ЖВМ и Мф, 12, № 12 (1972).

**К гл. 8, § 26, п. 1.** Идея замораживания коэффициентов во внутренних точках предложена в цитированной выше статье Неймана и Рихтмайера (см. примечание к § 21).

**К гл. 8, § 26, п. 2.** Признак К. И. Бабенко и И. М. Гельфанда был доложен в их совместном с О. В. Локуциевским докладе на конференции по функциональному анализу в Москве в 1956 году. См. также [2] и примечания к гл. 14, помещенные ниже.

**К гл. 8, § 27.** Существует очень экономный по числу арифметических действий алгоритм вычисления коэффициентов конечного ряда Фурье, называемый быстрым преобразованием Фурье. См., например, [4] или [15].

В частности, отметим оценку погрешности разностного решения уравнения Пуассона, найденную Е. А. Волковым (Тр. Матем. ин-та им. В. А. Стеклова 117 (1972)) в условиях отсутствия аппроксимации оператора Лапласа со вторым порядком в числе слоев сетки, неограниченно растущем при измельчении шага. Эта оценка в то же время является более сильной, чем равномерная оценка второго порядка, так как она устанавливает дополнительное убывание погрешностей вблизи границы области.

Конечные ряды Фурье для анализа нестационарных разностных уравнений, по-видимому, впервые использовала О. А. Ладыженская. С помощью этого аппарата ею была найдена сходящаяся неявная разностная схема для гиперболических по И. Г. Петровскому систем. По-видимому, это был первый пример сходящейся неявной разностной схемы (О. А. Ладыженская, Автореферат, канд. дисс., ЛГУ, март 1949). См. также [13].

**К гл. 9.** См. книги [1]—[3], [9], [13], [14], [21], [26] и имеющуюся там библиографию; в сборниках статей и журналах постоянно появляются новые работы по численным методам механики сплошных сред.

**К гл. 10.** Схема переменных направлений (12) из § 32 построена Д. Писманом и Г. Рэкфордом в 1956 году (см., например, [5] или [28]); схема расщепления (7) из § 31 предложена Н. Н. Яненко, ДАН СССР 125, № 6 (1959). В настоящее время схемы расщепления построены для многих основных задач математической физики. См., например, [5], [15], [23], [27], [28], монографию Е. Г. Дьяконова, Разностные методы решения краевых задач, ч. 1 (1971) и ч. 2 (1972), изд-во МГУ, и имеющуюся там библиографию.

Видоизменение метода переменных направлений, получающееся путем объединения его с вариационным методом Рунта, предложено и использовано для вычисления собственных значений сильноэллиптических операторов и для решения разностного уравнения Лапласа в работах: Г. П. Прокопов, ЖВМ и МФ 8, № 1 (1968); С. К. Годунов и Г. П. Прокопов, ЖВМ и МФ 9, № 2 (1969). С. К. Годунов, В. В. Огунов, Г. П. Прокопов, Сб. «Дифференц. уравнения с частными производными», Труды симпозиума, посвященного 60-летию акад. С. Л. Соболева, 1970. Оригинальная конструкция локально-одномерных схем предложена И. В. Фрязиновым (ЖВМ и МФ 13, № 1, 3, 1973).

**К гл. 10, § 33.** Относительно метода крупных частиц О. М. Белоцерковского и Ю. М. Давыдова и его приложений, помимо работы, цитированной в § 33, см. монографию [3]; текст обзорного доклада О. М. Белоцерковского и В. Е. Яницкого на IV Всесоюз. конф. по динамике разреженного газа в 1975 году в Звенигороде; текст лекции О. М. Белоцерковского на Кармановских чтениях 1976 года в Брюсселе.

**К гл. 11, § 34.** Для разностного уравнения Пуассона в прямоугольнике самым экономным способом вычисления решения является быстрое преобразование Фурье (см. примечание к § 27). Разностными схемами для уравнений Лапласа и Пуассона в криволинейных областях, начиная от работы Л. А. Люстерника 1924 года, занимались многие авторы. См., например, [4], [16], [23] и имеющуюся там библиографию.

Оценки погрешности, выражающиеся непосредственно через исходные данные, получены для ряда схем, аппроксимирующих задачи Дирихле, Неймана и смешанную краевую задачу для уравнений Лапласа и Пуассона на прямоугольнике, прямоугольном параллелепипеде и некоторых треугольниках см. Е. А. Волков, Тр. Матем. ин-та им. В. А. Стеклова 74 (1966), 105 (1969), И. А. Султанова, ЖВМ и МФ 11, № 5 (1971) и библиограф. там же. Е. А. Волков

установил также (Тр. Матем. ин-та им. В. А. Стеклова 128 (1972)), что если разностный оператор в граничных узлах удовлетворяет некоторому условию адекватности стандартному пятиточечному разностному оператору Лапласа, то разностное решение уравнения Пуассона, продолженное с сетки на замкнутую область с криволинейной границей, при достаточной гладкости данных задачи аппроксимирует со вторым порядком относительно шага искомое решение вместе с производными до порядка  $n$  включительно,  $n \geq 0$  любое.

**К гл. 11, § 35.** Идея рассмотрения решений стационарных задач как предела решений нестационарных при возрастании времени впервые использована в 30-х годах А. Н. Тихоновым.

Однa из разностных схем установления для расчета стационарного сверхзвукового обтекания тел газом предложена С. К. Годуновым, А. В. Забродным и Г. П. Прокоповым, ЖВМ и МФ 1, № 6 (1961), см. [9]. Интересно отметить, что обоснование устойчивости этой схемы, описанное в работе К. А. Багриновского и С. К. Годунова, ДАН СССР 115, № 3 (1957), использует расщепление разностных операторов. Имеется ряд работ многих авторов в направлении расчета стационарных задач установлением.

Один из первых эффективных методов ускорения сходимости при решении разностного уравнения Пуассона указал Л. А. Люстерник, Тр. Матем. ин-та им. В. А. Стеклова 20 (1947).

**К гл. 11, § 36.** Многочлены Чебышева для выбора оптимального набора итерационных параметров используются в различных задачах, начиная с работ А. А. Абрамова, М. К. Гавурина, Фландерса и Шортли, относящихся к 1950 году.

Новые результаты, библиография и обзоры с различных точек зрения итерационных методов решения разностных эллиптических краевых задач содержатся в книгах [5], [16], [23], [28]; в монографиях Е. Г. Дьяконова, Итерационные методы решения разностных аналогов краевых задач для уравнений эллиптического типа, Киев, 1970 (ротaпpинт); Г. И. Марчука и Ю. А. Кузнецова, Итерационные методы и квадратичные функционалы, Новосибирск, «Наука» СО, 1972 (ротaпpинт); в обзорной статье Р. П. Федоренко, УМН 28, № 2 (1973) и в др.

**К гл. 12.** Основная идея получения вариационно-разностных схем содержится в работе Р. Куранта (Courant R., Bull. Amer. Math. Soc. 49, № 1 (1943)). Независимо в инженерных расчетах прочности часто пользовались без теоретического обоснования различными реализациями вариационно-разностных схем под названием метода конечных элементов.

Систематическому изложению основ теории вариационно-разностных схем и некоторых их приложений посвящена монография Л. А. Оганесяна, В. Я. Ривкинца и Л. А. Руховца «Вариационно-разностные методы решения эллиптических уравнений», ч. 1 и 2, Тр. семинара по дифференц. уравнениям, Ин-т физики и математики АН Литовской ССР, в. 5, Вильнюс, 1973 и в. 8, Вильнюс, 1974. Ротaпp., которую мы использовали при работе над этой главой. См. также, например, [12], [18], [25].

В настоящее время вариационно-разностные схемы реализованы в виде хорошо отработанных программ на быстродействующие вычислительные машины для ряда задач теории упругости. См., например, [12]. Имеются численные реализации проекционно-разностного метода и для некоторых других задач, не только эллиптических. Ряд работ последнего времени помещен в двух сборниках «Вариационно-разностные методы в математической физике», Новосибирск, 1974 и Новосибирск, 1976.

**К гл. 13, § 42.** Стационарные решения часто используют для выяснения характера сходимости вблизи границ. См., например, С. К. Годунов, Матем. сб. 47 (89), 3 (1957).

**К гл. 13, § 43, п. 4.** Здесь использован написанный А. Ф. Филипповым п. 4 из § 6 книги [22].



К гл. 13, § 43, п. 5. Выбор скалярного умножения  $(u, v)_{B_k}$  по формуле (21), по-видимому, впервые предложен Н. Миньо в 1953 году в частном случае разностного аналога уравнения теплопроводности с переменными коэффициентами, а в более общей форме — в § 15 книги [22], где имеется также модифицированное изложение упомянутой работы Н. Миньо.

К гл. 13, § 43, п. 6. Первый из критериев устойчивости А. А. Самарского, приведенных в этом пункте, получается из теоремы 5, п. 6, § 1 гл. VI книги [23], если вместо гильбертова пространства рассматривать евклидово и положить  $\rho = 1$ . См. также п. 7, § 1, гл. VI книги [23].

К гл. 14, § 44. Понятие спектра семейства разностных операторов введено в книге [10], где авторы с помощью этого понятия, в частности, обосновали признак К. И. Бабенко и И. М. Гельфанда устойчивости нестационарных задач на отрезке. Там же доказано, что расположение спектра семейства операторов в единичном круге необходимо для устойчивости.

Теорема 2 получена В. С. Рябеньким, ДАН СССР 185, № 2 (1969).

К гл. 14, § 46. Понятие ядра спектра семейства операторов введено В. С. Рябеньким, ДАН СССР 185, № 2 (1969). Там же сформулированы теоремы 1—4.

Теорема А. В. Соколова для случая скалярных коэффициентов  $A_k, B_k$  опубликована в ДАН СССР 208, № 2 (1973). Доказательство в общем случае матричных коэффициентов содержится в его статье, Тр. Моск. матем. общ. 35, Изд-во МГУ, 1976.

К гл. 14, § 47. Здесь изложена статья В. С. Рябенького, ДАН СССР 193, 3 (1970).

К Дополнению. Метод внутренних граничных условий (МВГУ) предложен В. С. Рябеньким, Докт. дисс., Ин-т прикл. математики АН СССР (1969). В пп. 1—9 и 11 изложена часть статьи В. С. Рябенького, УМН 26, № 3 (1971). В этой статье даны также некоторые приложения МВГУ к исследованию и вычислению решений разностных краевых задач в простых и составных областях.

Содержание п. 10 опубликовано в докладе В. С. Рябенького на конференции, посвященной семидесятилетию акад. И. Г. Петровского в МГУ (январь 1976).

К п. 2 Дополнения. А. Я. Белянков, Матем. заметки 18, № 5 (1975), доказал существование фундаментального решения, растущего при  $\|n\|^2 = n_1^2 + \dots + n_s^2 \rightarrow \infty$  не быстрее некоторой степени  $\|n\|$ .

Им же построено так называемое циклическое фундаментальное решение, позволяющее строить внутренние граничные условия и допускающее эффективное построение с помощью быстрого преобразования Фурье, статья в сб. «Задачи механики и матем. физики», посвященном памяти акад. И. Г. Петровского, «Наука», 1976.

А. В. Забродин и В. В. Огнева, препринт ИПМ АН СССР (1973), использовали модифицированный ими вариант МВГУ для расчета нелинейной задачи теплопроводности на графах.

## ЛИТЕРАТУРА

1. Алалыкин Г. Б., Годунов С. К., Киреева И. Л., Плинер Л. А., Решение одномерных задач газовой динамики в подвижных сетках, М., «Наука», 1970.
2. Бабенко К. И., Воскресенский Г. П., Любимов А. Н., Русанов В. В., Пространственное обтекание гладких тел идеальным газом, М., «Наука», 1964.
3. Белоцерковский О. М. и др., Численное исследование современных задач газовой динамики, М., «Наука», 1974.
4. Бахвалов Н. С., Численные методы, М., «Наука», 1975.
5. Вазов В., Форсайт Дж., Разностные методы решения уравнений в частных производных, М., ИЛ, 1963.
6. Гавурич М. К., Лекции по методам вычислений, М., «Наука», 1971.
7. Гельфонд А. О., Исчисление конечных разностей, М., «Наука», 1967.
8. Годунов С. К., Уравнения математической физики, М., «Наука», 1971.
9. Годунов С. К., Забродин А. В., Иванов М. Я., Крайко А. Н., Прокопов Г. П., Численное решение многомерных задач газовой динамики, М., «Наука», 1976.
10. Годунов С. К., Рябенский В. С., Введение в теорию разностных схем, М., Физматгиз, 1962.
11. Дьяченко В. Ф., Основные понятия вычислительной математики, М., «Наука», 1972.
12. Зенкевич О., Метод конечных элементов в технике, М., «Мир», 1975.
13. Ладыженская О. А., Краевые задачи математической физики, М., «Наука», 1973.
14. Любимов А. Н., Русанов В. В., Течение газа около тупых тел, ч. 1, М., 1970.
15. Марчук Г. И., Методы вычислительной математики, Новосибирск, «Наука», 1978.
16. Марчук Г. И., Лебедев В. И., Численные методы в теории переноса нейтронов, М., Атомиздат, 1971.
17. Михлин С. Г., Численная реализация вариационных методов, М., «Наука», 1968.
18. Одэн Дж., Конечные элементы в нелинейной механике сплошных сред, М., «Мир», 1976.
19. Петровский И. Г., Лекции об уравнениях с частными производными, М., Физматгиз, 1961.
20. Рахтмайер Р. Д., Мортон К., Разностные методы решения краевых задач, М., «Мир», 1972.
21. Рождественский Б. Л., Янепко Н. Н., Системы квазилинейных уравнений и их приложения в газовой динамике, М., «Наука», 1968.
22. Рябенский В. С., Филиппов А. Ф., Об устойчивости разностных уравнений, М., Гостехиздат, 1956.

23. Самарский А. А., Введение в теорию разностных схем, М., «Наука», 1971.
24. Самарский А. А., Гулин А. В., Устойчивость разностных схем, М., «Наука», 1973.
25. Самарский А. А., Андреев В. Б., Разностные методы для эллиптических уравнений, М., «Наука», 1976.
26. Самарский А. А., Попов Ю. П., Разностные схемы газовой динамики, М., «Наука», 1975.
27. Тихонов А. Н., Самарский А. А., Уравнения математической физики, М., «Наука», 1972.
28. Яценко Н. Н., Метод дробных шагов решения многомерных задач математической физики, Новосибирск, «Наука», 1967.
29. Стренг Г., Фикс Дж., Теория метода конечных элементов, М., «Мир», 1977.

## ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Абсолютное ядро, свойства** 414  
**Адамс схема** 160, 162  
**Алгоритм вычислительно неустойчивый, пример** 53  
— для вычисления решения резонансной задачи Дирихле 305, 307, 309, 313, 324  
— прогонки 51, 52  
**Аппроксимационная вязкость** 237, 238  
**Аппроксимация** 92, 172  
— дифференциальной краевой задачи разностной схемой 91, 92, 96, 108, 172, 176  
— порядка  $h^k$  96—98, 172, 174  
— производной 77, 78, 103
- Бабенко — Гельфанда признак** 243  
**Белоцерковского — Давыдова метод** крупных частиц 296, 431
- Вариационные методы оценки собственных значений** 392  
— — решения краевых задач 327  
**Вариационная постановка задачи Дирихле** 329  
— — третьей краевой задачи 331  
**Вариационно-разностная схема** 321  
— — —, определение 344, 345  
— — — для задачи Дирихле 346  
— — — третьей краевой задачи 354  
**Выглаживание решения разностной задачи** 237
- Галеркина метод** 341, 344  
**Границы сеточной области** 420  
**Граничные условия внутренние** 423  
**Групповая разностная схема** 281  
— — разностной схемы, примеры построения 204, 205
- Дифференциальная краевая задача, обобщенное решение** 269—275  
— — —, символическая запись 12, 84  
— — — эллиптическая 289  
**Дифференциальное приближение разностного уравнения** 238, 430  
— уравнение дивергентного вида 280  
**Дугласа — Рэкфорда метод** 320
- Задача Коши для волнового уравнения, аппроксимирующая разностная схема** 229  
— — — уравнения теплопроводности, аппроксимирующие разностные схемы 188, 189  
— модельная 209, 251  
— о распаде разрыва 273, 281  
— эволюционная 221, 361  
**Замораживание коэффициентов во внутренних точках** 240
- Инварианты Римана** 374  
**Интегральная формула Коши, разностный аналог** 421, 422  
— — типа Коши, разностный аналог 421, 422  
**Интегральное представление решения разностной задачи Коши** 232, 233  
**Итерационные методы, выбор точности** 311  
— параметры, оптимальный выбор 309, 311  
— —, нумерация 316, 320  
— —, чебышевский набор 313—315  
**Итерационный процесс Ричардсона** 313  
**Итерация с переменным шагом** 313
- Колмогоровский поперечник** 339, 349  
**Конечных элементов метод** 327, 432

- Коши интегральная формула, разностный аналог 421—422  
 — разностная задача 221, 239
- Краевая задача дифференциальная — см. Дифференциальная краевая задача
- — общая для систем разностных уравнений на сеточном отрезке, условие хорошей обусловленности 48
- — разностная — см. Разностная краевая задача
- Критерии Самарского устойчивости разностных схем 394, 395
- Критерий хорошей обусловленности разностной краевой задачи 42, 44, 60
- Куранта, Фридриха и Леви условие 211, 212, 217, 219
- Лакса теорема эквивалентности 430
- Линейное пространство нормированное 87
- Мера обусловленности система линейных уравнений 40
- Метод внутренних граничных условий 419, 427
- Дугласа — Рэкфорда 320
- конечных разностей 11, 84, 277
- — элементов 327, 432
- крупных частиц Белоцерковского — Давыдова 296, 439
- неопределенных коэффициентов 191
- Ньютона 169
- прогонки 51, 53, 67, 70, 169, 429
- —, варианты 429
- —, история 429
- —, обоснование 56—65
- —, оценка прогоночных коэффициентов 65
- релаксационный Федоренко 323—326
- сеток 84
- стрельбы 55, 167
- установления 304
- характеристик 277
- Модальная задача 209, 251, 261, 268
- Невязка 92
- Неймана метод исследования эволюционных разностных задач 221
- спектральный признак устойчивости разностной задачи Коши 222, 223, 229, 338
- Норма 87, 88, 101, 115—118, 363, 366, 387
- энергетическая 256, 394, 395
- Ньютона метод 169
- Обобщение решения дифференциального уравнения 268, 269, 273—275
- Обусловленность, хорошая обусловленность 39
- Оператор граничного проектирования 423, 421
- перехода 127, 136, 366, 369
- —, конструирование 375
- —, оценка норм степеней 366, 387, 400
- —, оценки собственных значений 391—393
- —, признаки самосопряженности 390
- Ошибка окружения 53, 57, 107, 144, 317
- Парсевалья равенство 232, 252
- Поперечник колмогоровский 339, 349
- Порядок разностного уравнения 18, 21
- точности разностной схемы 72, 75, 90, 121, 145
- Преобразование Фурье быстрое 431
- Признак Бабенко — Гельфанда устойчивости нестационарных задач на отрезке 343, 431
- устойчивости разностной схемы, достаточный 134, 234, 368, 372, 374, 387, 389, 394
- — — —, необходимый 368, 369
- — разностных схем решения задачи Коши 123, 234, 235
- — спектральный — см. Спектральный признак устойчивости
- хорошей обусловленности 40—42, 44, 60
- Признаки самосопряженности разностных операторов 390
- Принцип замороженных коэффициентов 236, 240, 241
- максимума 177, 269, 900
- Прогонка 51, 52
- Проекционно-разностная схема 327, 344, 345
- Производная, замена разностными отношениями 103
- Разностная задача Коши 221
- — —, анализ устойчивости 224—231

- Разностная задача Коши, интегральное представление решения 232, 233
- — —, спектральный признак устойчивости Неймана 222, 223, 229, 388
- — —, условия устойчивости 221, 233, 235
- —, краевая общая 424
- — —, приемы построения 189
- — —, признаки хорошей обусловленности 40—43
- — — с возмущенными коэффициентами, оценка решений 56, 58, 59
- — —, символическая запись 88
- — — устойчивая 125, 176
- — —, хорошо обусловленная 39, 56, 64
- схема абсолютно устойчивая 267
- — Адамса 160, 162
- схема Годунова 281
- — дивергентная 280—283
- — для волнового уравнения 229
- — — задачи Дирихле 190, 217, 218, 298, 302, 303
- — — — о колебаниях струны 260
- — — интегрального уравнения 114
- — — краевой задачи 91, 92, 168, 172, 176
- — — системы уравнений акустики 374
- — — уравнения с разрывными коэффициентами 429
- — — с частными производными 171
- — — — теплопроводности 188, 189, 254, 257, 305, 362, 372—373
- —, каноническая запись 125, 133, 365
- — консервативная 280
- — пегибкая 188
- — неустойчивая 81
- — неявная 189, 265, 266
- — переменных направлений 288, 296, 491
- —, порядок точности 72, 75, 100, 145
- —, построение граничных условий 204
- —, — на основе интегрального закона сохранения 278, 283
- — предиктор-корректор — см. Разностная схема с пересчетом
- —, приведение к каноническому виду 365, 366
- Разностная задача, приемы построения 103, 105, 157, 172, 173, 183, 204, 274, 276, 278, 284, 430
- —, признаки устойчивости 123, 134, 146, 138, 211, 219, 221, 222—225, 368, 371—389
- —, проверка сходимости 71, 91, 118, 172
- —, разбиение на подсистемы 100, 102
- —, расщепления 284, 286, 289, 297, 431
- — Рунге — Кутта 105, 157, 158, 162
- — с искусственной вязкостью 276
- — с пересчетом 105, 200—202, 282, 430
- —, свойство дивергентности 279, 280, 282, 283
- — сходящаяся 87, 89, 108, 171, 172, 176
- —, точность задания 145
- — установления 306, 307, 309, 395
- — устойчивая 106, 107, 124, 175, 176
- —, устойчивость относительно ошибок округления 429
- — Эйлера 105, 109, 157, 158
- — экономичная 289
- — явная 262, 266
- Разностное уравнение 15, 16
- — второго порядка 18, 25
- — —, фундаментальное решение 29
- —, дифференциальное приближение 237, 430
- —, скорость сходимости решения 75, 108
- Разностный аналог интегральных формул Коши и типа Коши 386
- Распад разрыва 273, 281
- Расщепление по физическим факторам 297
- Релаксационный метод Федоренко 323—326
- Рунге — Кутта варианты 374
- Рунге метод 335, 344
- Рунге — Кутта итерационный процесс 313, 314
- Рунге — Кутта схема 105, 157, 158, 162
- Ряд Фурье для сеточной функции 250
- Самарского критерий устойчивости разностных схем 394, 395
- Самосопряженность оператора 393

- Сетка 11, 16, 84, 278  
 Сеточная функция 38, 84, 86—88  
 — —, разложение в конечный ряд Фурье 252, 261  
 Скорость сходимости решения разностного уравнения 75, 108, 176  
 Сохоцкого — Племеля условие 419, 423  
 Спектр семейства операторов, алгоритм вычисления 403, 404, 411  
 — — —, определение 398  
 — — —, ядра 412  
 Спектральный критерий ограниченности степеней самосопряженного оператора 389  
 — признак Неймана устойчивости разностной задачи Коши 222, 388  
 — — ограниченности норм степеней оператора 387, 389  
 — — устойчивости разностной схемы 136, 138, 139, 221, 223, 229, 387—389, 398  
 Схема переменных направлений 288, 309, 320  
 Сходимость порядка  $h^k$  90, 108, 172  
 — разностной схемы 87, 89, 108, 121, 176
- Т**  
 Теорема о включении ядра спектра в спектр 413  
 — о связи между аппроксимацией, устойчивостью и сходимостью 108, 116, 429  
 — о совпадении ядра спектра со спектром 414, 432  
 — о структуре спектра семейства операторов 401  
 — об инвариантности ядра показателя 415  
 — об устойчивости возмущенной схемы 145  
 — эквивалентности Лакса 430  
 Теоремы о свойствах абсолютного ядра 413—414  
 Точка спектра оператора 309  
 — — семейства операторов 398, 400
- У**  
 Условие Куранта, Фридрихса и Леви 211, 212, 217, 219  
 — на линии разрыва решения 271  
 — Сохоцкого — Племеля 419, 429  
 Устойчивость алгоритмов 53, 67—70, 147, 148, 317—320, 415, 416, 417  
 — внутренних граничных условий 389  
 — нелинейных задач, прием исследования 155, 241  
 — разностной задачи Коши, необходимое и достаточное условие 233, 235  
 — — — относительно возмущения начальных данных 222  
 — — схемы 81, 106—108, 124, 175, 176, 387, 389, 390  
 — — —, количественная характеристика 149
- Ф**  
 Федоренко релаксационный метод 323—326  
 Фридрихса неравенство 334  
 Фундаментальное решение 22, 29, 30, 385  
 — — ограниченное 31—33  
 — —, оценка 30—35  
 — — разностного уравнения 22, 29, 420  
 — —, условие ограниченности 22  
 Фурье быстрое преобразование 431  
 — ряд для сеточной функции 250
- Х**  
 Характеристическое уравнение 26
- Ч**  
 Частная производная, замена разностными отношениями 183, 184
- Э**  
 Эволюционная задача 221, 222, 228, 386, 401, 402  
 Эйлера схема 195, 109, 157, 158  
 Энергетическое неравенство 255
- Я**  
 Ядра показателя спектра семейства операторов 413, 415, 433

*Сергей Константинович Годунов,  
Виктор Соломонович Рябенский*

**РАЗНОСТНЫЕ СХЕМЫ**

**Введение в теорию**

**М., 1977 г., 440 стр. с илл.**

**Редактор *Г. Я. Пирогова***

**Техн. редактор *Л. В. Лихачева***

**Корректор *Т. С. Вайсберг***

---

**Сдано в набор 20.12.76. Подписано к печати 28.04.77. Бумага 60×90<sup>1</sup>/<sub>16</sub>, тип. № 3. Физ. печ. л. 27,5. Условн. печ. л. 27,5. Уч.-изд. л. 25,25. Тираж 25000 экз. Цена книги 1 р. 14 к. Заказ № 425.**

---

**Издательство «Наука»  
Главная редакция физико-математической литературы  
117071, Москва, В-71, Ленинский проспект, 15**

---

**Ордена Трудового Красного Знамени  
Ленинградская типография № 2  
имени Евгении Соколовой Союзполиграфпрома  
при Государственном комитете Совета Министров СССР  
по делам издательств, полиграфии и книжной торговли,  
198052, Ленинград, Л-52, Измайловский проспект, 29,**