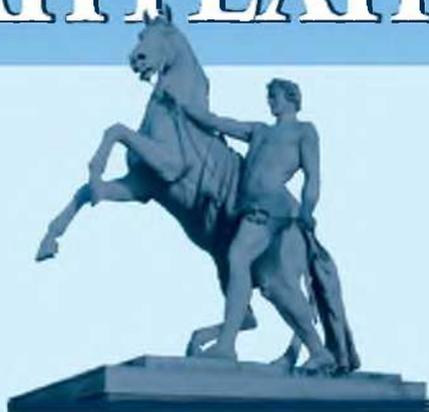


Ю. П. ПЕТРОВ, В. С. СИЗИКОВ

КОРРЕКТНЫЕ,
НЕКОРРЕКТНЫЕ
И
ПРОМЕЖУТОЧНЫЕ
ЗАДАЧИ
С
ПРИЛОЖЕНИЯМИ



ИЗДАТЕЛЬСТВО ПОЛИТЕХНИКА



Издательская и полиграфическая
деятельность с 1938 года

Основная тематика:
машиностроение
приборостроение
медицина

Издательство выпускает
монографии, справочники,
учебники, словари,
альбомы и журналы
в различных областях науки,
технологии и техники



ОАО «Издательство «ПОЛИТЕХНИКА»»

191023, Санкт-Петербург, Инженерная ул., д. 6.

Тел.: (812) 312-44-95, 710-62-73, тел./факс: (812) 312-57-68.

<http://www.polytechnics.ru> E-mail: gfm@polytechnics.ru, 710-62-73@polytechnics.ru



ОАО «Издательство «ПОЛИТЕХНИКА»» предлагает серию научно-технических журналов



«Металлообработка» (подписной индекс в агентстве «Роспечать» № 14250)

Для технологов предприятий, ученых НИИ и преподавателей вузов машиностроительной отрасли.

Тематика: новые технологии обработки металлов резанием, давлением, электрофизическими и электрохимическими методами; свойства материалов; заводской опыт.

Основан в 2001 г. **Входит в перечень ВАК.**
E-mail: mo@polytechnics.ru.



«Теплоэнергоэффективные технологии» (подписной индекс в агентстве «Роспечать» № 18341)

Для специалистов энергетической и строительной отраслей.

Тематика: производство, транспортировка, учет, распределение и потребление тепловой энергии; инженерная инфраструктура и аудит — статьи и нормативные документы.

Основан в 1995 г.
E-mail: teploenergo@polytechnics.ru.



«Биотехносфера» (подписной индекс в агентстве «Роспечать» № 45886)

Для специалистов медико-технического профиля.

Тематика: создание наукоемкой биомедицинской техники; исследования и разработки в области биоинженерии, биомедицинской информатики, бионанотехнологий, медико-технического менеджмента и образования.

Основан в 2009 г. **Входит в перечень ВАК.**
E-mail: bts@polytechnics.ru.

Издательство приглашает авторов и рекламодателей для размещения статей и рекламы по тематике журналов.

Все выпуски журналов можно приобрести в бумажном и электронном виде формата pdf на сайте издательства www.polytechnics.ru.

Электронный аналог печатного издания: Петров Ю. П., Сизиков В. С. Корректные, некорректные и промежуточные задачи с приложениями: учебное пособие для вузов. — 2-е изд., испр. и доп. — СПб.: Политехника, 2012. — 261 с.: ил.

УДК 517.983.54; 519.6

ББК 22.19; 22.161.6

ПЗ0



ПОЛИТЕХНИКА
ИЗДАТЕЛЬСТВО

Санкт-Петербург 2012

www.polytechnics.ru

*Рекомендовано УМО по оптическому
и приборостроительному образованию
в качестве учебного пособия при изучении дисциплин:
«Теория управления», «Обработка изображений»,
«Основы томографии», «Высшая математика»,
«Прикладная математика»*

Рецензент:

доктор физико-математических наук,
профессор Апарцин А. С. (г. Иркутск, СЭИ СО РАН)

Одобрено на заседаниях кафедры Моделирования электромеханических и компьютерных систем фак-та ПМ-ПУ СПбГУ и кафедры Измерительных технологий и компьютерной томографии фак-та ТМиТ СПбГИТМО(ТУ)

Петров, Ю. П., Сизиков, В. С.

ПЗ0

Корректные, некорректные и промежуточные задачи с приложениями: учебное пособие для вузов. — 2-е изд., испр. и доп. — СПб.: Политехника, 2012. — 261 с.: ил.

ISBN 5-7325-0761-2

Изложены понятия корректных и некорректных задач, а также задач, промежуточных между корректными и некорректными. Приведены примеры подобных математических задач: системы линейных алгебраических уравнений, системы обыкновенных дифференциальных уравнений, дифференциальные уравнения в частных производных, интегральные уравнения, а также примеры прикладных задач из теории управления, обработки изображений и томографии. Показано, что преобразования уравнений, эквивалентные в классическом смысле, могут переводить корректное уравнение в некорректное и наоборот. Введено понятие преобразований, эквивалентных в расширенном смысле. Изложены устойчивые методы регуляризации Тихонова и решения на компакте. Приведены результаты решения численных примеров. Данная книга может рассматриваться как учебное пособие (повышенной трудности), так и монография.

Для студентов, магистров, аспирантов, преподавателей и научных сотрудников в области фундаментальной и прикладной математики.

УДК 517.983.54; 519.6

ББК 22.19; 22.161.6

ISBN 5-7325-0761-2

© Издательство «Политехника», 2012

ОГЛАВЛЕНИЕ

СПИСОК СОКРАЩЕНИЙ	5
ПРЕДИСЛОВИЕ	6
ЧАСТЬ I. ТРИ КЛАССА ЗАДАЧ МАТЕМАТИКИ, ФИЗИКИ И ТЕХНИКИ	8
Глава 1. ПРОСТЕЙШИЕ НЕКОРРЕКТНЫЕ ЗАДАЧИ	8
1.1. Постановка проблемы. Примеры.....	8
1.2. Определения	12
1.3. Дальнейшие примеры и методы подхода к некорректным задачам.....	13
1.4. Некорректные задачи синтеза оптимальных систем управления.....	19
1.5. Некорректные задачи вычисления собственных значений систем линейных однородных уравнений	28
1.6. Решение систем дифференциальных уравнений. Всегда ли решения непрерывно зависят от параметров?.....	31
1.7. Заключение.....	39
Глава 2. КЛАСС ЗАДАЧ, ПРОМЕЖУТОЧНЫХ МЕЖДУ КОРРЕКТНЫМИ И НЕКОРРЕКТНЫМИ	41
2.1. Обнаружение третьего класса задач математики, физики и техники и его значение	41
2.2. Преобразования, эквивалентные в классическом смысле.....	42
2.3. Обнаружившиеся парадоксы	44
2.4. Преобразования, эквивалентные в расширенном смысле.....	46
2.5. Задачи, промежуточные между корректными и некорректными.....	49
2.6. Приложения к системам управления и другим объектам, описываемым дифференциальными уравнениями	54
2.7. Приложения к практике вычислений.....	62
2.8. Заключение по главам 1 и 2	67
Глава 3. ИЗМЕНЕНИЕ ЧУВСТВИТЕЛЬНОСТИ К ПОГРЕШНОСТЯМ ИЗМЕРЕНИЯ ПРИ ИНТЕГРАЛЬНЫХ ПРЕОБРАЗОВАНИЯХ, ИСПОЛЬЗУЕМЫХ ПРИ РАСЧЕТЕ СУДОВ И СИСТЕМ СУДОВОЙ АВТОМАТИКИ	70
3.1. Применение интегральных преобразований для решения практических задач.....	70
3.2. Свойства корреляционных функций	74
3.3. Свойства спектров	79
3.4. Корректность интегральных преобразований	83
3.5. Задачи, мало чувствительные к погрешностям спектров.....	87
3.6. Дифференцирование функций, отягощенных помехами	94
3.7. Предсказание будущего	100
ЛИТЕРАТУРА к Части I	111

ЧАСТЬ II. УСТОЙЧИВЫЕ МЕТОДЫ РЕШЕНИЯ ОБРАТНЫХ ЗАДАЧ..... 114

Глава 4. РЕГУЛЯРНЫЕ МЕТОДЫ РЕШЕНИЯ НЕКОРРЕКТНЫХ ЗАДАЧ

.....	116
4.1. Элементы функционального анализа.....	116
4.2. Некоторые сведения из линейной алгебры.....	122
4.3. Основные типы уравнений и преобразований.....	129
4.4. Корректность и некорректность по Адамару.....	137
4.5. Классические методы решения интегральных уравнений Фредгольма	
I рода.....	148
4.6. Методы наименьших квадратов Гаусса и псевдообратной матрицы	
Мура-Пенроуза.....	152
4.7. Метод регуляризации Тихонова.....	159
4.8. Метод решения на компакте.....	178

Глава 5. ОБРАТНЫЕ ЗАДАЧИ РЕКОНСТРУКЦИИ ИЗОБРАЖЕНИЙ И ТОМОГРАФИИ.....188

5.1. Реконструкция смазанных изображений.....	188
5.2. Реконструкция дефокусированных изображений.....	198
5.3. Задача рентгеновской томографии.....	204
5.4. Задача синтеза магнитного поля в ЯМР-томографе.....	212

ЛИТЕРАТУРА к Части II.....224

Приложение. О «ГРАММАТИКЕ» НАУКИ.....227

ЛИТЕРАТУРА к «ПРИЛОЖЕНИЮ».....256

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ.....257

СПИСОК СОКРАЩЕНИЙ

- БПФ — быстрое преобразование Фурье
- ДПФ — дискретное преобразование Фурье
- КТХ — космический телескоп «Хаббл»
- МНК — метод наименьших квадратов (Гаусса)
- МПОМ — метод псевдообратной матрицы (Мура–Пенроуза)
- НПФ — непрерывное преобразование Фурье
- ОДУ — обыкновенное дифференциальное уравнение
- ОПФ — обратное преобразование Фурье
- ПФ — преобразование Фурье
- ПХ — преобразование Хартли
- РА — регуляризирующий алгоритм
- РО — регуляризирующий оператор
- РТ — рентгеновская томография
- СЛАУ — система линейных алгебраических уравнений
- ЧХ — частотная характеристика (спектрометра)
- ЯМР — ядерный магнитный резонанс

ПРЕДИСЛОВИЕ

Настоящая книга является учебным пособием для студентов, аспирантов и преподавателей вузов, а также монографией для научных сотрудников в области фундаментальной и прикладной математики, имеющих дело с решением алгебраических, дифференциальных, интегральных и операторных уравнений. Более конкретно, книга посвящена одной из важнейших проблем прикладной математики – исследованию и обеспечению устойчивости решений уравнений с учетом погрешностей исходных данных, параметров и коэффициентов математической модели исследуемого объекта, аппаратной функции, начальных условий и т.д., а также с учетом погрешностей вычислений, например, погрешностей округлений.

До последнего времени считалось, что все задачи математики, физики и техники достаточно поделить на два класса – *задачи корректные*, в которых малым погрешностям исходных данных соответствуют малые погрешности решений, и *задачи некорректные*, в которых малым погрешностям исходных данных соответствуют сколь угодно большие погрешности решений. Некорректные задачи требуют для своего решения специальных (устойчивых, регулярных) методов, часть которых рассмотрена в настоящей книге. Однако целесообразно ввести также *третий класс задач*, которые можно назвать *промежуточными между корректными и некорректными*. Это – задачи, меняющие свою корректность в результате эквивалентных преобразований, а также задачи, являющиеся корректными или некорректными в зависимости от типов используемых функциональных пространств. С точки зрения приложений, задачи, меняющие свою корректность в результате эквивалентных преобразований, требуют к себе особо пристального внимания, так как их решение классическими методами может привести к грубым ошибкам в расчетах и тем самым стать причиной аварий и даже катастроф.

Книга состоит из двух частей, дополняющих друг друга.

В первой части рассматриваются общие свойства всех трех классов задач математики, физики и техники, а также подходы к их решению.

Во второй части изложен ряд устойчивых методов решения обратных некорректных (а также плохо обусловленных) задач, проиллюстрированный численными примерами.

Текст и формулы данной книги набраны в редакторе Word97 (и Word2000) в виде структуры с главным документом (управляющим файлом) и вложенными документами (файлами списка сокращений,

предисловия, частей, глав, литературы, приложения и предметного указателя). Что касается рисунков, то для их построения были использованы графические редакторы Grapher, PBbrush, Word, Visual C++, MathCAD, PhotoShop, CorelDraw, а также FineReader, причем для построения рисунков 4.2–4.12, 5.18–5.20 были предварительно выполнены расчеты на языке MS Fortran 5 (и Fortran 90), а для построения рисунков 5.2–5.5, 5.7 и 5.8 – расчеты на Visual C++.

Разработан также электронный вариант данной книги в виде файла pet&siz.pdf (с гипертекстовыми ссылками и другим сервисом) и помещен в Internet.

Авторы благодарят рецензента проф. А.С. Апарцина за просмотр рукописи и ценные замечания. Авторы благодарят И.А. Белова за помощь при оформлении структуры книги и за расчеты рисунков 5.2–5.5, 5.7, 5.8.

Первая часть книги и приложение написаны д.т.н., проф. СПбГУ Петровым Ю. П. (<http://www.petrov1930.narod.ru>, petrov1930@mail.ru), а вторая часть – д.т.н., проф. СПбГИТМО(ТУ) Сизиковым В. С. (<http://de.ifmo.ru/curs/svs.html>, sizikov2000@mail.ru).

Все пожелания и замечания по поводу данного учебного пособия–монографии авторы просят направлять по адресам: 190005, С-Петербург, ул. 2-я Красноармейская, 15, кв. 11, Петрову Ю. П.; 197183 С-Петербург, ул. Сестрорецкая, 9, кв. 8, Сизикову В. С.; а также по e-mail или по факсу: 812-2335952.

Часть I

ТРИ КЛАССА ЗАДАЧ МАТЕМАТИКИ, ФИЗИКИ И ТЕХНИКИ

Глава 1

ПРОСТЕЙШИЕ НЕКОРРЕКТНЫЕ ЗАДАЧИ

1.1. Постановка проблемы. Примеры

Необходимость исследования некорректных задач связана с одной из основных проблем прикладной математики – обеспечения надежных результатов вычислений при учете неизбежных погрешностей в задании коэффициентов и параметров математической модели, по которой производятся расчеты.

Действительно, коэффициенты математической модели, коэффициенты уравнений или системы уравнений, по которым производятся расчеты, получены, как правило, на основе измерений и поэтому имеют ограниченную точность. Кроме того, параметры реального процесса или технического объекта, который мы рассчитываем, не остаются идеально постоянными, они испытывают малые неконтролируемые изменения, вариации, точная величина которых обычно неизвестна.

Поэтому мы будем различать номинальные значения коэффициентов математической модели (будем обозначать их a_{in}) и реальные, истинные значения (обозначим их a_{ip}). Номинальные значения – это те, которые вводятся в вычислительную машину, участвуют во всех расчетах, а истинные, реальные значения a_{ip} неизвестны нам. Мы можем лишь утверждать, что эти неизвестные значения заключены в некоторых пределах, удовлетворяя неравенствам:

$$(1 - \varepsilon) a_{in} \leq a_{ip} \leq a_{in} (1 + \varepsilon), \quad (1.1)$$

где ε – числа, малые в сравнении с единицей. Их точные значения нам неизвестны; известны лишь их оценки.

Произведения $\pm \varepsilon \cdot a_{in}$ будем называть *вариациями* (или погрешностями) номинальных коэффициентов.

Поскольку все расчеты ведутся по номинальным коэффициентам (а реальные, истинные коэффициенты, как уже говорилось, нам неизвестны), то необходимо исследовать, как влияют вариации коэффициентов (а также вариации параметров, начальных условий, граничных

условий и т.д.) на точность наших расчетов.

Существуют задачи, в которых погрешности решений имеют тот же порядок, что и погрешности коэффициентов: это – наиболее легкий случай. Существуют задачи, где погрешности решений больше погрешности коэффициентов. Здесь нужно особенно тщательно следить за точностью, искать методы решения, уменьшающие погрешности.

Наконец, существуют задачи, где даже при сколь угодно малых, неизбежных на практике, погрешностях коэффициентов, параметров, начальных или граничных условий погрешности решения велики.

Такие задачи называют *некорректными* (точное определение дадим позже) и они особенно трудны для решения. Однако подобные задачи встречаются часто и методы подхода к ним необходимо изучать.

Начнем с простых примеров.

Рассмотрим систему двух уравнений с двумя переменными x и y :

$$\begin{cases} a_{11}x + a_{12}y = b_1, \\ a_{21}x + a_{22}y = b_2. \end{cases} \quad (1.2)$$

Каждое уравнение является уравнением прямой на плоскости (x, y); решение системы (1.2), т.е. значения x и y , обращающие уравнения в тождества, являются координатами точки пересечения этих прямых. Согласно известным формулам Крамера, решения x, y выражаются через определители:

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \quad (1.3)$$

Так, для системы уравнений

$$\begin{cases} x + 2y = 3, \\ 2x + y = 3 \end{cases} \quad (1.4)$$

определитель системы равен $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = 3$, определитель

$\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = \begin{vmatrix} 3 & 2 \\ 3 & 1 \end{vmatrix} = -3$, определитель $\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix} = -3$. Система

имеет решение $x=1, y=1$.

Исследуем влияние вариации коэффициента при x в первом из уравнений (1.4) на погрешность решения.

Для системы

$$\begin{cases} (1+\varepsilon)x + 2y = 3, \\ 2x + y = 3 \end{cases} \quad (1.5)$$

решением будут числа $x = \frac{3}{3-\varepsilon}$, $y = \frac{3-3\varepsilon}{3-\varepsilon}$.

Если $|\varepsilon| \leq 0.01$, то

$$\begin{aligned} 0.996 \leq x \leq 1.0034, \\ 0.986 \leq y \leq 1.0034. \end{aligned} \quad (1.6)$$

Мы убедились, что в данном случае малые погрешности в коэффициенте привели к малым погрешностям в решении. Однако существуют системы, в которых относительная погрешность решения существенно больше относительной погрешности коэффициентов. Такие системы называют часто «плохо обусловленными» системами (точное определение мы дадим позже).

Так, например, для системы

$$\begin{cases} 1.1x + y = 1.1, \\ (1+\varepsilon)x + y = 1 \end{cases} \quad (1.7)$$

определитель равен

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 1.1 & 1 \\ 1+\varepsilon & 1 \end{vmatrix} = 0.1 - \varepsilon.$$

Система (1.7) имеет решение $x = \frac{1}{10-\varepsilon}$, $y = -\frac{11\varepsilon}{1-10\varepsilon}$ и

если $|\varepsilon| \leq 0.001$, то $0.99 \leq x \leq 1.01$,

если $|\varepsilon| \leq 0.01$, то $0.909 \leq x \leq 1.11$,

если $|\varepsilon| \leq 0.1$, то $0.5 \leq x \leq \infty$.

В данном примере погрешность решений x и y больше погрешности коэффициентов. Кроме того, она быстро возрастает с ростом ε и если $|\varepsilon| \leq 0.1$, то погрешность уже может быть сколь угодно велика.

В случае с системой (1.2) все достаточно очевидно: погрешности решений велики тогда, когда мал определитель

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}. \quad (1.8)$$

Малость определителя (1.8) говорит о том, что прямые пересекаются друг с другом под очень малым углом и поэтому даже небольшая вариация коэффициента, изменяющая этот угол, приводит к большому изменению координат точки пересечения.

Рассмотрим теперь предельный случай, когда определитель (1.8)

при номинальных значениях коэффициентов равен нулю. В этом случае уже сколь угодно малые вариации коэффициентов могут привести к большим и даже к коренным изменениям решений.

Рассмотрим систему

$$\begin{cases} x + y = b_1, \\ x + y = 1, \end{cases} \quad (1.9)$$

для которой

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0.$$

Если $b_1 \neq 1$, то решений не существует (действительно, в этом случае прямые параллельны). Если же $b_1 = 1$, то решений много: в этом случае прямые совпадают и любая пара чисел $x = 1 - y$ является решением.

Рассмотрим теперь систему

$$\begin{cases} (1 + \varepsilon)x + y = b_1, \\ x + y = 1, \end{cases} \quad (1.10)$$

т.е. рассмотрим влияние на решение системы (1.9) вариации коэффициента a_{11} . Вычитая из первого уравнения второе, находим

$$x = \frac{b_1 - 1}{\varepsilon}, \quad y = 1 - \frac{b_1 - 1}{\varepsilon}. \quad (1.11)$$

Если $b_1 \neq 1$, то теперь для любого ε (даже сколь угодно малого) решение существует, но оно целиком зависит от неизвестной нам погрешности ε . Практического смысла это решение не имеет, хотя оно существует для любого $\varepsilon \neq 0$. Если же $b_1 = 1$, то в этом случае, также вычитая второе из уравнений (1.10) из первого, получим $\varepsilon \cdot x = 0$, откуда вытекает решение

$$x = 0, \quad y = 1, \quad (1.12)$$

справедливое для всех ε (в том числе и сколь угодно малых ε), кроме единственного значения $\varepsilon = 0$, которому соответствуют бесконечное множество решений (любая пара $x = 1 - y$ при $\varepsilon = 0$ является решением). Будет решение единственным, или решений будет много – зависит от неизвестного нам значения ε .

Поэтому решения системы (1.2) при определителе (1.8), равном нулю, сами по себе практического смысла не имеют.

Для вырожденных систем существует метод псевдообратной матрицы Мура-Пенроуза [24, с.189], дающий единственное (но, вообще говоря, неустойчивое) решение, называемое нормальным решением.

1.2. Определения

Рассмотрев простейший пример, мы должны теперь определить, какие задачи считать корректными и какие – некорректными (или – как их еще называют – некорректно поставленными задачами).

Существует определение, приведенное в наиболее авторитетной книге о некорректных задачах – Тихонов А.Н., Арсенин В.Я. «Методы решения некорректных задач» М., Наука, третье издание, 1986. (Первое издание было опубликовано в 1974 г., второе – в 1979.) Согласно [20], задача называется некорректной, если она не удовлетворяет хотя бы одному из трех условий:

1. Решение существует.

2. Решение единственно.

3. Сколь угодно малым вариациям коэффициентов, параметров, начальных или граничных условий соответствуют сколь угодно малые изменения решений.

Это определение не идеально (вряд ли все многочисленные задачи, не имеющие решений, т.е. не удовлетворяющие условию 1, целесообразно относить к некорректным), однако оно общепринято и мы будем его придерживаться.

Отметим, что из определения вариаций коэффициентов и параметров через неравенства (1.1) следует, что если номинальное значение коэффициента, т.е. значение a_n , равно нулю, то и вариация его равна нулю, т.е. «нуль не варьируется», нуль остается нулем. Это означает, что при анализе корректности мы рассматриваем только относительные вариации коэффициентов, а не абсолютные. Мы исключаем из рассмотрения те случаи, когда номинальное значение коэффициента равно нулю, а проварьированное значение стало сколь угодно малым, но не равным нулю. Такое ограничение необходимо, иначе очень широкий, почти необъятный круг задач незаслуженно попадет в некорректные. Так, например, уравнение $a_1x + a_0 = 0$ можно, естественно, рассматривать как уравнение $0 \cdot x^2 + a_1x + a_0 = 0$, т.е. как уравнение $a_2x^2 + a_1x + a_0 = 0$, в котором коэффициент $a_2 = 0$. Если после вариации он примет значение $a_2 = \varepsilon$, то уже при сколь угодно малых ε уравнение будет иметь два корня, не близких друг к другу, и мы задачу нахождения корней полинома $a_1x + a_0 = 0$ и вообще любого полинома

$$a_nx^n + a_{n-1}x^{n-1} + \dots + a_0 = 0$$

должны были бы отнести к некорректным задачам, что совершенно нецелесообразно.

Отказавшись от рассмотрения «вариаций нуля», мы выводим из

круга некорректных задач многочисленные так называемые «сингулярно-возмущенные» системы дифференциальных уравнений, т.е. системы с малыми параметрами при старших производных – типа уравнения

$$\varepsilon \ddot{x} + a_1 \dot{x} + a_0 = 0$$

и ему подобных. Это – интересные задачи, но к некорректным они не относятся. Если даже ε – сколь угодно малая величина, то она (разумеется, только условно говоря) «в бесконечно большое число раз» больше, чем точный нуль. Поэтому переход от точного нуля к $\varepsilon \neq 0$ не следует считать вариацией нулевого коэффициента, малым изменением его.

Мы убедимся далее, что и при отказе от «вариаций нуля» некорректных задач все равно остается очень много.

1.3. Дальнейшие примеры и методы подхода к некорректным задачам

К некорректным задачам относится очень широкий круг задач на максимумы и минимумы (экстремальные задачи).

Пример. Какова минимальная длина изгороди, необходимая для плотного, без щелей, огораживания участка земли площадью s ?

На первый взгляд задача совсем проста: известно, что минимум длины изгороди достигается в том случае, если огораживаемый участок имеет форму круга с периметром $p = 2\pi R$ (где R – радиус круга) и площадью $s = \pi R^2$. Исключая R , получим $p_{\min} = \sqrt{4\pi} \cdot \sqrt{s}$.

Однако площадь s не может быть измерена идеально точно. Если истинная площадь больше номинальной s_n хотя бы на малую величину Δs , то изгородь длиной $p_{\min} = \sqrt{4\pi} \cdot \sqrt{s_n}$ не удастся замкнуть, задача огораживания решена не будет. Мы убеждаемся, что рассматриваемая задача, как и впрочем любая задача об элементе, доставляющем минимум – является задачей некорректной. При сколь угодно малой погрешности в условии решение может исчезнуть.

На этом простом примере удобно разъяснить общий подход к решению некорректных задач: задачу некорректную, не имеющую поэтому практического смысла, надо заменить на близкую к ней задачу корректную.

В данном случае корректная задача такова: если площадь земельного участка s_n известна с погрешностью, не превышающей Δs , то какой запас длины изгороди Δp следует прибавить к минимальной длине p_{\min} для того, чтобы задача огораживания была всегда выпол-

нима? Из условия $p_{\min} + \Delta p = \sqrt{4\pi} \cdot \sqrt{s_n + \Delta s}$ сразу находим $\Delta p = \sqrt{4\pi} \cdot (\sqrt{s_n + \Delta s} - \sqrt{s_n})$.

Поскольку погрешность Δs может быть различной, равной $\Delta_i s$, то полученная формула дает нам для каждого $\Delta_i s$ последовательность корректных задач, имеющих смысл. В пределе, при $\Delta s \rightarrow 0$ получаем предельное, уже не имеющее практического смысла, решение исходной некорректной задачи.

Рассмотренный простой пример дает понятие о простейшем случае так называемой регуляризации некорректной задачи: исходную некорректную задачу, не имеющую практического смысла, мы заменяем на последовательность корректных задач с некоторым параметром, предельное значение которого соответствует исходной некорректной задаче. В данном простейшем случае этот параметр равен Δs . В дальнейшем будут рассмотрены другие методы регуляризации.

Так же, как и в рассмотренном примере, некорректность будет возникать в любой задаче на экстремум, точнее, в задаче о вычислении того элемента, на котором будет достигаться минимум некоторой функции, функционала и т.п. Сколь угодно малые изменения в условии приведут к тому, что найденного минимального элемента (в нашем примере – длины изгороди $p_{\min} = \sqrt{4\pi} \cdot \sqrt{s_n}$) уже не хватит. Задача потеряет смысл, что и свидетельствует о ее некорректности.

Однако, поскольку для всех задач об элементе, доставляющем экстремум, причина некорректности ясна и очевидна, то их издавна решали «с запасом», не оговаривая специально, что, например, в задаче о длине изгороди нужно за искомое условие брать не номинальное значение площади s_n , а максимально возможное, т.е. брать $s = s_n + \Delta s$, где Δs – максимально возможная погрешность в измерении и задании s . Ввиду простоты подхода к решению, экстремальные задачи издавна успешно решали без выделения их в особый класс некорректных задач. Этот класс был выделен французским математиком Жаком Адамаром (1865 – 1963) в 1902 г. Адамар привел пример задачи теплопроводности, в которой сколь угодно малая ошибка в задании граничного условия приводит к большой ошибке в решении. Однако пример Адамара относился к довольно сложной математической модели, описываемой дифференциальными уравнениями с частными производными. Достаточно сложны были и примеры, рассматриваемые позже в известном учебнике [20], книгах [6], [10], [21] и т.п.

Поэтому у студентов, инженеров и научных работников – у всех лиц, ведущих математические расчеты, часто возникает мнение, что

некорректные задачи – это нечто, относящееся к очень «высокой» математике, относящееся к сложным и редко встречающимся задачам.

На самом деле это не так. Некорректные задачи встречаются часто, на каждом шагу, и о свойствах некорректных задач, об опасности ошибок, с ними связанных, надо всегда помнить. Развитие представлений о некорректных задачах изложено в книге [18].

Мы уже упоминали, что некорректными являются все те многочисленные задачи на экстремум, в которых нужно найти элемент, реализующий экстремум.

К некорректным относятся и задачи на вычисление корней полиномов, если эти корни кратные, а физический смысл имеют лишь вещественные решения.

Рассмотрим простейший полином второй степени:

$$x^2 + 2x + 1.$$

Любой студент скажет, что этот полином имеет двухкратный корень $x_1 = x_2 = -1$ и вычислить его можно по элементарной формуле:

$$x_{1,2} = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - c} = -1 \pm \sqrt{1-1} = -1.$$

Однако если коэффициент при последнем члене полинома равен не точно единице, а $1 + \varepsilon$, что всегда возможно, поскольку все коэффициенты известны лишь с ограниченной точностью, то вещественное решение при сколь угодно малых $\varepsilon > 0$ сразу исчезает, в этом случае

$$x_{1,2} = -1 \pm \sqrt{1 - (1 + \varepsilon)} = -1 \pm \sqrt{-\varepsilon}.$$

Если нас интересуют только вещественные решения, то уже при сколь угодно малых $\varepsilon > 0$ решение исчезает. Задача вычисления вещественных кратных корней некорректна.

Некорректны многие краевые задачи для обыкновенных дифференциальных уравнений. Рассмотрим уравнение

$$\frac{d^2 x}{dt^2} - x = 0 \tag{1.13}$$

с краевыми условиями: $x(t=0) = 0$, $x(t=a) = b$. Уравнение (3.13) имеет общее решение

$$x = c_1 \sin t + c_2 \cos t,$$

где c_1 и c_2 – постоянные интегрирования (действительно, функции $c_1 \sin t$ и $c_2 \cos t$, как легко проверить, удовлетворяют уравнению (3.13); поэтому удовлетворяет ему и их линейная комбинация). Из первого краевого условия находим, что $c_2 = 0$. Из второго краевого усло-

вия $x(t=a)=b$ находим, что $c_1 = \frac{b}{\sin a}$. Таким образом, заданным краевым условиям удовлетворяет решение

$$x_1(t) = \frac{b}{\sin a} \cdot \sin t.$$

Однако, если краевое условие относится не к точке $t = a$, а к точке $t = a + \epsilon$ (что вполне возможно, поскольку малые погрешности в краевых условиях неизбежны), то на самом деле решение имеет вид

$$x_2(t) = \frac{b}{\sin(a + \epsilon)} \cdot \sin t.$$

Если, например, $a = \pi - \epsilon$, то модуль разности между $x_1(t)$ и $x_2(t)$ может быть сколь угодно велик даже для сколь угодно малых ϵ .

Таким образом, мы убеждаемся, что некорректных задач очень много.

Отметим, что в отличие от краевой задачи, задача Коши для одного дифференциального уравнения любого порядка корректна. Однако в системах дифференциальных уравнений недавно открыты интересные явления изменения корректности, о которых мы расскажем в главе второй.

Подходы к решению некорректных задач покажем на примере задачи определения места корабля по пеленгам. Если берег виден, то судоводитель измеряет углы между направлением на север и направлениями на два маяка А и В, отмеченные на карте. Затем он прочерчивает карандашом на карте эти направления (пеленги). Место корабля находится на пересечении двух пеленгов (рис. 1.1). Таким образом, судоводитель решает графически, решает построением, систему двух линейных уравнений – типа уравнений (1.2).

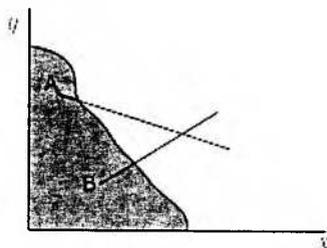


Рис. 1.1

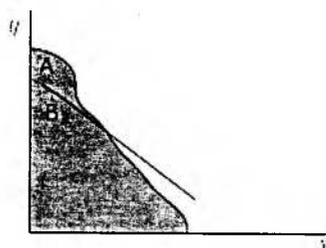


Рис. 1.2

Разумеется, углы измеряются с конечной точностью, проведение линии на карте тоже не обходится без погрешности, поэтому место корабля определяется приближенно, но обычно с достаточной для су-

доводителя точностью.

Если два маяка расположены близко друг от друга, то направления на них близки, карандашные линии будут пересекаться под очень острым углом, а могут и вообще в пределах точности чертежа слиться в одну прямую, как это показано на рис. 1.2.

Рассмотрим именно этот случай – случай, когда судоводителю приходится решать некорректную задачу, например, задачу решения системы уравнений

$$\begin{cases} x + y = 1, \\ x + y = 1, \end{cases} \quad (1.14)$$

определитель которой равен нулю и которая, поэтому, фактически сводится к одному уравнению

$$x + y = 1.$$

Место корабля по двум совпадающим линиям (с уравнениями $x + y = 1$ и $x + y = 1$), естественно, не определить – оно может быть любым на линии $x + y = 1$. Этого и следовало ожидать, поскольку сама по себе некорректная задача как таковая практического смысла не имеет.

Однако некорректную задачу можно различными способами регуляризовать.

Вот один из способов (кстати, довольно широко рекомендуемый): ищут не любое решение системы (1.14), а решение с наименьшей нормой. (Этот метод, дающий решение с минимальной нормой; называемое еще нормальным решением, рассмотрен более подробно в [24], стр. 179 и 189). Одна из возможных норм – это сумма квадратов переменных x и y . Приходим к задаче: найти значения x и y , доставляющие минимум норме:

$$F = x^2 + y^2$$

при условии, что x и y связаны уравнениями (1.14), или (что то же самое) связаны уравнением $x + y = 1$.

Учитывая, что из этого уравнения следует $x = 1 - y$, приведем норму к виду $F = (1 - y)^2 + y^2$ и взяв производную $\frac{\partial F}{\partial y}$ сразу устанавливаем, что минимум нормы будет достигаться в точке $x = 0.5$, $y = 0.5$.

Таким образом, введя норму $F = x^2 + y^2$ мы перешли от некорректной задачи к корректной (нетрудно проверить, что при малых вариациях коэффициентов системы (1.14) решение $x = 0.5$, $y = 0.5$ изменится мало).

Однако для интересующей нас практической задачи определения

места корабля решение $x = 0.5, y = 0.5$ ничего не дает. К месту корабля это решение никакого отношения не имеет. Поэтому не всякий метод регуляризации всегда пригоден для решения конкретной практической задачи. Надо искать более подходящий метод.

Один из таких методов – использование дополнительной информации. Пусть два маяка расположены так близко, что пеленги на них сливаются на карте в одну линию, уравнения пеленгующих линий принимают вид (1.14) и поэтому задача нахождения точки пересечения некорректна (рис. 1.2), но есть возможность оценить расстояния до маяков, которые и являются дополнительной информацией. Тогда мы на карте циркулем откладываем это расстояние и получаем на прямой $x + y = 1$ истинное положение корабля.

Еще один метод использования дополнительной информации – взять, если можно, пеленг на третий маяк, расположенный в стороне от первых двух. Точка пересечения пеленга на этот маяк с ранее полученной прямой $x + y = 1$ даст место корабля (рис 1.1).

Таким образом, для задачи определения места корабля в тех случаях, когда она оказывается некорректной, хорошим методом регуляризации оказывается использование дополнительной информации.

Методов регуляризации много, но подобрать подходящий метод для встретившейся вам конкретной задачи не очень легко.

Вообще на «регуляризацию» не следует смотреть как на «волшебный ключик», доставляющий решение там, где ранее решение не имело смысла. Решение самой по себе некорректной задачи, как уже говорилось, практического смысла не имеет. Регуляризацию следует рассматривать скорее не как метод решения, а как метод подхода к некорректным задачам, когда мы некорректную задачу заменяем близкой к ней корректной задачей – путем использования дополнительной информации, введения нормы и поиска решения с минимальной нормой, введения последовательности корректных задач, приближающихся в пределе к некорректной и т.п.

Различных подходов к некорректным задачам много и мы рассмотрим далее дополнительные примеры.

Главное же, о чем нужно всегда помнить – это то, что перед решением любой задачи нужно проверить, корректна задача или нет. Если некорректную задачу решать обычными методами, как задачу корректную, то почти наверняка придем к грубой ошибке.

Разумеется, это не означает, что каждую встретившуюся задачу надо проверять индивидуально. Можно основываться на общих теоремах. Так, давно доказано, что задача вычисления корней полинома любой степени вида

$$a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0$$

в поле комплексных чисел корректна. Однако надо помнить, что некорректных задач много и для многих из них общих теорем нет. Мы увидим далее, что, например, такая распространенная задача как решение систем обыкновенных дифференциальных уравнений, в общем случае некорректна.

1.4. Некорректные задачи синтеза оптимальных систем управления

Наиболее примечательная встреча с некорректными задачами произошла в 70-х годах XX века при синтезе оптимальных систем управления. Незадолго до этого (см., например, публикации [7, 8, 9]) была разработана теория и предложены алгоритмы синтеза оптимальных систем управления, которые позволяли существенно повысить эффективность работы многих объектов промышленности и транспорта, испытывающих влияние неполностью известных возмущающих сил.

Рассматривались объекты управления, описываемые следующими дифференциальными уравнениями с постоянными коэффициентами

$$A(D)x = B(D)u + \varphi(t). \quad (1.15)$$

В уравнении (1.15) через $A(D) = a_n D^n + a_{n-1} D^{n-1} + \dots + a_0$ и $B(D) = b_m D^m + b_{m-1} D^{m-1} + \dots + b_0$ обозначены полиномы от оператора дифференцирования $D = \frac{d}{dt}$, x – регулируемая величина (точнее, ее

отклонение от желаемого значения), u – управляющее воздействие, $\varphi(t)$ – возмущающее воздействие, неполностью известная нам стационарная случайная функция времени, относительно которой нам известна, как правило, только ее спектральная плотность мощности (называемая обычно коротко спектром). Спектр случайной функции вычисляется несложной обработкой наблюдений за возмущающими воздействиями и затем аппроксимируется дробно-рациональной четной функцией:

$$S = \frac{a_p \omega^{2p} + a_{p-1} \omega^{2p-2} + \dots + a_0}{b_q \omega^{2q} + b_{q-1} \omega^{2q-2} + \dots + b_0}, \quad (1.16)$$

где ω – переменная, имеющая размерность частоты, т.е. $\frac{1}{\text{сек}}$.

Критерием качества управления являлся функционал вида

$$J = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (\lambda_0 x^2 + u^2) dt = \lambda_0 \langle x^2 \rangle + \langle u^2 \rangle, \quad (1.17)$$

где член $\langle x^2 \rangle$ отражал точность регулирования, а член $\langle u^2 \rangle$ косвенно учитывал неизбежные ограничения на величину управляющего воздействия. Неотрицательное число λ_0 являлось множителем Лагранжа, которое зависело от ограничений на управление и вычислялось по правилам, приведенным в публикациях [7, 8, 9, 13].

Еще в 50-х годах XX века было доказано, что существует легко реализуемое линейное управление вида

$$W_1(D)x = W_2(D)u, \quad (1.18)$$

(где $W_1(D)$ и $W_2(D)$ – полиномы от оператора дифференцирования $D = \frac{d}{dt}$), которое обеспечивает устойчивость замкнутой системы (т.е.

устойчивость решений системы уравнений (1.15)–(1.18) рассматриваемых совместно) и обеспечивает минимум критерия качества (1.17). Задача оптимизации сводилась тем самым к задаче поиска оптимальных операторов $W_1(D)$ и $W_2(D)$ и их последующей реализации с помощью технических устройств, называемых регуляторами.

Пример. Для танкеров типа «Казбек» (водоизмещение – 16000 тонн, скорость – 14 узлов) математическая модель движения по курсу под действием руля имела вид:

$$(T_1^2 D^2 + T_2 D)x = u + \varphi(t), \quad (1.19)$$

(т.е. $A(D) = T_1^2 D^2 + T_2 D$; $B(D) = 1$), где T_1 и T_2 – постоянные времена в секундах (для танкеров типа «Казбек» $T_1 = 26.3$ сек, $T_2 = 17.3$ сек), x – отклонение судна от заданного курса в градусах, u – управляющее воздействие, отклонение руля от диаметральной плоскости в градусах, $\varphi(t)$ – возмущающее воздействие на корабль, момент сил ветра и морского волнения, измеряемое в градусах отклонения руля, создающего момент той же величины.

Для танкера «Казбек» минимум потери скорости обеспечивал легко реализуемый регулятор вида (1.18):

$$u = -(2.5 + 43.6D)x \quad (1.20)$$

(т.е. $W_1(D) = -2.5 + 43.6D$; $W_2(D) = 1$).

Регулятор (1.20) является параллельным соединением усилительного и дифференцирующего звена и поэтому легко реализуем.

Для других объектов управления также удавалось построить регуляторы (1.18) и эти регуляторы обеспечивали очень существенное улучшение качества работы этих объектов по сравнению с теми регу-

ляторами, которые стояли на этих объектах ранее, до появления теории оптимального управления.

Коэффициенты оптимальных полиномов $W_1(D)$ и $W_2(D)$ в регуляторах (1.18) вычислялись по достаточно сложным алгоритмам. Приведем для примера один из самых простых алгоритмов, пригодный для частного случая $B(D) = 1$. Алгоритм состоит из следующих шагов:

1. После замены в спектре (1.16) переменной $j\omega$ на переменную s : $j\omega = s$ спектр факторизуется

$$S(s) = S_1(s) \cdot S_1(-s), \quad (1.21)$$

т.е. представляется в виде произведения двух симметричных множителей $S_1(s)$ и $S_2(s)$, в первый из которых входят все множители вида $s - a_i$, где a_i – корни с отрицательными вещественными частями, а во второй входят все множители того же вида, но с положительными вещественными частями корней.

Поскольку спектр (1.16) является четной функцией, то его факторизация (т.е. разложение на симметричные множители) всегда осуществима, но требует вычисления корней числителя и знаменателя спектра (1.16), что предполагает, разумеется, использование вычислительной техники.

2. На втором шаге факторизуется четный полином:

$$A(s)A(-s) + \lambda_0 = G(s) \cdot G(-s). \quad (1.22)$$

3. На третьем шаге выполняется разложение на дроби (сепарация):

$$\frac{A(-s)}{G(-s)} \cdot S_1(s) = M_0 + M_+ + M_-, \quad (1.23)$$

где M_0 – целый полином, M_+ и M_- – правильные дроби с полюсами, соответственно, в левой и правой полуплоскости комплексного переменного s . Операция требует, естественно, применения вычислительной техники и за вычетом простейших полиномов $A(s)$ результат операции не может быть выписан в явном виде.

4. Четвертый этап уже не сложен: с использованием M_0 и M_+ составляем равенство для полиномов $W_1(s)$ и $W_2(s)$ оптимального регулятора (1.18):

$$\frac{W_1(s)}{W_2(s)} = A(s) - \frac{G(s) \cdot S_1(s)}{M_0 + M_+}. \quad (1.24)$$

Несмотря на сложность вычислительного алгоритма, простота реализации оптимальных регуляторов, а главное – существенность улучшения качества работы самых различных объектов после перехода на оптимальное управление – способствовала тому, что теория оптималь-

ного управления стала широко использоваться; вышли первые монографии, посвященные изложению теории и ее приложений, а вскоре эта теория стала преподаваться в высших учебных заведениях, была включена в курсы теории управления, теории автоматического управления, теории автоматических систем читаемые в университетах и высших технических учебных заведениях.

Затем последовал неожиданный тяжелый удар: произошло несколько аварий, связанных с тем, что некоторые из оптимальных систем, построенных в полном соответствии с рекомендациями недавно разработанной теории оптимального управления, оказались способными терять устойчивость при малых отклонениях параметров объекта управления или регулятора от номинальных значений.

Эти аварии нанесли в 60-х годах XX века роковой удар по репутации теории оптимального управления и по ее практическому применению, тем более, что причина аварий долгое время оставалась совершенно загадочной. Действительно, к этому времени было уже хорошо известно, что оптимальные системы мало чувствительны к вариациям параметров (поскольку на экстремали обращается в нуль главная, линейная, часть приращения функционала) и поэтому они должны обладать параметрической устойчивостью (отметим, что «параметрически устойчивыми» называют системы, сохраняющие устойчивость при вариациях параметров). Поэтому отсутствие параметрической устойчивости выглядело очень странно, тем более, что оно наблюдалось не у всех систем. Одни оптимальные системы обладали параметрической устойчивостью, другие – не обладали и подметить какую-нибудь закономерность долго не удавалось.

Первоначально подозрение пало на алгоритм: предполагали, что параметрическая неустойчивость связана с недостатками алгоритма синтеза и поэтому велись интенсивные поиски алгоритма, обеспечивающего параметрическую устойчивость. В 1965-77 гг. вышло несколько монографий, посвященных новым методам синтеза, но каждый раз обнаруживалось, что новые алгоритмы обладают теми же недостатками в отношении параметрической устойчивости, что и прежде.

Перелом произошел в 1973 году. В начале года, в публикации [11] П.В. Надеждин обнаружил, что еще один алгоритм синтеза, незадолго до того опубликованный тремя украинскими математиками [7], не обеспечивает параметрической устойчивости, и считал это недостатком алгоритма, который можно устранить.

Однако в том же 1973 году, в монографии [12] было показано, что алгоритма, способного всегда обеспечить параметрическую устойчивость, существовать не может, поскольку для ряда объектов управле-

ния минимум критерия качества лежит на границе устойчивости и никакой алгоритм не может этого изменить. Это удалось доказать потому, что в монографии [12] для ряда спектров возмущающих воздействий были впервые получены математические модели оптимальных регуляторов в конечном виде, в виде формулы, а не алгоритма. Это сразу прояснило дело.

В частности, для широко используемого при оптимизации судов и судовых систем спектра Рахманина-Фирсова:

$$S = \langle \varphi^2 \rangle = \frac{\alpha^2 + \beta^2}{(\alpha^2 + \beta^2 + \omega^2)^2 - 4\beta^2\omega^2} \quad (1.25)$$

(где α и β – параметры, зависящие от интенсивности волнения) и объектов управления вида (1.15) при $B(D) = 1$ в [12] была получена следующая формула для оптимального регулятора

$$u_{\text{опт}} = \left[A(D) - \frac{G(D)}{a + bD} \right] x, \quad (1.26)$$

в которой полином $G(D)$ определяется на основании формулы (1.22) и имеет поэтому ту же степень (обозначим ее n), что и полином $A(D)$, старший член полинома $G(D)$ равен $a_n D^n$, а коэффициенты a и b вычисляются через вещественную и мнимую часть комплексного числа:

$$\frac{A(\alpha - j\beta)}{G(\alpha - j\beta)} = K_1 + jK_2, \quad (1.27)$$

причем $a = K_1 + \frac{\alpha}{\beta} K_2$, $b = \frac{K_2}{\beta}$.

Если параметры объекта управления в точности равны расчетным, номинальным значениям, то замкнув объект управления (1.15) оптимальным регулятором (1.27), получим математическую модель замкнутой системы:

$$G(D)x = (a + bD)\varphi(t). \quad (1.28)$$

Из формулы (1.28) следует, что в этом случае замкнутая система устойчива, поскольку полином $G(D)$ – гурвицев. Однако если коэффициенты объекта управления отличаются от номинальных даже на сколь угодно малые числа ε_i , т.е. его математическая модель имеет вид:

$$\left[(a_n + \varepsilon_n)D^n + (a_{n-1} + \varepsilon_{n-1})D^{n-1} + \dots + a_0 + \varepsilon_0 \right] x = u + \varphi, \quad (1.29)$$

то уравнение замкнутой системы примет теперь вид:

$$\left[\varepsilon_n D^n + \varepsilon_{n-1} D^{n-1} + \dots + \varepsilon_0 + \frac{G(D)}{a + bD} \right] x = \varphi(t), \quad (1.30)$$

а после домножения на $(a + bD)$ и с учетом того, что старший член полинома $G(D)$ равен $a_n D^n$, уравнение (1.30) примет вид

$$\left[\varepsilon_n b D^{n+1} + \varepsilon_n a D^n + a_n D^n + \dots \right] x = (a + bD) \varphi(t), \quad (1.31)$$

где точками обозначены члены со степенями меньшими, чем n . Из формулы (1.31) сразу видно, что даже при сколь угодно малом ε_n замкнутая система может оказаться неустойчивой, поскольку нарушится необходимое условие устойчивости в том случае, если знак произведения $\varepsilon_n b$ будет противоположен знаку членов полинома $G(D)$. Поскольку знак ε_n может быть любым, то система (1.31) параметрически неустойчива.

Это означает, что для объектов управления вида (1.15) при $B(D) = 1$ и возмущающем воздействии со спектром (1.25) задача синтеза регулятора, доставляющего минимум критерию качества (1.17) некорректна: при номинальных значениях параметров замкнутая система устойчива и критерий (1.17) достигает минимума, но уже при сколь угодно малой вариации коэффициента a_n устойчивость теряется, а в неустойчивой системе функционал (1.17) вообще не имеет конечно-го значения (т.е. «стремится к бесконечности»).

Именно во-время незамеченная некорректность задачи синтеза оптимального регулятора для некоторых спектров возмущающих воздействий и привела в свое время к нескольким неожиданным авариям и, самое главное, привела к долготелней потере доверия к теории оптимального управления и к возможности ее практического использования.

В то же время после того, как некорректность задачи была обнаружена и причины ее в монографии [12] были выяснены, оказалось совсем несложным синтезировать параметрически устойчивые регуляторы и системы, близкие к оптимальным. Первый метод синтеза, обеспечивающий параметрическую устойчивость, был предложен еще в 1973 г. в [12]. Он оказался не самым удачным и исследования продолжались. В 1977 г. в [13] был опубликован более совершенный метод, а также простой критерий различения корректности или некорректности задачи синтеза оптимального регулятора для объектов управления (1.15): если выполнено неравенство

$$p \geq m + q - 1, \quad (1.32)$$

то задача синтеза корректна, если не выполнено – то некорректна. В неравенстве (1.32) m – это степень полинома $B(D)$ в математической

модели (1.15), а p и q – половинные степени числителя и знаменателя в аналитической аппроксимации (1.16) спектра $S(\omega)$ возмущающего воздействия $\varphi(t)$. За формулой (1.32) закрепилось впоследствии название «критерий Ю.Петрова» или «неравенство Ю.Петрова».

Если критерий Ю.Петрова не выполнен, то задача синтеза оптимального регулятора некорректна, но она может быть легко регуляризована. Поясним методику регуляризации на примере.

Рассмотрим простой объект управления

$$4Dx = (D + 1)u + \varphi(t) \quad (1.33)$$

с критерием качества

$$J = 9 \langle x^2 \rangle + \langle u^2 \rangle \quad (1.34)$$

при возмущающем воздействии со спектром

$$S(\omega) = \frac{2}{\pi} \cdot \frac{1}{1 + \omega^2} \quad (1.35)$$

и поставим задачу – найти математическую модель регулятора, обеспечивающего минимум критерия (1.34). Используя алгоритм синтеза, описанный выше (а более подробно изложенный в [13], [14]), мы найдем математическую модель оптимального регулятора

$$(3D - 5)u = 12(D + 4)x. \quad (1.36)$$

Замкнув регулятором (1.36) объект управления (1.33), убедимся, что замкнутая система устойчива, а вычисляя критерий качества (1.34) в замкнутой оптимальным регулятором системе найдем его минимальное возможное значение $J_{\min} = 0.436$.

Однако на самом деле решение (1.36) смысла не имеет, поскольку рассматриваемая нами задача некорректна. Действительно, для объекта управления (1.33) и спектра (1.35) имеем: $m=1$, $p=0$, $q=1$, а это означает, что неравенство Ю. Петрова (1.32) не выполнено, параметрической устойчивости нет.

Чтобы убедиться в этом, достаточно учесть неизбежные малые отклонения коэффициентов реального объекта управления от расчетных значений. Если мы замкнем регулятором (1.36) объект управления (1.33), в котором учтены возможные вариации коэффициентов и он поэтому принял вид:

$$4(1 + \varepsilon)Dx = (D + 1)u + \varphi(t) \quad (1.37)$$

(для простоты учитываем вариацию только одного коэффициента), то характеристический полином замкнутой системы примет вид

$$\Delta = -3\varepsilon D^2 + (20 + 5\varepsilon)D + 12. \quad (1.38)$$

При $\varepsilon = 0$ он имеет единственный отрицательный корень $D_1 = -0.6$, но уже при сколь угодно малых $\varepsilon > 0$ полином (1.38) имеет два корня, второй из которых имеет положительную вещественную

часть и замкнутая система при сколь угодно малых $\varepsilon > 0$ теряет устойчивость. При $\varepsilon \leq 0$ устойчивость сохраняется.

Если бы мы замкнули регулятором (1.36) реальный объект управления (1.37), то поскольку знак ε не предсказуем, то замкнутая система могла бы оказаться и неустойчивой (при $\varepsilon > 0$), и устойчивой (при $\varepsilon \leq 0$). Если оказалось, что $\varepsilon > 0$, то система окажется неустойчивой и будет забракована, что сразу подорвет доверие к теории оптимального управления. Еще хуже будет, если окажется $\varepsilon \leq 0$. В этом случае изготовленная система может успешно пройти испытания, может неопределенно долгое время исправно работать, а затем, при неизбежном в ходе эксплуатации малом «дрейфе» параметров, может в самый неожиданный момент времени потерять устойчивость и создать аварийную ситуацию.

Разумеется, для такого простого объекта управления как (1.33) разобраться в возможности потери устойчивости еще нетрудно. Но такие же явления возникали и в гораздо более сложных системах. Поэтому неудивительно, что практическое применение теории оптимального управления на долгие годы застопорилось.

А между тем использование теории некорректных задач позволяет легко решить эту проблему. Главное – это найти простой критерий различения корректных и некорректных задач. Как только критерий [32] был найден и опубликован в [13], дальнейшее уже было нетрудно: для обеспечения параметрической устойчивости аналитическую аппроксимацию спектра, используемую при синтезе регулятора, следует выбирать так, чтобы неравенство Ю. Петрова оказалось выполненным.

Конкретно для объекта управления (1.33) достаточно выбрать для спектра возмущающих воздействий аппроксимацию

$$S_{\varphi}(\omega) = \frac{2}{\pi} \cdot \frac{1 + k^2 \omega^2}{1 + \omega^2}, \quad (1.39)$$

для которой $p=1$, $q=1$ и поэтому неравенство (32) выполнено. Таким образом, для $k \neq 0$ задача синтеза оптимального регулятора корректна. Используя тот же алгоритм синтеза, что и ранее, находим математическую модель оптимального регулятора:

$$[(3 - 11k)D - (5 + 3k)]u = 12[(1 + 3k)D + 4]x. \quad (1.40)$$

Вычисляя характеристический полином замкнутой системы, получаем его в виде

$$\Delta = (20k - 3\varepsilon + 11\varepsilon k)D^2 + (20 + 12k + 5\varepsilon + 3\varepsilon k)D + 12. \quad (1.41)$$

Если коэффициент k конечен, а ε – сколь угодно малая величина, то замкнутая система устойчива и поэтому задача синтеза – как и следовало ожидать, на основе неравенства Ю. Петрова – корректна.

Что же касается конечных, а не сколь угодно малых вариаций ϵ , то здесь дело сложнее: для очень малых k устойчивости может и не быть. Из формулы (1.41) следует, что, например, при $k=0.01$ устойчивость сохранится лишь для $\epsilon < 0.069$, но уже при $k=0.1$ устойчивость сохранится для $\epsilon < 1.05$, т.е. практически уже для любых вариаций.

В то же время не следует увлекаться большими значениями коэффициента k в аппроксимации (1.39). Дело в том, что если спектр возмущающего воздействия соответствует формуле (1.35), а мы при расчете регулятора использовали формулу (1.39), то величина критерия (1.34) неизбежно возрастет и это возрастание будет тем больше, чем больше k . Так, при $k=0$, имеем $J=0.4336$, при $k=0.1$ будет $J=0.4374$ (увеличение на 0.88%), а при $k=0.3$ величина J возрастет до $J=0.4481$. Увеличение критерия качества является как бы «платой» за обеспечение параметрической устойчивости.

Введением параметра k в формулу (1.39) мы регуляризуем исходную некорректную задачу, превращаем ее в последовательность корректных задач, решения которых в пределе, при $k \rightarrow 0$, стремятся к решению исходной некорректной задачи.

Заметим, что если $k \neq 0$, но является очень малой величиной, то задача синтеза уже не является некорректной, но остается плохо обусловленной: если, например, $k=0.01$, то уже при $\epsilon > 0.069$ устойчивость теряется, что на практике также может привести к аварии. Однако при $k=0.1$ сохранение устойчивости будет гарантировано.

Предложение обеспечивать корректность задачи синтеза оптимального регулятора и параметрическую устойчивость замкнутой системы путем введения в аналитическую аппроксимацию спектра множителей вида $1+k^2\omega^2$ было сделано и обосновано в 1977 г. в монографии [13]. Оно позволило синтезировать регуляторы, пригодные для практического использования и существенно улучшающие качество работы различных объектов управления. Недоверие к теории оптимального управления было преодолено и она стала широко использоваться на практике, принося большой экономический эффект. Практические применения методики синтеза оптимальных регуляторов отражены в публикациях [1, 5, 7, 8, 9, 12, 13, 14] и многих других. Упомянем лишь один пример: для уменьшения качки судов широко используются управляемые бортовые рули. Переход на оптимальное управление рулями позволил повысить эффективность умерения качки в 2,7 раза по сравнению с ранее используемым законом управления ([1], стр. 231). При этом оптимальный регулятор несложен, его математическая модель имеет вид

$$(0.035B + 0.85)u = -(4.15D^3 + 10.91D^2 + 6.15D + 3.35)x \quad (1.42)$$

и регулятор легко реализуется.

Столь же несложно реализуются и регуляторы, обеспечивающие эффективное управление морскими судами ([1], стр. 224), регуляторы возбуждения синхронных машин в энергетике, обеспечивающие хорошее качество напряжения при колебаниях нагрузки ([1], стр. 209–215) и т.д.

Все эти регуляторы обеспечивают параметрическую устойчивость замкнутой системы. Заметим, что уже после 1977 г., после публикации [13], наибольшее распространение получил другой метод регуляризации – вместо критерия качества (1.17) стали синтезировать регуляторы, доставляющие минимум функционалу

$$J = \langle u^2 \rangle + k_0 \langle x^2 \rangle + k_1 \langle \dot{x}^2 \rangle + \dots + k_n \langle (x^{(n)})^2 \rangle, \quad (1.43)$$

где дополнительные члены $k_1 \langle \dot{x}^2 \rangle + \dots + k_n \langle (x^{(n)})^2 \rangle$ вводятся только для преодоления некорректности. Алгоритм синтеза регулятора, доставляющего минимум функционалу (1.43) не намного сложнее, чем для функционала (1.17), параметрическая устойчивость обеспечивается, но обеспечивается ценой более значительной жертвы в критерии качества по сравнению с методикой изменения аналитической аппроксимации спектра.

Но это уже детали. Главное – это то, что выявление причин параметрической неустойчивости, выявление некорректности ряда задач оптимизации, позволило широко использовать ее для улучшения качества работы многих объектов промышленности и транспорта, принесло большой экономический эффект.

1.5. Некорректные задачи вычисления собственных значений систем линейных однородных уравнений

Многие важные задачи физики и техники приводят, как известно, к необходимости вычислять значения параметра (традиционно обозначаемого буквой λ), при которых существуют ненулевые решения системы линейных однородных уравнений с параметром. В качестве примера рассмотрим систему:

$$\begin{cases} (\lambda^2 - 2\lambda)x_1 + (1 - 3\lambda)x_2 = 0, \\ \lambda x_1 - 3x_2 = 0. \end{cases} \quad (1.44)$$

Поскольку система (1.44) однородна, то она имеет, разумеется, тривиальное решение $x_1 = x_2 = 0$. Однако, при некоторых значениях параметра λ система может иметь и ненулевые решения. Так, система (1.44) имеет ненулевые решения при $\lambda = 0$. Тогда она принимает вид

$$\begin{cases} 0 \cdot x_1 + x_2 = 0, \\ 0 \cdot x_1 - 3x_2 = 0 \end{cases} \quad (1.45)$$

и любые пары чисел (в которых $x_2 = 0$, x_1 – любое) обращают уравнения (1.44) при $\lambda=0$ в тождество.

Значения параметра λ , при которых система линейных однородных уравнений с параметром имеет ненулевые решения, называются *собственными значениями* (иногда их называют *собственными числами*). Вычисление собственных значений является этапом в решении систем линейных дифференциальных уравнений с постоянными коэффициентами и поэтому часто встречается в практике инженерных расчетов.

Рассмотрим для примера систему уравнений с переменными y_1 и y_2 :

$$\begin{cases} \ddot{y}_1 - 2\dot{y}_1 = 3y_2 - y_2, \\ \dot{y}_1 = 3y_2. \end{cases} \quad (1.46)$$

Предположим, что решениями системы (1.46) являются функции

$$\begin{aligned} y_1 &= x_1 e^{\lambda t}, \\ y_2 &= x_2 e^{\lambda t}, \end{aligned} \quad (1.47)$$

где в данном случае x_1 и x_2 – постоянные интегрирования (т.е. постоянные числа). Подставив (1.47) в (1.46), получим

$$\begin{cases} e^{\lambda t} [(\lambda^2 - \lambda)x_1 + (1 - 3\lambda)x_2] = 0, \\ e^{\lambda t} [\lambda x_1 - 3x_2] = 0. \end{cases} \quad (1.48)$$

Сократив на неравную нулю функцию $e^{\lambda t}$, приходим к системе (1.44). Мы убеждаемся, что ненулевые решения (1.47) могут быть лишь для тех значений λ , при которых имеет ненулевые решения система (1.44), т.е. для собственных значений системы. Именно эти значения λ , и только они, могут стоять в показателях экспонент решений системы линейных дифференциальных уравнений с постоянными коэффициентами.

Вот почему неотъемлемым этапом решения систем линейных дифференциальных уравнений с постоянными коэффициентами является нахождение собственных значений линейной однородной системы алгебраических уравнений – это справедливо не только для рассмотренного нами примера, но и для любой системы.

В свою очередь, собственные значения равны корням определителя полиномиальной матрицы – для системы (1.44) это будет матрица

на. Действительно рассмотрим вариацию хотя бы только одного коэффициента и вместо определителя (1.50) вычислим определитель

$$\Delta = \begin{vmatrix} \lambda^2 - 2\lambda & 1 - 3\lambda \\ \lambda & -3(1 + \varepsilon) \end{vmatrix} = -3\varepsilon\lambda^2 - 6\varepsilon\lambda + 5\lambda. \quad (1.55)$$

Сразу убедимся, что определитель (1.55) при сколь угодно малом ε имеет уже не один, а два корня: $\lambda_1 = 0$, $\lambda_2 = 2 + \frac{5}{3\varepsilon}$, причем при $\varepsilon \rightarrow 0$ второй корень отнюдь не стремится к первому и исчезает лишь при точном равенстве $\varepsilon = 0$.

Для определителей полиномиальных матриц причина некорректности сразу видна: некорректность возникает там, где при номинальных значениях коэффициентов происходит сокращение членов с наибольшей степенью параметра λ . Понятно, что даже при сколь угодно малых вариациях коэффициентов сокращения уже не происходит, в результате при вариациях параметров в определителе изменяется степень полинома от параметра λ и у полинома появляется новый корень. Здесь все достаточно прозрачно и ясно. Однако некорректность задачи вычисления определителей некоторых полиномиальных матриц означает, что будут некорректными некоторые из гораздо более важных и часто встречающихся задач – задач решения систем обыкновенных дифференциальных уравнений. Некорректность возможна даже при решении самого простого и наиболее часто встречающегося в приложениях класса дифференциальных уравнений, а именно линейных уравнений с постоянными коэффициентами.

1.6. Решение систем дифференциальных уравнений. Всегда ли решения непрерывно зависят от параметров?

Поскольку коэффициенты дифференциальных уравнений в математических моделях природных процессов и технических систем получаются на основе опыта и измерения и почти всегда известны лишь с ограниченной точностью, то необходимым (хотя и недостаточным!) условием практической пригодности результатов расчета является непрерывная зависимость решений от коэффициентов и параметров. Если непрерывной зависимости решений от параметров нет, то мы совсем не можем гарантировать, что неизбежные малые отклонения истинных значений коэффициентов от номинальных не приведут к большим ошибкам в решениях.

В этом случае даже экспериментальная проверка ничего не гарантирует – при отсутствии непрерывной зависимости решений от пара-

метров совсем рядом с экспериментально проверенными значениями могут лежать значения коэффициентов и параметров, для которых решения уравнений будут совершенно другими.

Поэтому важнейшей теоремой теории дифференциальных уравнений, теоремой, лежащей в основе всех практических приложений теории, является известная теорема о непрерывной зависимости решений обыкновенных дифференциальных уравнений от параметров. Эта теорема приводится во всех достаточно подробных курсах дифференциальных уравнений, в которых приводятся и доказательства этой теоремы – для одного уравнения n -го порядка и для системы n уравнений в нормальной форме Коши, т.е. в форме

$$\begin{cases} \dot{x}_1 = f(x_1, \dots, x_n), \\ \dots\dots\dots \\ \dot{x}_n = f(x_1, \dots, x_n). \end{cases} \quad (1.56)$$

Для справедливости теоремы достаточно, чтобы для правых частей выполнялись условия Липшица. На практике они почти всегда выполняются.

Поскольку почти любую систему уравнений, разрешенных относительно старших производных, можно привести к нормальной форме, то обычно теорему о непрерывной зависимости решений от параметров толкуют расширительно, считают ее справедливой для любых систем. На самом деле это не так.

Рассмотрим систему дифференциальных уравнений (1.46), в которую введем параметр, т.е. рассмотрим систему

$$\begin{cases} \ddot{y}_1 = 2y_1 + 3\dot{y}_2 - y_2, \\ \dot{y}_1 = my_2 \end{cases} \quad (1.57)$$

и исследуем зависимость решений от параметра m .

Характеристический полином системы (1.57) равен определителю

$$\Delta = \begin{vmatrix} \lambda^2 - 2\lambda & 1 - 3\lambda \\ \lambda & -m \end{vmatrix} = (3 - m)\lambda^2 + (2m - 1)\lambda, \quad (1.58)$$

который при $m \neq 3$ имеет два корня: $\lambda_1 = 0$, $\lambda_2 = -\frac{2m-1}{3-m}$, а при $m = 3$ – один корень $\lambda_1 = 0$. Сразу видно, что значение $m = 3$ будет особым. Обозначим $m = 3(1 + \epsilon)$ и исследуем зависимость решений от ϵ . Общее решение системы (1.57) имеет вид

$$y_1(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} = C_1 + C_2 e^{\left(2 + \frac{5}{3\epsilon}\right)t}. \quad (1.59)$$

Тот же вид имеет и решение $y_2(t)$. Постоянные интегрирования в

формуле (1.59) определяются из начальных условий. Пусть начальные условия таковы, что $C_1 = 0$, $C_2 = 1$. Проанализируем – чему будет равно решение $y_1(t)$, например, при значении времени $t=1$ в зависимости от ε . Если ε – малое отрицательное число, то показатель экспоненты $2 + \frac{5}{3\varepsilon}$ является большим по абсолютной величине отрицательным числом, откуда следует, что $y_1(t)$ будет близко к нулю. Если же ε является малым положительным числом, то показатель экспоненты $2 + \frac{5}{3\varepsilon}$ велик и решение $y_1(t)$ при $t=1$ будет очень большим числом – тем большим, чем меньше ε . При $\varepsilon \rightarrow 0$ (т.е. при $m \rightarrow 3$) будет $y_1(t) \rightarrow \infty$.

Таким образом, в точке $m = 3$ зависимость решения $y_1(t)$ от параметра m терпит разрыв, непрерывность нарушается.

Теперь посмотрим, к чему это приводит на практике. Пусть номинальное значение параметра m в системе уравнений (1.57) равно 2.999, а точность его задания (определяемая, например, точностью измерения) равна двум тысячным. Вычисляя, например, значение решения $y_1(t)$ при $t=1$ (или, для любого другого значения t) при номинальном значении параметра $m=2.999$, мы получим, что с точностью до четвертого знака $y_1(t) = 0$. Однако с учетом точности задания параметра m в уравнениях (1.57) его истинное значение может быть равно, например, 3.001, и тогда истинное значение решения $y_1(t)$ будет очень большим числом, ошибка расчета будет велика и может привести к самым печальным последствиям, к авариям и даже катастрофам. Это связано с тем, что при $m = 3$ задача нахождения решения системы (1.57) некорректна.

Такие же грубые ошибки в расчетах будут возникать во всех тех системах уравнений, характеристический полином которых будет изменять свою степень при некоторых критических значениях параметра из-за сокращения коэффициентов при старшей степени. Если значение параметра не в точности равно критическому, но близко к нему, то в характеристическом полиноме изменения степени не произойдет, но коэффициент при старшей степени будет малой величиной, т.е. полином примет вид

$$\Delta = \varepsilon \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0, \quad (1.60)$$

где ε – величина много меньшая, чем коэффициенты $a_{n-1}, a_{n-2}, \dots, a_0$. Полином (1.60) будет иметь один большой (по абсолютной величине)

корень $\lambda_n \approx -\frac{a_{n-1}}{\varepsilon}$, а остальные $n-1$ корней будут близки к корням полинома (1.60) без первого члена. Это легко доказать, поделив полином (1.60) на двучлен:

$$\frac{\varepsilon}{a_{n-1}}\lambda + 1. \quad (1.61)$$

В остатке получим полином, близкий (для малых ε) к полиному (1.60) без первого члена. Знак большого корня $\lambda_n \approx -\frac{a_{n-1}}{\varepsilon}$ будет зависеть от знака ε и поэтому корень изменит свой знак при переходе ε через значение нуля. Это говорит о том, что при $\varepsilon < 0$ в решении будет присутствовать очень быстро возрастающий член, а при $\varepsilon = 0$ зависимость решения от параметра терпит разрыв. Все это мы видели на примере системы (1.57).

Таким образом, мы приходим к выводу: если при некотором критическом значении параметра в характеристическом полиноме системы дифференциальных уравнений происходит сокращение коэффициентов при члене со старшей степенью, то при этом значении параметра задача решения системы дифференциальных уравнений некорректна: уже при сколь угодно малой вариации параметра решение для любых конечных значений аргумента t может измениться в любое число раз. Причина некорректности – разрыв непрерывности в зависимости решений от параметра.

Если значение параметра не в точности равно критическому, но отличается от него на малое (но конечное) значение, то задача решения системы уравнений перестает быть некорректной, но остается плохо обусловленной: малые (но теперь уже только конечные) изменения параметра могут изменить решение в любое число раз.

Заметим, что встреча с плохо обусловленной задачей в данном случае опасней, чем встреча с задачей некорректной: о некорректности задачи будет свидетельствовать легко заметное изменение степени характеристического полинома при вариациях параметров, а в плохо обусловленной задаче степень характеристического полинома и порядок системы при вариациях параметра не изменяются и поэтому вероятность грубой ошибки при расчете больше.

Разумеется, есть простой способ обнаружения некорректных или плохо обусловленных задач – это повторение решения при немного измененных значениях коэффициентов и параметров. Однако в системах уравнений высокого порядка, когда число коэффициентов велико, этот простой метод становится очень трудоемким.

Изложенный выше материал позволяет предложить простой кри-

терий выделения подозрительных на плохую обусловленность систем дифференциальных уравнений: в таких системах старший член характеристического полинома (или старший член и несколько членов меньшей степени) существенно меньше остальных членов. Это может (но не обязательно!) говорить о том, что старший член характеристического полинома образовался как малая разность больших членов в исходной системе и их малые вариации могут привести к большим (относительным) изменениям старшего члена или же могут привести к изменению его знака. Мы уже видели это на примере системы (1.57).

Любопытно отметить, что некорректные задачи, которые можно, как мы уже говорили, рассматривать как предельный случай задач плохо обусловленных, выделить легче: в некорректных задачах решения систем дифференциальных уравнений при номинальных значениях коэффициентов происходит точное сокращение старших членов характеристического полинома и его степень понижается. Понижение степени является простым диагностическим признаком некорректности. Так, в системе (1.57) характеристический полином – равный определителю (1.58) – должен иметь вторую степень, и когда при $m = 3$ его степень понижается, это обстоятельство сразу и громко говорит о том, что при $m = 3$ задача становится некорректной. Для задач плохо обусловленных все сложнее, степень характеристического полинома у них не меняется, а малость коэффициента при его старшем члене говорит лишь о том, что данная система подозрительна в отношении ее плохой обусловленности по отношению к вариациям ее коэффициентов.

Теперь рассмотрим кажущееся противоречие между известной теоремой о непрерывной зависимости решений от параметров для одного дифференциального уравнения любого порядка и отсутствием этой непрерывной зависимости у некоторых систем дифференциальных уравнений. Кажущееся противоречие связано с тем, что систему уравнений можно привести к одному уравнению путем эквивалентных преобразований, не изменяющих решений, и поэтому, на первый взгляд, кажется, что раз теорема о непрерывной зависимости решений от параметра справедлива для одного уравнения, то она должна быть справедлива и для эквивалентной ему системы. На самом деле это не так.

Система (1.57) при $m = 3$ имеет решение $y_1 = C_1$ и простым эквивалентным преобразованием приводится к уравнению

$$(3 - m)y_1 + (2m - 1)\dot{y}_1, \quad (1.62)$$

которое при $m = 3$ переходит в уравнение

$$5\dot{y}_1 = 0, \quad (1.63)$$

имеющее то же решение $y_1 = C_1$. Решение уравнения (1.63) зависит от

его единственного коэффициента $2m-1$, равного 5 при $m=3$, а также и от параметра m непрерывно, но из этого факта нельзя сделать никаких заключений относительно непрерывной зависимости от параметра m решений системы (1.57). Мы уже убедились, что для системы (1.57) зависимость решений от параметра m в точке $m=3$ терпит разрыв, а формула (1.62) показывает причину этого: при $m=3$ коэффициент перед членом со старшей производной обращается в нуль, и, следовательно, член со старшей производной, в коэффициент перед которой входит параметр m , при $m=3$ просто исчезает. Таким образом, из справедливости теоремы о непрерывной зависимости решений одного дифференциального уравнения от параметра совсем не следует справедливость этой теоремы для всех систем дифференциальных уравнений.

Аналогичный вывод справедлив и для системы дифференциальных уравнений, записанных в нормальной форме Коши, в форме (1.56). Для систем, записанных в нормальной форме, теорема о непрерывной зависимости решений от параметров справедлива, но это не дает оснований утверждать справедливость этой теоремы для систем уравнений, не записанных в нормальной форме – даже если эти системы можно привести к нормальной форме путем эквивалентных преобразований (об эквивалентных преобразованиях и их свойствах мы будем более подробно говорить во второй главе).

Однако в учебниках и учебной литературе по дифференциальным уравнениям приводятся доказательства теоремы о непрерывной зависимости решений от параметров для систем в нормальной форме Коши и для одного уравнения, но ничего не говорится о том, что для систем, записанных не в нормальной форме, непрерывной зависимости может и не быть.

В результате выпускники высших учебных заведений и даже почти все выпускники физико-математических и математико-механических факультетов университетов – уверены, что непрерывная зависимость решений от коэффициентов существует у всех систем дифференциальных уравнений и эта ложная уверенность может стать причиной грубых ошибок при практических расчетах.

Поэтому необходимо предупреждение о существовании систем дифференциальных уравнений, в которых может отсутствовать непрерывная зависимость решений от параметров.

Приведем примеры подобных систем:

1. Система

$$(D^3 + 4D^2 + 2D + 4)x_1 = (D^2 - D + 2)x_2, \quad (1.64)$$

$$(D^2 - 6D + 6)x_1 = (D - 25)x_2 \quad (1.65)$$

приводилась в работе [17]. Легко убедиться, что коэффициенты системы соответствуют точке разрыва непрерывности в зависимости решений от коэффициента при Dx_1 в уравнении (1.65) и от коэффициента D^2x_2 в уравнении (1.64).

2. Система

$$[TD^3 + (2 + 2T)D^2 + (4 + T)D + s]x_1 = (D^2 + 2D + 1)x_2, \quad (1.66)$$

$$(D + 1)x_2 = (D^2 + 4D + 5)x_1, \quad (1.67)$$

исследованная ранее в работах [15] и [17], описывает процессы, протекающие в системе регулирования частоты вращения электропривода постоянного тока. В системе (1.66)–(1.67) переменная x_1 является отклонением частоты вращения от желаемого значения, переменная x_2 – отклонение тока якоря от номинального значения – является управляющим воздействием. Уравнение (1.67) является математической моделью оптимального регулятора – оптимального для возмущающих воздействий со спектром

$$S = \frac{1}{(1 + \omega^2)^2}. \quad (1.68)$$

Параметр T является механической постоянной времени электропривода, а поскольку в уравнениях (1.66)–(1.67) время t выражено в долях от номинального значения механической постоянной, то номинальное значение $T=1$. Легко убедиться, что это же значение $T=1$ является точкой разрыва непрерывности решения $x_1(t)$ от параметра T . Действительно, характеристический полином системы (1.66)–(1.67) равен

$$\Delta = (1 - T)\lambda^4 + (4 - 3T)\lambda^3 + (8 - 3T)\lambda^2 + (8 - T)\lambda + 3 \quad (1.69)$$

и при сколь угодно малых отклонениях параметра T от номинального значения $T=1$ в решениях системы возникает экспоненциально растущий четвертый член, растущий тем быстрее, чем меньше разность $T-1$.

Для T , сколь угодно близких к $T=1$, скорость роста этого члена может быть сколь угодно велика. Поэтому для T , близких к $T=1$, результаты расчета по математической модели (1.67)–(1.67) заведомо недостоверны. Решения дифференциальных уравнений (1.66)–(1.67) могут не иметь ничего общего с реальным поведением электропривода. На примере уравнений (1.67)–(1.67) видна разница между вариацией какого-либо коэффициента дифференциального уравнения и вариацией параметра того или иного технического устройства или физического процесса. Вариация параметра T – механической постоянной времени электропривода – может приводить к разным по величине

вариациям сразу нескольких коэффициентов в системе дифференциальных уравнений (в уравнении (1.66) – коэффициентов при D^3 , D^2 и D) и это необходимо учитывать при расчетах.

Пример с системой (1.66)–(1.67) и многочисленные подобные примеры, приведенные в публикациях [14, 15, 17], показывают, что потеря непрерывной зависимости решений от параметров особенно часто происходит в системах, среди которых встречаются так называемые неканонические уравнения, у которых порядок производных в правой части уравнения выше, чем в левой – примером как раз является уравнение (1.67). Неканонические уравнения, и системы уравнений, в которые они входят, изучались В.А. Стекловым в 1927 г. (смотри публикацию [25]). В последние десятилетия им уделялось незаслуженно мало внимания, а ведь они довольно часто встречаются в приложениях и обладают интересными свойствами.

Нетрудно установить простой критерий различения корректных и некорректных задач решения систем линейных дифференциальных уравнений с постоянными коэффициентами. Для этого достаточно составить «матрицу степеней», т.е. матрицу, каждый следующий элемент которой является старшим членом операторного полинома в системе (1.53) и посмотреть, будет ли происходить сокращение старших членов в определителе матрицы.

Для системы (1.57) «матрица степеней» принимает вид:

$$\begin{pmatrix} D^2 & -3D \\ D & -m \end{pmatrix} \quad (1.70)$$

и в ее определителе происходит сокращение членов, содержащих D^2 при $m=3$. При $m=3$ задача нахождения решения системы некорректна.

Для системы (1.66)–(1.67) «матрица степеней» имеет вид:

$$\begin{pmatrix} TD^3 & -D^2 \\ D^2 & -D \end{pmatrix} \quad (1.71)$$

и в ее определителе сокращение членов, содержащих D^4 , происходит при номинальном значении $T=1$ параметра T – «механической постоянной времени электропривода».

Таким образом, нетрудно ввести очень простой критерий различения некорректных задач от корректных в очень важной для практики проблеме решения систем линейных дифференциальных уравнений с постоянными коэффициентами (этот же критерий легко распространить на многие нелинейные задачи). Гораздо более трудно найти критерий различения хорошо и плохо обусловленных задач. Малость старшего члена характеристического полинома системы дифференциальных уравнений наводит на подозрения о плохой обусловленности

системы, не более. Так, в одном из учебных пособий приводится пример системы $\dot{x} = Ax$, где матрица A имеет вид

$$A = \begin{pmatrix} 0 & 3.99 & 0 & 0 & 0 & 0 & 0 \\ -2575 & -127.4 & 2710 & 0 & 0 & 0 & 0 \\ 0 & 0 & -300 & 4200 & 520.8 & -520.8 & -692.7 \\ 0 & 0 & 0 & 0 & 15.8 & -15.8 & -21 \\ 0 & 0 & 0 & 0 & 0 & 0 & -21.3 \\ 0 & 10 & 0 & 0 & 0 & -500 & 0 \\ 24 & 0 & 0 & 0 & 0 & 0 & -200 \end{pmatrix}$$

и характеристический полином равен

$$\Delta = \lambda^7 + 1.127 \cdot 10^3 \lambda^6 + 4.427 \lambda^5 + 9.389 \cdot 10^7 \lambda^4 + 1.181 \cdot 10^{10} \lambda^3 + 7.838 \cdot 10^{11} \lambda^2 + 1.326 \cdot 10^{13} \lambda + 1.84 \cdot 10^{14}.$$

Первый член полинома много меньше остальных, и поэтому он может быть малой разностью больших чисел и может изменить знак при малых (но конечных!) вариациях коэффициентов матрицы, но это подозрение нуждается в проверке.

1.7. Заключение

На основе материала, изложенного в главе 1, можно сформулировать следующие выводы:

1. Помимо известных некорректных задач в области математической физики, в области уравнений в частных производных и интегральных уравнений, существует много более простых, но не менее важных некорректных задач среди алгебраических уравнений, дифференциальных уравнений, задач на экстремумы и т.п.

2. Каждая неожиданная встреча с некорректной задачей может привести к грубой ошибке в расчетах и стать причиной аварий и даже катастроф. Об этом рассказано в §4, а более подробно с примерами и фактами – в монографии [17].

3. Для предотвращения ошибок рекомендуется перед решением любой задачи удостовериться – относится она к классу корректных, или к классу некорректных задач. Такая проверка в настоящее время проводится далеко не всегда, а желательно проводить ее всегда. Прямая проверка (через повторные расчеты с учетом вариаций параметров), разумеется, громоздка, поэтому очень полезно найти простые критерии различения корректных и некорректных задач. Для проблемы синтеза оптимальных задач управления таким критерием является простое неравенство (1.32) – критерий Ю.Петрова.

Для систем линейных дифференциальных уравнений с постоянными коэффициентами простой критерий вытекает из рассмотренных в §6 «матриц степеней».

В общем случае подобными критериями могут являться также замкнутость (или незамкнутость) оператора, ограниченность (или неограниченность) обратного оператора и т.д.

Глава 2

КЛАСС ЗАДАЧ, ПРОМЕЖУТОЧНЫХ МЕЖДУ КОРРЕКТНЫМИ И НЕКОРРЕКТНЫМИ

2.1. Обнаружение третьего класса задач математики, физики и техники и его значение

После того, как в 1902 г. французский математик Жак Адамар доказал существование класса некорректных задач, на протяжении девяти десятилетий XX века считалось общепризнанным, что все задачи математики, физики и техники делятся на два класса – на издавна известный класс корректных задач и класс задач некорректных, исследовать который начали с 1902 года и в исследование которого внесли важный и существенный вклад многие российские ученые – смотри публикации [6, 10, 20, 23, 24, 27, 28, 29, 30].

В ходе исследований, проводимых на кафедре моделирования электромеханических систем факультета прикладной математики – процессов управления Санкт-Петербургского Государственного Университета (СПбГУ), было обнаружено существование третьего класса задач – задач, меняющих свою корректность в ходе их решения. Эти задачи нельзя отнести ни к корректным, ни к некорректным и их поэтому следует выделить в особый третий класс – класс задач, промежуточных между корректными и некорректными – смотри публикацию [16].

Важность этого третьего класса заключается в том, что его существование заставляет совсем по другому решать важную проблему различения корректных и некорректных задач. После 1902 года, после опубликования первых работ Ж. Адамара и вплоть до последнего десятилетия XX века, все верили, что корректность задачи достаточно проверить один раз – проверить перед ее решением. На самом деле оказалось, что такая проверка не всегда достаточна, поскольку существуют – как будет показано далее – задачи, изменяющие свою корректность в ходе совершенно эквивалентных преобразований, использующихся в ходе их решения. Поэтому для обеспечения достоверности результатов расчета необходимо либо проверять корректность много раз, после каждого преобразования, используемого при решении (а это, разумеется, очень громоздко), либо надо внимательно проанализировать те эквивалентные преобразования, которые используются при решении почти всех задач. К анализу эквивалентных преобразований мы и пе-

рейдем. Мы убедимся – как будет далее показано, – что эквивалентные преобразования полезно разделить на преобразования, эквивалентные в классическом смысле, и преобразования, эквивалентные в расширенном смысле. Мы покажем, что только преобразования, эквивалентные в расширенном смысле, обязаны не изменять корректности решаемых задач, и что только их можно спокойно применять при анализе всех вопросов, связанных с корректностью (в частности – при анализе параметрической устойчивости систем управления).

2.2. Преобразования, эквивалентные в классическом смысле

Эквивалентные преобразования (их называют еще равносильными преобразованиями) входят в программу средних школ, и теория эквивалентных преобразований считалась окончательно построенной и завершенной еще в восемнадцатом веке. Простейшие эквивалентные преобразования – это прибавление к правой и левой частям уравнения одного и того же числа, перенос членов из левой части в правую с изменением знака, умножение всех членов уравнения на число, не равное нулю, подстановка, т.е. замена любого члена на равное ему выражение и т.п.

Определение: *эквивалентными (равносильными) системами уравнений* называют системы, множества решений которых совпадают (Математическая энциклопедия, том 4, стр. 800, М.: Советская энциклопедия, 1984). Такое определение дает и "Математический энциклопедический словарь", стр. 511, издание 1995 г.

Эквивалентными преобразованиями называют те, при которых исходная и преобразованная системы эквивалентны между собой.

Из определения сразу видно, что эквивалентность или неэквивалентность систем зависит от решаемой задачи и от того, в каком множестве отыскиваются решения.

Так, например, если поставлена задача найти значения x , удовлетворяющие системе

$$\begin{cases} 3x + y = 4, \\ x + y = 2, \end{cases} \quad (2.1)$$

а отыскивать значение y не требуется, то системе (2.1) эквивалентно уравнение $2x=2$, получившееся вычитанием второго уравнения из первого. Если же задано найти и значения x и значения y , удовлетворяющие системе (2.1), то уравнение $2x=2$ уже не эквивалентно ей.

Уравнения $x^2 - 1 = 0$ и $x^4 - 1 = 0$ эквивалентны в поле действительных чисел (оба имеют корни $x_1 = 1$, $x_2 = -1$) и не эквивалентны в

поле комплексных чисел, поскольку уравнение $x^4 - 1 = 0$ имеет еще и комплексные корни $x_3 = j$, $x_4 = -j$.

При решении уравнений и систем уравнений надо тщательно следить за эквивалентностью используемых преобразований. При многих преобразованиях в преобразованной системе возникают лишние решения, не удовлетворяющие исходной системе, или же теряется часть решений. Такие преобразования, разумеется, не эквивалентны, но (в отличие от совершенно ошибочных преобразований, меняющих все множество решений), их полезно выделить в отдельный класс не полностью эквивалентных преобразований, эквивалентность которых можно восстановить, если убрать лишние корни и восстановить недостающие.

Так, уравнение $x - 2 = 0$ и уравнение $x^2 - 3x + 2 = 0$, полученное из первого умножением на двучлен $x - 1$, разумеется, не эквивалентны: первое имеет один корень $x_1 = 2$, второе – два: $x_1 = 2$, $x_2 = 1$. Но можно записать: уравнение $x^2 - 3x + 2 = 0$ после вычеркивания лишнего корня $x=1$ (и только после его вычеркивания!) эквивалентно уравнению $x - 2 = 0$. Использование таких не полностью эквивалентных преобразований (разумеется, после исключения лишних корней) полезно при решении некоторых практических задач.

Кроме простейших, известных со средней школы, эквивалентных преобразований, рассмотрим еще правила почленного дифференцирования и умножения на полиномы от оператора дифференцирования, исследованные Л. Эйлером в XVIII веке.

Уравнение

$$\dot{x} + x = 0, \quad (2.2)$$

с начальным условием $x(0) = 0$ имеет единственное решение $x = 0$, которое можно получить из общего решения уравнения (2.2): $x = C_1 e^{-t}$, определив постоянную C_1 из начального условия. В нашем случае $C_1 = 0$.

Продифференцировав все члены уравнения (2.2), мы приходим к уравнению второго порядка:

$$\ddot{x} + \dot{x} = 0, \quad (2.3)$$

в общее решение которого

$$x(t) = C_1 e^{-t} + C_2 \quad (2.4)$$

входят две постоянные интегрирования и для определения конкретного решения требуется знание двух начальных условий. К имеющемуся начальному условию $x(0) = 0$ нужно добавить второе. Это второе начальное условие, условие для производной \dot{x} при $t = 0$ следует добав-

лять, разумеется, не произвольно, а по четкому правилу. Правило это заключается в том, что мы вычисляем, чему будет равно значение производной \dot{x} при $t=0$ для уравнения (2.2) и именно это значение используем как второе начальное условие для продифференцированного уравнения, для уравнения (2.3). Поскольку нами было найдено решение уравнения (2.2): $x=0$, то из него следует, что $\dot{x}(0)=0$. Добавив это второе начальное условие, мы из общего решения уравнения (2.3) – решения (2.4) – найдем конкретное решение уравнения (2.3): $x(t)=0$, т.е. то же самое, что и решение уравнения (2.2). Таким образом, при правильном назначении дополнительных начальных условий почленное дифференцирование является эквивалентным преобразованием.

Отметим, что против этого положения иногда выставляют возражение, основанное на не совсем четком понимании термина «общее решение дифференциального уравнения». Часто говорят «после почленного дифференцирования уравнение (2.2), перейдя в уравнение (2.3), стало иметь другое общее решение – решение (2.4), а значит уравнения (2.2) и (2.3) не эквивалентны между собой». Однако, пришедший из XVIII века термин «общее решение» не совсем удачен. Правильнее говорить не «общее решение», а «семейство решений» – семейство, зависящее от постоянных интегрирования. Определяя эти постоянные из начальных условий, мы получаем конкретное решение.

Поэтому нет ничего удивительного в том, что после почленного дифференцирования семейство возможных решений изменяется, но само конкретное решение при правильном выборе дополнительных начальных условий остается неизменным.

Почленное дифференцирование, как известно, используется во многих методах интегрирования дифференциальных уравнений.

2.3. Обнаружившиеся парадоксы

Несмотря на кажущуюся простоту эквивалентных преобразований, изучаемых еще в средней школе, в двадцатом веке обнаружили связанные с ними сложности, далеко не сразу получившие правильное объяснение.

Началось все с теории управления, с предложенного в 60-х годах XX века А.М. Летовым метода синтеза систем управления, получившего тогда широкую известность под названием «аналитическое конструирование оптимальных регуляторов» или просто «аналитическое конструирование» [8, 9]. Заметим, что с математической точки зрения «регулятором» называют любую зависимость между управляющим воздействием (управлением) «и» и регулируемыми переменными x_i – типа зависимостей (20), (26), (36), (40) из главы первой. Затем эти за-

висимости реализуются в технических устройствах, также называемых «регуляторами».

А.М. Летов установил, что для весьма распространенных линейных объектов управления, математической моделью которых является система линейных дифференциальных уравнений в нормальной форме

$$\begin{cases} \dot{x}_1 = a_{11}x_1 + \dots + a_{1n}x_n + b_1u, \\ \dots \\ \dot{x}_n = a_{n1}x_1 + \dots + a_{nn}x_n + b_nu \end{cases} \quad (2.5)$$

(где x_i – регулируемые переменные, u – управление), можно построить очень простой регулятор вида:

$$u = -k_1x_1 - \dots - k_nx_n, \quad (2.6)$$

который (при соответствующем выборе коэффициентов усиления в регуляторе – коэффициентов k_1, k_2, \dots, k_n) обеспечивает устойчивость замкнутой системы, ее параметрическую устойчивость и минимум квадратичной формы от регулирующих переменных и управления – например, формы:

$$J = \int_0^{\infty} (u^2 + c_1x_1^2 + \dots + c_nx_n^2) dt. \quad (2.7)$$

А.М. Летовым была разработана методика вычисления коэффициентов усиления k_i в регуляторе (2.6), обеспечивающих минимум формы (2.7) для любых c_i , а также обеспечивающих устойчивость замкнутой системы. Эта методика была названа им «аналитическим конструированием» и описана в статьях [8] и книге [9].

Поскольку минимум квадратичных форм – типа формы (2.7) – часто неплохо отражает реальные требования к качеству работы объектов управления и обеспечивает хорошие переходные процессы, а регуляторы вида (2.6) легко реализовывались, то «аналитическое конструирование» регуляторов быстро приобрело самую широкую популярность. Регуляторы вида (2.6) стали использоваться для оптимизации самых различных объектов управления, стало публиковаться много статей (только в 1960–1968 г.г. – более 200 статей), посвященных различным проблемам методики «аналитического конструирования».

При практическом применении методики быстро выявился интересный парадокс: в реальных объектах управления очень часто далеко не все регулируемые переменные x_1, x_2, \dots, x_n в уравнении (2.5) доступны для измерения и использования в регуляторе (2.6). Часть переменных не измерима. Наиболее простым и естественным выходом из

этого положения было исключение неизмеряемых переменных из уравнений (2.5)–(2.6) с помощью эквивалентных преобразований. После этого исключения в уравнения объекта управления (2.5) и регулятора (2.6) входили только измеряемые переменные и их производные. Поскольку преобразования были эквивалентными, решения $x_i(t)$ оставались неизменными, то ожидалось, что устойчивость замкнутой системы и ее параметрическая устойчивость измениться не должны. Эти ожидания не оправдались: устойчивость сохранялась, переходные процессы сохранялись, но параметрическая устойчивость часто не сохранялась. Системы, спроектированные и изготовленные по всем рекомендациям методики «аналитического конструирования» часто теряли устойчивость при неизбежных в ходе эксплуатации малых отклонениях параметров объекта управления или регулятора от расчетных значений и создавали аварийные ситуации. Это вызвало резкое падение интереса к «аналитическому конструированию» и оно быстро сошло на нет, тем более, что причины потери параметрической устойчивости долго оставались непонятными (см., например, статью [11], в которой обсуждалась проблема потери параметрической устойчивости, и дискуссию, возникшую на страницах наиболее авторитетного в области управления журнала «Автоматика и телемеханика» после публикации этой статьи; несмотря на накал дискуссии, к сожалению, не доведенной до конца, тогда, в 1973 г., понимания сущности проблемы и причины изменения параметрической устойчивости при эквивалентных преобразованиях достигнуть не удалось).

2.4. Преобразования, эквивалентные в расширенном смысле

Проблема возможного изменения параметрической устойчивости после эквивалентных преобразований существенно прояснилась после того, как в монографии [14] был поставлен простой вопрос: а что, собственно, означает утверждение: «система управления параметрически устойчива (стабильна) или – что то же самое – задача проверки устойчивости для данной системы корректна»? Ведь фактически это утверждение не о самой системе, не об ее решениях. Это утверждение об окрестности системы, утверждение о том, что у других систем, близких к исходной в пространстве параметров (но не тождественных ей) все решения устойчивы. Когда мы говорим: решение уравнения $\dot{x} + x = 0$ параметрически устойчиво, задача проверки устойчивости для этого уравнения корректна, то это равнозначно утверждению: все решения семейства уравнений $(1 + \varepsilon_1)\dot{x} + (1 + \varepsilon_2)x = 0$, которое является

окрестностью уравнения $\dot{x} + x = 0$ при малых ε_1 и ε_2 , устойчивы. Если в окрестности рассматриваемой нами системы лежит хотя бы одна система, решения которой не устойчивы, то рассматриваемая система параметрически неустойчива, задача проверки устойчивости для нее некорректна.

А из этого сразу следует, что эквивалентные преобразования, которые, по определению, не меняют решений преобразуемой системы, совсем не обязаны всегда, во всех случаях, сохранять свойства окрестности преобразуемой системы, а значит и не обязаны сохранять параметрическую устойчивость преобразуемой системы, не обязаны сохранять корректность рассматриваемой задачи. Они могут сохранять корректность – и очень часто сохраняют – но могут и не сохранять. А это значит, что примеры несохранения параметрической устойчивости при эквивалентных преобразованиях, с которыми сталкивались практики и исследователи еще в 70-х годах и которые вызывали тогда недоумение и острые дискуссии, на самом деле не являются чем-то необычным и парадоксальным. Так вполне может быть.

Вот один из примеров.

Система двух дифференциальных уравнений

$$\begin{cases} (D^3 + 4D^2 + 5D + 2)x_1 = (D^2 + 2D + 1)x_2, \\ (D^2 + 4D + 5)x_1 = (D + 1)x_2, \end{cases} \quad (2.8)$$

где $D = \frac{d}{dt}$ – оператор дифференцирования, описывающая процессы, протекающие в системе управления частотой вращения электропривода постоянного тока (x_1 – частота вращения, x_2 – управление, ток якоря), имеет характеристический полином

$$\Delta = \lambda^3 + 5\lambda^2 + 7\lambda + 3 \quad (2.9)$$

с корнями $\lambda_1 = -3$, $\lambda_2 = \lambda_3 = -1$ и является устойчивой.

Вводя новые переменные x_3 и x_4 с помощью, например, равенств

$$\begin{cases} x_3 = \dot{x}_1 + 2x_1 - x_2, \\ x_4 = \dot{x}_3, \end{cases} \quad (2.10)$$

мы приводим систему (2.8), как и полагается, к виду четырех дифференциальных уравнений первого порядка относительно x_1 , x_2 , x_3 , x_4 :

$$\begin{cases} \dot{x}_1 = -2x_1 + x_2 + x_3, \\ \dot{x}_3 = x_4, \\ \dot{x}_4 = -x_3 - 2x_4, \\ 0 = x_1 + x_2 + 2x_3 + x_4, \end{cases} \quad (2.11)$$

причем последнее дифференциальное уравнение оказывается вырожденным, превращается в конечное соотношение между переменными x_1, x_2, x_3, x_4 . Более подробно переход от из уравнений (2.8) к последнему из уравнений (2.11) будет рассмотрен еще раз в §5.

Система (2.11) имеет тот же характеристический полином (2.9), что и система (2.8), она также, как и система (2.8), устойчива, семейства решений обеих систем имеют одинаковый вид:

$$x_1(t) = C_1 e^{-3t} + (C_2 t + C_3) e^{-t}. \quad (2.12)$$

Системы (2.8) и (2.11) эквивалентны. И в то же время система (2.11) параметрически устойчива, а система (2.8) – нет: при сколь угодно малых вариациях некоторых коэффициентов (например, коэффициента при $D^2 x_2$ в правой части первого из уравнений (2.8)) система (2.8) может потерять устойчивость.

Примеров, подобных системам (2.8)–(2.11) в работах [16, 17] было приведено много и эти примеры говорят о желательности введения нового математического понятия: помимо прежнего классического понятия эквивалентных преобразований, не изменяющих решений преобразуемых систем, которые мы в дальнейшем будем называть преобразованиями, эквивалентными в классическом смысле, введем понятие преобразований, эквивалентных в расширенном смысле.

Преобразованиями, эквивалентными в расширенном смысле, назовем преобразования, которые:

- 1) во-первых, эквивалентны в классическом смысле,
- 2) во-вторых, не изменяют корректности рассматриваемой задачи.

Большинство эквивалентных преобразований эквивалентны и в классическом смысле, и в расширенном и поэтому необходимость в новом понятии долго не осознавалась. Однако примеры, приведенные в [14, 15, 17], показали, что существует (хотя и редко) простые преобразования, эквивалентные в классическом смысле, но не в расширенном.

Важность введенного в [15] понятия преобразований, эквивалентных в расширенном смысле, заключается в том, что как традиционные методы проверки сохранения устойчивости при вариациях параметров, так и новейшие методы [18], использующие, например, теорему В. Л. Харитоновна [22], основаны на исследованиях характеристического полинома замкнутой системы. Между тем уже пример с системами (2.8) и (2.11), когда две системы, имеющие один и тот же характеристический полином, различаются по параметрической устойчивости, говорит о том, что традиционные методы расчета устойчивости заведомо неполны, не могут во всех случаях гарантировать правильного результата. Недостаточно исследовать характеристический полином – надо

проверить, какими преобразованиями получен он из первичных, исходных уравнений исследуемой системы. Если эти преобразования эквивалентны в расширенном смысле – все хорошо. Если они эквивалентны в классическом смысле, но не в расширенном, то любое исследование характеристического полинома не может спасти от возможной ошибки в расчетах, причем ошибки с тяжелыми последствиями, ошибки, которая может стать причиной аварий и даже катастроф. Об авариях, связанных с ошибками в расчетах устойчивости, более подробно рассказано в [17].

Исследования, проведенные в [16, 17], показали, что наличие у системы характеристического полинома, имеющего корни только с отрицательными вещественными частями, наличие матрицы коэффициентов (при записи системы в нормальной форме Коши) с собственными значениями с отрицательными вещественными частями, или существование у исследуемой системы функции Ляпунова не гарантирует правильного, надежного ответа на вопрос о параметрической устойчивости.

Наиболее простой путь обеспечения надежности расчетов – это проверка используемых в ходе расчета преобразований на эквивалентность в расширенном смысле. Но для этого необходима разработка теории подобных преобразований, которая еще далека от завершения.

2.5. Задачи, промежуточные между корректными и некорректными

Из существования преобразований, изменяющих корректность (т.е. преобразований, эквивалентных в классическом смысле, но не в расширенном) следует возможность существования задач, которые нельзя отнести ни к корректным, ни к некорректным. Такие задачи целесообразно выделить в отдельный, *третий класс*.

Пример – уже рассмотренная задача проверки устойчивости для математической модели электропривода постоянного тока. Если математическая модель записана в нормальной форме Коши, то задача проверки устойчивости корректна. Если математическая модель записана в эквивалентной (в классическом смысле) форме системы уравнений (2.8), то та же задача проверки устойчивости уже не корректна.

Многочисленные примеры задач, относящихся к третьему классу, возникают там, где при решении используются цепочки преобразований.

Рассмотрим систему линейных однородных уравнений с параметром λ :

$$\begin{cases} (1-\lambda)x_1 + x_2 + 2x_3 = 0, \\ x_1 + (1-\lambda)x_2 + 3x_3 = 0, \\ x_1 + x_2 = 0 \end{cases} \quad (2.13)$$

и поставим уже знакомую нам по главе первой задачу: отыскать собственные значения параметра λ , т.е. те значения, при которых система (2.13) имеет ненулевые решения x_1, x_2, x_3 .

Отметим, что в третьем из уравнений (2.13) параметр λ не входит. Подобные системы, где в одно (или несколько) уравнений не входит λ , часто встречаются в механике. Уравнения, не содержащие λ , соответствуют голономным (т.е. не содержащим производных) уравнениям связи между переменными.

Как уже было обосновано в первой главе, собственные значения совпадают с корнями определителя системы (2.13):

$$\Delta = \begin{vmatrix} 1-\lambda & 1 & 2 \\ 1 & 1-\lambda & 3 \\ 1 & 1 & 0 \end{vmatrix}. \quad (2.14)$$

Вычисляя определитель (2.14) разложением по минорам последней строки, находим, что $\Delta = -5\lambda$ и, следовательно, система (2.13) имеет единственное собственное значение $\lambda_1 = 0$.

Задача вычисления собственных значений для системы (2.13) корректна. Это можно доказать наиболее простым, хотя и громоздким способом, составив проварьированный определитель Δ_ε , в котором к каждому из ненулевых коэффициентов системы (2.13) добавлена его вариация. С учетом вариаций коэффициентов определитель (2.14) принимает вид:

$$\Delta_\varepsilon = \begin{vmatrix} 1+\varepsilon_1 - (1+\varepsilon_2)\lambda & 1+\varepsilon_3 & 2(1+\varepsilon_4) \\ 1+\varepsilon_5 & 1+\varepsilon_6 - (1+\varepsilon_7)\lambda & 3(1+\varepsilon_8) \\ 1+\varepsilon_9 & 1+\varepsilon_{10} & 0 \end{vmatrix}. \quad (2.15)$$

Вычисляя определитель (2.15) разложением по минорам последней строки, находим, что (пренебрегая произведениями малых величин $\varepsilon_i \cdot \varepsilon_j$)

$$\Delta_\varepsilon = \varepsilon_3 + \varepsilon_5 - \varepsilon_1 - \varepsilon_6 + (5 + \varepsilon_2 + 2\varepsilon_4 + \varepsilon_7 + \varepsilon_9 + \varepsilon_{10})\lambda, \quad (2.16)$$

откуда определяем собственное значение

$$\lambda_1 = \frac{\varepsilon_1 - \varepsilon_3 - \varepsilon_5 + \varepsilon_6}{5 + \varepsilon_2 + 2\varepsilon_4 + \varepsilon_7 + \varepsilon_9 + \varepsilon_{10}}. \quad (2.17)$$

Мы убеждаемся, что при малых ε_i собственное значение изменя-

ется мало, задача его определения – корректна.

Однако, как известно, вычисление определителя разложением по минорам реально применяется лишь для определителей не выше четвертого-пятого порядков, поскольку с ростом порядка определителя количество необходимых вычислений возрастает необычайно быстро. Поэтому для систем линейных однородных уравнений с большим числом уравнений обычно прибегают к сокращению числа уравнений путем исключения части переменных. Покажем процесс исключения переменных на примере простой системы (2.13). Для исключения переменной x_1 берут сперва первое и второе из уравнений (2.13), домножают первое уравнение на -1 , второе – на $1-\lambda$ и складывают их. Члены, содержащие x_1 , сокращаются и получаем уравнение

$$(\lambda^2 - 2\lambda)x_2 + (1 - 3\lambda)x_3 = 0, \quad (2.18)$$

содержащее только x_2 и x_3 . Далее берут второе и третье уравнение системы, домножают второе уравнение на -1 и складывают с третьим. Члены, содержащие x_1 , снова сокращаются и мы получаем уравнение

$$\lambda x_2 - 3x_3 = 0, \quad (2.19)$$

содержащее только x_2 и x_3 . В целом, после исключения x_1 , получаем систему двух уравнений (2.18) и (2.19) с двумя переменными, определитель которой

$$\Delta = \begin{vmatrix} \lambda^2 - 2\lambda & 1 - 3\lambda \\ \lambda & -3 \end{vmatrix} = 5\lambda \quad (2.20)$$

имеет порядок меньше, чем у системы (2.13), но имеет тот же единственный корень $\lambda_1 = 0$, что и определитель (2.14).

Система (2.18)–(2.19) эквивалентна в классическом смысле системе (2.13) по отношению к задаче отыскания собственных значений и имеет то же собственное значение $\lambda_1 = 0$, что и система (2.13).

Однако для системы (2.18)–(2.19) задача отыскания собственных значений некорректна. Мы уже рассматривали систему (2.18)–(2.19) в главе первой (см. в главе первой систему (46) с определителем (50)) и убедились, что, если, например, коэффициент при x_3 в уравнение (2.19) изменится на сколь угодно малую величину и станет равным $-3(1+\epsilon)$, то система (2.18)–(2.19) будет иметь уже два собственных значения: $\lambda_1 = 0$, $\lambda_2 = 2 + \frac{5}{3\epsilon}$, причем второе значение будет существенно отличаться от первого даже при сколь угодно малом ϵ и исчезнет только при точном равенстве $\epsilon = 0$.

Таким образом, система (2.18)–(2.19) по отношению к задаче оты-

скания собственных значений эквивалентна системе (2.13) в классическом смысле и не эквивалентна – в расширенном.

С подобными примерами систем, эквивалентных в классическом смысле, но не в расширенном, приходится часто сталкиваться при исключении переменных из систем линейных однородных уравнений с параметром, при решении задачи отыскания собственных значений параметра, при вычислении частот малых колебаний различных механических и электрических систем.

Рассмотрим теперь, к каким практическим последствиям это приводит. Если все коэффициенты системы – небольшие целые числа (как у системы (2.13)), то некорректность задачи для системы небольшого порядка не опасна. Однако при последовательных умножениях одних коэффициентов на другие, необходимых при исключении переменных, начинают быстро накапливаться погрешности округления. Действительно, если, например, машина считает с точностью до восьми десятичных знаков, а коэффициенты имеют по восемь значащих цифр, то при их перемножении возникает число с шестнадцатью значащими цифрами. При округлении его до восьми цифр возникает неизбежная погрешность. В результате уже в процессе вычислений новые коэффициенты системы начинают приобретать вариации ϵ_i , не зависящие от исходных погрешностей в задании исходных коэффициентов. И если получившаяся после преобразований (например, после исключения x_1) система некорректна, то даже сколь угодно малые погрешности могут привести к неверному ответу.

Так, например, если мы будем решать задачу определения собственных значений для системы, аналогичной системе (2.13), но с коэффициентами в виде десятичных дробей с полным использованием разрядной сетки вычислительной машины и будем вычислять собственные значения путем последовательного исключения переменных, начиная с x_1 , то мы можем получить ошибочный результат – можем получить фиктивное второе собственное значение, величина которого будет зависеть от погрешности округления.

Такие же фиктивные собственные значения могут возникать и для более сложных систем – если в ходе преобразований математической модели встретятся преобразования, эквивалентные в классическом смысле, но не в расширенном.

Таким образом, была открыта еще одна возможная причина ошибок в вычислениях – изменение корректности задачи в ходе преобразований, использованных при ее решении. Если эта причина известна и осознана, то бороться с ней, получить правильный ответ относительно нетрудно. Опасно незнание, опасна неожиданная для работающего с

вычислительной техникой встреча с непонятной ему причиной ошибок. Вот почему хотя бы краткие сведения о существовании третьего класса задач математики, физики и техники – задач, промежуточных между корректными и некорректными – и краткие сведения о преобразованиях, эквивалентных в расширенном смысле, должны быть включены в учебный процесс. Более детальное изложение этих вопросов дано в монографии [17].

Изложенный материал показывает, что для суждения о корректности или некорректности задачи нужно рассматривать триаду:

1. Математическая модель.
2. Поставленная задача.
3. Применяемый метод решения.

Одна и та же задача для одной математической модели, для одной системы уравнений будет корректной, для другой модели – некорректной. Это достаточно ясно. Менее очевидно, что для одной и той же математической модели, для одной и той же решаемой в этой модели задачи, один метод решения делает задачу корректной, другой метод приводит к некорректности.

Пример: для системы (2.13) задача отыскания собственных значений методом вычисления определителя (2.14) разложением по минорам корректна, а при использовании метода последовательного исключения переменных – некорректна.

Поэтому для большей точности лучше говорить не о корректности задачи, а о корректности триады.

Мы уже упоминали, что существуют преобразования, эквивалентные в расширенном смысле и преобразования, эквивалентные в классическом смысле, но не в расширенном, т.е. преобразования, не изменяющие решений преобразованной системы, но изменяющие корректность задачи.

Как различать эти преобразования? Различение нужно проводить в составе триады. Если брать преобразования изолированно, то вряд ли можно найти преобразования, которые всегда, в составе любых триад, были бы эквивалентными в расширенном смысле. Даже, казалось бы, самые простые и «невинные» преобразования могут изменить корректность.

Вот пример: мы уже преобразовывали второе из уравнений (2.8) к новым переменным x_1, x_2, x_3, x_4 , где переменные x_3 и x_4 определяются равенствами (2.10). Рассмотрим это преобразование более подробно. Будем пользоваться только двумя видами преобразований: переносом членов из левой части в правую с изменением знака и разбиением членов, т.е., например, вместо $4Dx_1$ будем писать $2Dx_1 + 2Dx_1$.

После таких совершенно «невинных» и повсеместно используемых преобразований второе из уравнений (2.8) принимает вид:

$$[(D^2 + 2D)x_1 - Dx_2] + [(2D + 4)x_1 - 2x_2] + x_2 + x_1 = 0.$$

Теперь в первой квадратной скобке мы узнаем (с учетом равенств (2.10)) переменную x_4 , во второй квадратной скобке узнаем член $2x_3$, а в целом второе из уравнений (2.8) переходит в последнее из уравнений (2.11). Но мы уже убедились, что преобразования уравнений (2.8) в уравнения (2.11) изменяет корректность. Таким образом, даже простые переносы членов из левой части в правую с изменением знака могут оказаться эквивалентными в классическом смысле, но не в расширенном. Все зависит от свойств конкретной триады.

Свойства преобразований, эквивалентных в расширенном смысле, к сожалению, более сложны, чем простые свойства и правила преобразований, эквивалентных в классическом смысле и изучаемых еще в средней школе.

Если какое либо преобразование (например, умножение всех членов на число, не равное нулю) эквивалентно в классическом смысле, то оно будет эквивалентно всегда, во всех задачах, для всех уравнений. Одно и то же преобразование для одной системы эквивалентно в расширенном смысле, для другой – нет. Эквивалентность в расширенном смысле можно исследовать только в составе уже упоминавшейся триады: 1) математическая модель; 2) рассматриваемая задача; 3) метод решения.

Утверждения об эквивалентности в расширенном смысле могут быть лишь такого вида: для такой-то задачи и такой-то математической модели такое-то преобразование будет (или не будет) эквивалентным в расширенном смысле. Примеры мы далее приведем.

2.6. Приложения к системам управления и другим объектам, описываемым дифференциальными уравнениями

Обнаружение различий между преобразованиями эквивалентными в классическом смысле и в расширенном позволило более правильно подойти к расчету параметрической устойчивости решений систем дифференциальных уравнений, запасов устойчивости при вариациях параметров, других тонких вопросов, возникающих в теории управления, и избежать ошибок.

Пример. Уравнения одного из объектов управления – электропривода постоянного тока, работающего на исполнительный механизм, у которого колебания момента сопротивления имеют спектр

$$S = \frac{1}{(1+\omega^2)^2}, \quad (2.21)$$

могут быть, как известно (смотри, например, [14]), записаны в виде следующей системы дифференциальных уравнений:

$$\begin{cases} \dot{x}_1 = -2x_1 + x_2 + x_3, \\ \dot{x}_3 = x_4, \\ \dot{x}_4 = -x_3 - 2x_4, \end{cases} \quad (2.22)$$

где x_1 – отклонение частоты вращения от номинального значения, x_2 – управление (отклонение тока якоря от номинального значения), x_3 – отклонение момента сопротивления от номинального значения, x_4 – производная этого отклонения; независимой переменной является время t . Последние два уравнения системы (2.22) являются уравнениями так называемого «формирующего фильтра», формирующего спектр возмущающего воздействия (2.21) из «белого шума». В целом уравнения (2.22) совпадают с уже рассмотренными первыми тремя из уравнений (2.11).

Если управляющее воздействие формируется в функции переменных x_1, x_3, x_4 по закону

$$x_2 = -x_1 - 2x_3 - x_4, \quad (2.23)$$

то структурная схема системы управления имеет вид, показанный на рис. 2.1.

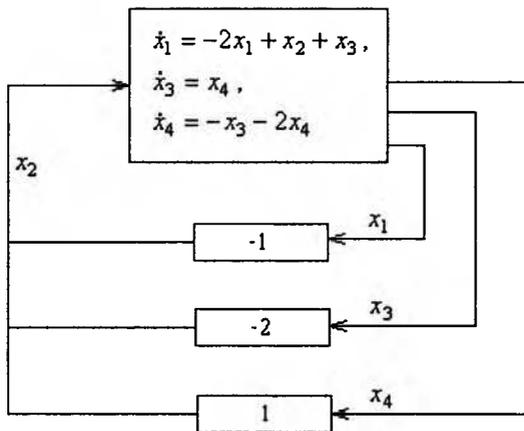


Рис. 2.1. Структурная схема объекта управления (2.22), замкнутого обратной связью (2.23)

Если переменные x_3 и x_4 , как это часто бывает, непосредственно неизмеримы, то можно исключить их из уравнения (2.23) путем эквивалентных (в классическом смысле) преобразований. После этого уравнение (2.23) принимает вид

$$(D^2 + 4D + 5)x_1 = (D + 1)x_2. \quad (2.24)$$

Теперь рассмотрим, какой будет структурная схема системы управления, в которой управляющее воздействие формируется на основе измерения одной переменной x_1 согласно уравнению (2.24), но переходные процессы (решения $x_1(t)$, $x_2(t)$, $x_3(t)$, $x_4(t)$) остаются теми же самыми, что и ранее. Новая структурная схема принимает вид, показанный на рис. 2.2.

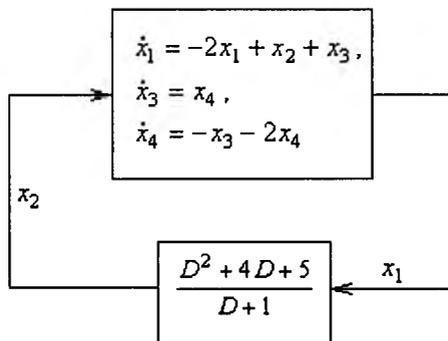


Рис. 2.2. Структурная схема объекта управления (2.22), замкнутая обратной связью (2.24)

Уже один взгляд на структурные схемы, изображенные на рис.2.1 и рис.2.2, сразу показывает, что системы уравнений (2.22)–(2.23) и (2.22)–(2.24) не тождественны друг другу. Они эквивалентны (в классическом смысле), но не тождественны (на рис.2.1 и рис.2.2 это сразу видно) и не эквивалентны в расширенном смысле. Они имеют одинаковые решения, но параметрическая устойчивость решений у этих систем различна. Введение нового математического понятия – эквивалентности в расширенном смысле – как раз и позволяет объяснить, почему при преобразованиях, эквивалентных в классическом смысле, но не в расширенном, изменяются некоторые тонкие, но важные свойства систем управления – такие, как параметрическая устойчивость.

Поэтому при исследовании этих свойств уже нельзя без дополнительных проверок пользоваться традиционными методами расчета и

привычными эквивалентными преобразованиями уравнений. Нужно дополнительно проверять – были ли использованные преобразования эквивалентны не только в классическом смысле, но и в расширенном (или же проверять корректность после каждого преобразования). Если этих дополнительных проверок не делать и продолжать некритически пользоваться традиционными методами, то возможны опасные ошибки в расчетах, которые могут стать, как уже говорилось, причиной неожиданных аварий и даже катастроф, подробнее рассмотренных в монографии [17].

Будут результаты расчета ошибочными или нет – зависит от того, какая из эквивалентных между собой в классическом смысле математических моделей лучше отражает свойства реального объекта. Если реальные соотношения между регулируемымыми переменными и управлением в системе управления электроприводом лучше описывает система уравнений (2.22)–(2.24), то расчет по системе уравнений (2.22)–(2.23) дает неверный ответ о параметрической устойчивости реальной системы управления, хотя обе системы (2.22)–(2.23) и (2.22)–(2.24) эквивалентны в классическом смысле, и для расчетов устойчивости (не параметрической устойчивости, а просто устойчивости по Ляпунову) их можно использовать равноправно, одну вместо другой.

Отметим еще одно обстоятельство, которое часто вызывает вопросы: проводя эквивалентные преобразования какой-либо системы уравнений, мы приходим к новой системе с совсем другими коэффициентами. Так, исключив x_2 из системы уравнений

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

мы приходим к уравнению $(a_{12}a_{21} - a_{11}a_{22})x_1 = a_{12}b_2 - a_{22}b_1$. Вместо четырех коэффициентов исходной системы у нас осталось два.

Можно ли, прибавляя сколь угодно малые вариации к коэффициентам преобразованной системы и исследуя их влияние на решения, считать, что мы получили правильный ответ на вопрос о влиянии на решения вариаций исходной системы?

Ответ не сложен: коэффициенты преобразованной системы являются функциями коэффициентов исходной системы, например, первый коэффициент преобразованной системы a_{1np} является некоторой функцией $k_1(a_1, a_2, \dots, a_n)$ от коэффициентов a_1, a_2, \dots, a_n исходной системы, т.е. $a_{1np} = k_1(a_1, a_2, \dots, a_n)$; аналогично $a_{2np} = k_2(a_1, a_2, \dots, a_n)$ и т.д. Здесь возможны два варианта: при номинальных значениях коэффициентов исходной системы эти функции:

- 1) непрерывны,

2) имеют разрыв.

В первом случае сколь угодно малым изменениям коэффициентов преобразованной системы соответствуют столь же малые изменения коэффициентов у исходной системы. Поэтому проверку корректности можно вести и по исходной системе, и по преобразованной – разумеется, если использованные преобразования были эквивалентны в расширенном смысле. В противном случае, как мы уже показывали, возможны ошибки.

Во втором случае сколь угодно малым изменениям коэффициентов исходной системы могут соответствовать конечные и даже большие изменения коэффициентов преобразованной системы. В этом случае сразу видно, что для исходной системы рассматриваемая задача была некорректной.

Пр и м е р . Систему уравнений

$$\begin{aligned}\ddot{x}_1 &= -\dot{x}_2 - x_2, \\ \dot{x}_2 &= -m\ddot{x}_1 + e^{-t}\end{aligned}$$

с параметром m и нулевыми начальными условиями: $x_1(0) = \dot{x}_1(0) = x_2(0) = \dot{x}_2(0) = 0$ можно путем введения новой переменной $x_3 = \dot{x}_1$ привести к нормальной форме Коши:

$$\begin{aligned}\dot{x}_1 &= x_3, \\ \dot{x}_2 &= \frac{m}{1-m}x_2 + \frac{1}{1-m}e^{-t}, \\ \dot{x}_3 &= -\frac{1}{1-m}x_2 - \frac{1}{1-m}e^{-t}.\end{aligned}$$

Сразу видно, что при значениях параметра $m=1$ непрерывная зависимость решения от параметра исчезает. Но это означает, что и в исходной системе уравнений при $m=1$ нет непрерывной зависимости решений от параметра m .

Это простейший и легко распознаваемый случай. Ранее мы привели более «коварные» примеры, когда в системе, приведенной к нормальной форме Коши, решения зависят от параметра непрерывно, а в исходной, не преобразованной, системе уравнений непрерывной зависимости решений от параметра все-таки не было.

Введение понятия о преобразованиях, эквивалентных в расширенном смысле, позволяет окончательно закрыть старую дискуссию о возможности сокращения одинаковых множителей в передаточных функциях систем управления. Пусть, например, передаточная функция имеет вид

$$\bar{u} = \frac{s+1}{(s+1)(s+2)} \bar{x}, \quad (2.25)$$

где s – символ преобразования по Лапласу. Допустимо ли сокращение на $s+1$? С одной стороны, такое сокращение является эквивалентным преобразованием, решений $u(t)$ оно не меняет, а вычисления упрощает. С другой стороны, использование подобных сокращений при анализе сохранения устойчивости после вариаций параметров иногда приводило к ошибочным результатам и это настораживало. Высказывались мнения о недопустимости сокращения в передаточных функциях, хотя это и усложняло вычисления.

После введения понятия о преобразованиях, эквивалентных в расширенном смысле, все стало ясным: сокращение числителя и знаменателя передаточной функции на равные множители является преобразованием, эквивалентным в классическом смысле, но совсем не обязательно – в расширенном. Если нас интересует вычисление решения, вычисление функции $u(t)$ для номинальных значений коэффициентов, то сокращение, безусловно, допустимо и существенно упрощает вычисления. Если же нас интересует параметрическая устойчивость, то сокращение может быть недопустимым.

То же самое относится и к любым объектам, математической моделью которых являются системы обыкновенных дифференциальных уравнений различных порядков. Если нас интересует только вычисление решений при номинальных значениях коэффициентов, то мы можем спокойно применять эквивалентные в классическом смысле преобразования и приводить исследуемые системы уравнений к нормальной форме Коши, поскольку такое преобразование позволяет использовать стандартное программное обеспечение, стандартные программы, составленные, естественно, для систем в нормальной форме. Если же мы хотим проследить, как влияют на поведение решений неизбежные малые неточности в значениях коэффициентов, полученных из опыта или измерения, то здесь использование преобразований, эквивалентных в классическом смысле, но не в расширенном, может дать неверный ответ. Так, для системы (2.22)–(2.24) при сколь угодно малых отклонениях коэффициентов при D^2x_1 и Dx_2 от номинальных значений поведение решений изменится очень резко, а в эквивалентной ей (в классическом смысле) системе в нормальной форме (2.22)–(2.23) при малых отклонениях любых коэффициентов от номинальных значений решения изменятся мало.

Для избежания ошибок исследование зависимости решений от почти неизбежных неточностей в задании коэффициентов нужно проводить по первичной, исходной, непосредственно вытекающей из за-

конов физики и математики математической модели, а если математическая модель все же для удобства исследования преобразуется, то нужно следить, чтобы преобразования были эквивалентными не только в классическом, но и в расширенном смысле.

Так, если мы систему, структурная схема которой показана на рис.2.2 и первичными уравнениями которой являются уравнения (2.22)–(2.24), запишем в нормальной форме, в форме уравнений (2.22)–(2.23), то результаты расчета – с учетом неточностей в задании коэффициентов – могут быть ошибочными. По расчету в нормальной форме система (2.23)–(2.24) параметрически устойчива, а на самом деле параметрической устойчивости нет.

Теперь отметим важность различия некорректных и плохо обусловленных задач (напомним – некорректными мы называем задачи, в которых конечные и даже большие изменения решений происходят при сколь угодно малых изменениях коэффициентов, параметров, начальных или граничных условий и т.п., а плохо обусловленными называем задачи, где конечные (и часто большие) изменения решений происходят при малых, но обязательно конечных изменениях коэффициентов, параметров и т.п.). Определение «плохо обусловленных» задач неизбежно носит расплывчатый характер, поскольку вопрос о том, какие именно изменения коэффициентов и решений считать «малыми» и какие «существенными» и «большими» зависит от особенностей конкретной задачи, от точности измерений и измерительных инструментов, требований к точности решений и т.п.

Отличие от этого некорректные задачи имеют точное определение (приведенное в первой главе) и выделить их проще, чем задачи плохо обусловленные.

Примером может служить система, состоящая из уравнения (2.24) и уравнения

$$(D^3 + 4D^2 + 5D + 2)x_1 = (kD^2 + 2D + 1)x_2. \quad (2.26)$$

При $k=1$ вычисление решений системы (2.24)–(2.26) и проверка ее устойчивости являются задачами некорректными: при сколь угодно малом отклонении коэффициента k от значения $k=1$ решения $x_1(t)$ и $x_2(t)$ для любых $t > 0$ могут измениться на любые величины; при $k=1$ система (2.24)–(2.26) асимптотически устойчива, при $k = 1 + \varepsilon$ она теряет устойчивость при сколь угодно малом ε , если $\varepsilon < 0$.

Но все эти неприятности зато относительно легко предвидеть, поскольку после исключения, например, переменной x_2 из системы (2.24)–(2.26) она перейдет в уравнение:

$$[(k - 1)D^4 + (4k - 3)D^3 + 5kD^2 + 7D + 3]x_1 = 0, \quad (2.27)$$

которое как раз при $k=1$ понижает порядок. Понижение порядка сразу сигнализирует об опасности и может свидетельствовать о некорректности задачи вычисления решений и проверки их устойчивости при $k=1$.

Теперь рассмотрим ту же систему (2.24)–(2.26) при $k=1.001$. После исключения x_2 система перейдет, как и полагается, в уравнение четвертого порядка:

$$(0.001D^4 + 1.004D^3 + 5.005D^2 + 7D + 3)x_1 = 0. \quad (2.28)$$

Понижения порядка нет, нет, следовательно, и «сигнала тревоги». Действительно, система (2.24)–(2.26) при $k=1.001$ устойчива и параметрически устойчива. Задача проверки устойчивости для системы (2.24)–(2.26) при $k=1.001$ – корректна; при сколь угодно малом отклонении коэффициента k от значения $k=1.001$ решения системы (2.24)–(2.26) изменятся на сколь угодно малые величины и останутся устойчивыми. Однако запас устойчивости по вариации параметра k у системы (2.24)–(2.26) при $k=1.001$ очень мал: достаточно коэффициенту k измениться на 0.11% и система (2.24)–(2.26) потеряет устойчивость. Поскольку гарантировать, что в ходе реальной эксплуатации вариации параметра k не превысят 0.11% обычно очень трудно, то с практической точки зрения систему (2.24)–(2.26) при $k=1.001$ следует считать параметрически неустойчивой, а точнее – имеющей недопустимо малый запас устойчивости по вариациям параметра k . В то же время предупреждения об опасности в виде понижения порядка уравнения, остающегося после исключения x_2 или x_1 , в этом случае не будет. Более того, вычисляя корни характеристического полинома системы (2.24)–(2.26) при $k=1.001$, получим, что они (с точностью до первой значащей цифры) равны $\lambda_1 = -3$, $\lambda_2 = \lambda_3 = -1$, $\lambda_4 = -1000$. Убеждаемся, что все корни характеристического полинома лежат в левой полуплоскости комплексного переменного далеко слева от мнимой оси и поэтому согласно традиционным методам проверки систему (2.24)–(2.26) при $k=1.001$ следовало бы признать устойчивой и обладающей хорошими запасами устойчивости. На самом деле это не так, но традиционное исследование характеристического полинома не подает сигнала опасности. Это связано с тем, что система (2.24)–(2.26) при $k=1.001$ не является некорректной, является лишь плохо обусловленной, но исследовать и проверять такие системы сложнее, чем системы некорректные и использовать для этого надо совсем другие методы.

В то же время выделение некорректных задач полезно и для решения задач плохо обусловленных. Если при каких-то конкретных значениях коэффициентов и параметров задача некорректна, то при коэффициентах и параметрах близких к этим критическим значениям зада-

ча очень часто оказывается плохо обусловленной. Близость к некорректным задачам позволяет выделить задачи плохо обусловленные.

2.7. Приложения к практике вычислений

Обнаружение различия между преобразованиями, эквивалентными в классическом смысле и в расширенном позволяет также раскрыть одну из причин ошибок в вычислениях. Эта причина заключается в том, что в ходе преобразований математической модели мы можем изначально корректную модель незаметно преобразовать в некорректную – и тогда любая сколь угодно малая погрешность округления при вычислениях сразу может привести к неверному ответу.

В качестве примера рассмотрим обобщенную задачу вычисления собственных значений матрицы и рассмотрим систему:

$$\left. \begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \dots + a_{2n}x_n &= 0, \\ \dots & \\ a_{n-1,1}x_1 + \dots + (a_{n-1,n-1} - \lambda)x_{n-1} + a_{n-1,n}x_n &= 0, \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= 0, \end{aligned} \right\} \quad (2.29)$$

т.е. систему линейных однородных уравнений с параметром λ , причем во все первые $n-1$ уравнений параметр λ входит, а в последнее – не входит. Ставится задача – найти собственные значения параметра λ , при которых система (2.29) имеет ненулевые решения.

Системы, подобные системе (2.29) с разным числом уравнений, не содержащих λ , встречаются в задачах нахождения частот собственных колебаний различных механических и электромеханических систем, систем управления и т.п. (Уравнения, не содержащие λ , отражают голономные связи между переменными, такие связи часто встречаются).

Предположим, что мы решили вычислять собственные значения путем последовательного исключения переменных из системы (2.29) путем домножений и сложений (заметим сразу, что существуют различные методы вычисления собственных значений; мы не рассматриваем сравнительных достоинств и недостатков различных методов, а сосредоточим свое внимание на том, что может произойти в ходе использования одного из возможных методов).

После исключения $n-2$ переменных, перед последним шагом исключения, мы приходим к системе двух уравнений:

$$\begin{aligned} A_1 x_{n-1} + A_2 x_n &= 0, \\ A_3 x_{n-1} + A_4 x_n &= 0, \end{aligned} \quad (2.30)$$

где A_1, \dots, A_4 – полиномы от переменной λ . Опираясь на формулы

Крамера, можно установить, что эти полиномы будут равны следующим определителям $n-1$ порядка:

$$A_1 = \begin{vmatrix} a_{11} - \lambda & \dots & a_{1n-1} \\ \dots & \dots & \dots \\ a_{n-1;1} & \dots & (a_{n-1;n-1} - \lambda) \end{vmatrix}, \quad (2.31)$$

$$A_2 = \begin{vmatrix} a_{11} - \lambda & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n-1;1} & \dots & a_{n-1;n} \end{vmatrix}, \quad (2.32)$$

$$A_3 = \begin{vmatrix} a_{21} & \dots & a_{2n-1} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn-1} \end{vmatrix}, \quad (2.33)$$

$$A_4 = \begin{vmatrix} a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}. \quad (2.34)$$

Таким образом, первый определитель составлен из коэффициентов, стоящих в первых $n-1$ уравнениях системы (2.29) при первых ее $n-1$ переменных – от x_1 до x_{n-1} . Определитель (2.32) будет равен определителю (2.31), в котором последний столбец заменен на столбец коэффициентов, стоящих в системе (2.29) перед переменной x_n .

Аналогично, определитель (2.33) составлен из коэффициентов, стоящих в системе (29), начиная со второй ее строки, перед переменными с индексами от x_2 до x_{n-1} , а определитель (2.34) отличается от определителя (2.33) тем, что последний столбец в нем заменен на столбец коэффициентов при переменной x_n .

Из системы двух уравнений (2.30) сразу видно, что собственные значения системы (2.29) будут находиться среди корней полинома

$$\Delta = A_1 A_4 - A_2 A_3. \quad (2.35)$$

Разлагая определители (2.31)–(2.34) по минорам, нетрудно будет выписать члены со старшими степенями параметра λ . Имеем:

$$\begin{aligned} A_1 &= (-1)^{n-1} \lambda^{n-1} + \dots, \\ A_2 &= (-1)^{n-2} a_{n-1,n} \lambda^{n-2} + \dots, \\ A_3 &= (-1)^{n-2} a_{n,1} \lambda^{n-2} + \dots, \\ A_4 &= (-1)^{n-3} (a_{n-1,n} a_{n,1} - a_{n-1,n} a_{n,n}) \lambda^{n-3} + \dots \end{aligned} \quad (2.36)$$

Подставляя полиномы (2.36) в (2.35), получим

В результате будем иметь $\Delta = a_{n-1,n}\lambda^{2n-3} \dots$, где точками обозначены члены более низких степеней, и поэтому малые ошибки округления уже не приведут к потере корректности. Поскольку при счете вручную вычисление собственных значений удобно начинать с уравнений, не содержащих λ , и пользуясь ими, исключить часть переменных, то после их исключения мы придем к классической проблеме собственных значений, т.е. к системе, состоящей из меньшего числа уравнений, но теперь уже параметр λ будет входить в каждое уравнение, и поэтому при дальнейших исключениях переменных изменений корректности не произойдет.

При вычислениях на программируемых вычислительных машинах важнее всего унификация программы и поэтому в системе, подобной системе (2.29) машина может начать исключать переменные в порядке их индексов и поэтому может встретиться с потерей корректности, т.е. с тем, с чем в эпоху ручного счета встречаться не приходилось.

Это обстоятельство подчеркивает необходимость тщательной проверки вычислительных алгоритмов при переходе на машинные вычисления. Тонкости, несущественные при ручном счете, могут стать причиной ошибок при машинных вычислениях.

Возвращаясь к проблеме различения преобразований, эквивалентных в классическом смысле и в расширенном, мы можем сказать, что эквивалентные в классическом смысле преобразования (домножения и сложения), используемые при последовательном исключении переменных в классической задаче о вычислении собственных значений матрицы эквивалентны и в расширенном смысле. Те же преобразования, используемые при последовательном исключении переменных в другой математической модели – в обобщенной задаче о собственных значениях (когда некоторые из уравнений не содержат параметра λ) уже не будут эквивалентными в расширенном смысле и могут изменить корректность решаемой задачи.

Это еще раз подчеркивает, что при рассмотрении корректности и ее возможного изменения в ходе решения нужно рассматривать триаду: решаемую задачу, математическую модель, метод решения.

Более подробно проблемы, возникающие в различных вариантах обобщенной задачи о собственных значениях, рассмотрены в монографии [17]. Там же изложен и метод «матриц степеней», позволяющий выявить задачи третьего класса, у которых в ходе преобразований, используемых при их решении, может изменяться корректность.

Если причина изменения корректности известна, то справиться с ней, избежать ошибок, совсем нетрудно. Опасно незнание, опасна неожиданная встреча с неизвестным ранее явлением. Нужно помнить о

следующей возможности: в ходе последовательных эквивалентных (в классическом смысле) преобразований математической модели, используемых в ходе вычислений, возможно изменение корректности решаемой задачи. Если такое изменение произошло, то даже сколь угодно малая погрешность округления может привести к ошибке.

Рассмотрим еще один пример, относящийся к вычислению функций, для которых известно их изображение по Лапласу. Пусть дано изображение

$$\bar{f}(s) = \frac{s+b}{(s+a)(s+b)}, \quad (2.39)$$

и нужно вычислить оригинал, т.е. функцию $f(t)$. Согласно известным рекомендациям операционного исчисления, изображение (2.39) разлагаем на простые дроби:

$$\frac{s+b}{(s+a)(s+b)} = \frac{A}{s+a} + \frac{B}{s+b} = \frac{As+Ab+Bs+Ba}{(s+a)(s+b)}, \quad (2.40)$$

а поскольку числители должны быть равны, то это дает два уравнения для определения двух неизвестных чисел A и B :

$$\begin{aligned} A+B &= 1, \\ Ab+Ba &= b, \end{aligned} \quad (2.41)$$

из которых находим $B=0$, $A=1$. Следовательно,

$$\bar{f}(s) = \frac{1}{s+a}, \quad f(t) = e^{-at}. \quad (2.42)$$

Вычисления можно упростить, числитель и знаменатель можно сократить на $s+b$ (т.е., собственно, умножить на число $\frac{1}{s+1}$, не равное нулю). Тогда сразу получаем равенства (2.42). Проведенное нами разложение на простые дроби (2.40) лишний раз подтвердило, что сокращение числителя и знаменателя изображения по Лапласу на равные множители является преобразованием, эквивалентным в классическом смысле. Будет ли оно эквивалентным в расширенном смысле? Прежде чем ответить на этот вопрос, уточним смысл утверждения: «в числителе и знаменателе изображения (2.39) присутствуют равные множители $s+b$ ». Числа « b », входящие в них, получены (в конечном счете) из опыта или измерения и мы можем быть уверены лишь в конечном числе десятичных знаков числа b . Пусть величина коэффициента « b » измерена хорошим прибором с точностью до четвертого знака и все четыре знака в числителе и знаменателе равны. Несмотря на равенство всех четырех знаков, мы можем лишь утверждать, что изображение (2.39) равно

$$\bar{f}(s) = \frac{s + b_0 + \varepsilon_1}{(s + a)(s + b_0 + \varepsilon_2)}, \quad (2.43)$$

где ε_1 и ε_2 – неизвестные числа, относительно которых известно лишь что $\frac{\varepsilon_1}{b_0} < 10^{-4}$ и $\frac{\varepsilon_2}{b_0} < 10^{-4}$. Разлагая изображение (2.43) на простые дроби

$$\frac{s + b_0 + \varepsilon_1}{(s + a)(s + b_0 + \varepsilon_2)} = \frac{A}{s + a} + \frac{B}{s + b_0 + \varepsilon_2} \quad (2.44)$$

и вычисляя числа A и B с учетом неизвестных ε_1 и ε_2 , убеждаемся, что

$$A = \frac{b_0 - a + \varepsilon_1}{b_0 - a + \varepsilon_2}, \quad B = \frac{\varepsilon_2 - \varepsilon_1}{b_0 - a + \varepsilon_2}, \quad (2.45)$$

т.е. в действительности число B не обязательно равно нулю и истинный оригинал $f_1(t)$ может быть равен не $f(t) = e^{-at}$, а

$$f_1(t) = e^{-at} + \frac{\varepsilon_2 - \varepsilon_1}{b_0 - a + \varepsilon_2} e^{-at} + \frac{\varepsilon_2 - \varepsilon_1}{b_0 - a + \varepsilon_2} e^{-(b_0 + \varepsilon_2)t}. \quad (2.46)$$

Пусть, например, $a = 1, b_0 = -2$ и $\varepsilon_2 - \varepsilon_1 \leq 10^6$; тогда при $t=1$ разность между $f_1(t)$ и $f(t)$ не превысит 10^{-5} , но уже при $t=20$ они могут различаться во много раз. Если же $a > 0$ и $b > 0$, то различие между $f_1(t)$ и $f(t)$ мало для любых t .

Сокращение числителя и знаменателя изображений по Лапласу является примером преобразования, эквивалентного в классическом смысле, но не в расширенном. Сокращение на равные множители в изображениях часто используется для упрощения вычислений, и в то же время может приводить к ошибкам. В руководствах по теории управления неоднократно делались предупреждения о том, что к сокращению одинаковых множителей в числителе и знаменателе изображений по Лапласу передаточных функций следует подходить осторожно. Мы убеждаемся, что эти предостережения были совершенно обоснованы, и мы видим теперь их истинную причину.

2.8. Заключение по главам 1 и 2

В первой и второй главах мы рассмотрели ту часть некорректных задач, которая обычно находится за пределами внимания исследователей. Чаще всего рассматривают гораздо более сложные некорректные задачи, требующие решения дифференциальных уравнений в частных производных, интегральных уравнений и т.п. Сложность этих задач не всегда позволяет выявить и выделить принципиальные вопросы.

В главах первой и второй мы убедились, что большое число более простых некорректных задач возникает в самых разнообразных областях приложений и для того, чтобы избежать ошибок, инженер должен быть знаком хотя бы с самыми простыми свойствами некорректных задач и методами подхода к их решению.

Важно отметить также, что некорректные задачи являются частным, предельным случаем более обширного, но гораздо менее определенного класса задач плохо обусловленных – задач, в которых существенные изменения решений порождаются малыми (но конечными) вариациями коэффициентов, параметров, начальных и граничных условий и т.п. В первой и второй главах мы убедились, что (в отличие от плохо обусловленных задач) для задач некорректных можно найти простые условия выявления некорректности, что сразу облегчает решение. А уже вокруг некорректных задач располагаются более сложно выделяемые плохо обусловленные задачи. После выделения некорректных задач уже проще подходить к задачам плохо обусловленным.

Другим отличием нашего подхода является выделение в особый, третий, класс тех задач, которые нельзя отнести ни к корректным, и к некорректным, поскольку они могут менять свою корректность в ходе эквивалентных преобразований, используемых при решении.

Наибольшее значение для самого широкого круга исследователей и инженеров, производящих расчеты, имеет недавнее обнаружение систем дифференциальных уравнений, не имеющих непрерывной зависимости решений от коэффициентов и параметров. До последнего времени огромная армия инженеров и научных работников, использующих в своих расчетах решения дифференциальных уравнений, не заботилась о проверке корректности решений, спокойно преобразовывала уравнения к нормальной форме Коши для того, чтобы пользоваться стандартными программами и считала, что надежность расчетов гарантируется теоремой о непрерывности решения по параметрам. Теперь мы видим, что это не так, что возможны особые случаи, в которых использование традиционных методов исследований и расчета приводит к грубым ошибкам (а следствием ошибок могут быть опасные аварии). Для избежания ошибок следует использовать несложные дополнительные проверки, описанные в главе второй (а более подробно – в монографии [17], в третьем ее издании).

Материал главы второй показывает, что выделение и исследование недавно обнаруженного третьего класса задач математики, физики и техники позволяет избежать ошибок в расчетах.

Особые случаи, особые системы дифференциальных уравнений, для которых применение обычных, традиционных методов решения приводит к серьезным ошибкам, встречаются не очень часто и поэтому

до последних лет их существование не замечалось.

Однако каждая неожиданная встреча с подобной особой системой может привести к ошибке в расчетах и может стать причиной аварий и даже катастроф (как уже не раз становилось, смотри [17], стр. 21–23, 107–108). Поэтому на особые случаи, на особые системы надо обращать самое серьезное внимание, тем более, что выделить особые системы совсем не трудно.

То же самое относится ко многим расчетам и методам расчета, использующим цепочки эквивалентных преобразований. И здесь неожиданная встреча с задачами, относящимися к третьему классу, классу задач, промежуточных между корректными и некорректными, может приводить (и приводит!) к грубым ошибкам в расчетах и может становиться причиной аварий. Примеры мы приводили. Поэтому задачам третьего класса надо уделять самое серьезное внимание.

Особо отметим, что исследование задач третьего класса, исследование преобразований математических моделей, эквивалентных не только в классическом смысле, но и в расширенном, исследование преобразований, изменяющих корректность решаемой задачи, – все эти исследования еще только начинаются и здесь открыто поле для интересной и плодотворной научно-исследовательской работы.

Глава 3

ИЗМЕНЕНИЕ ЧУВСТВИТЕЛЬНОСТИ К ПОГРЕШНОСТЯМ ИЗМЕРЕНИЯ ПРИ ИНТЕГРАЛЬНЫХ ПРЕОБРАЗОВАНИЯХ, ИСПОЛЬЗУЕМЫХ ПРИ РАСЧЕТЕ СУДОВ И СИСТЕМ СУДОВОЙ АВТОМАТИКИ

3.1. Применение интегральных преобразований для решения практических задач

В предыдущих главах мы рассмотрели влияние погрешностей измерения на точность решения различных технических задач. С особым вниманием мы рассмотрели острый случай некорректных задач, когда даже сколь угодно малые погрешности ведут к большим изменениям решений. Мы убедились, что существует целый класс задач, промежуточных между корректными и некорректными, таких задач, корректность которых, а значит и чувствительность к погрешностям, может изменяться в ходе эквивалентных преобразований, используемых при их решении. Такие задачи особенно трудны для решения и неожиданная для инженера встреча с подобной задачей может привести к грубым ошибкам в расчетах.

В настоящей главе мы рассмотрим изменения чувствительности к погрешностям измерения при интегральных преобразованиях. Особое внимание мы уделим известному «косинус преобразованию Фурье»

$$S_x(\omega) = \frac{2}{\pi} \int_0^{\infty} K_x(\tau) \cos \omega \tau d\tau, \quad (3.1)$$

которое преобразует функцию времени $K(\tau)$ в функцию частоты $S(\omega)$. Косинус-преобразование Фурье (3.1) широко (хотя все же еще недостаточно широко) применяется в расчетах систем судовой автоматики. Без него в расчетах судов и судовых систем никак не обойтись.

Действительно, возмущающие воздействия на суда и судовое системы порождаются ветром, морским волнением, вибрациями, предсказать величину которых мы принципиально не можем.

Рассмотрим одну из простейших задач, расчет бортовой качки судна, расположенного, как говорят моряки, «лагом к волне», т.е. гребни волны параллельны диаметральной плоскости судна. Уравнение бортовой качки несложно, оно имеет вид

$$(T_1^2 D^2 + T_2 D + 1)\theta = \varphi(t), \quad (3.2)$$

где T_1 и T_2 – постоянные времени, измеряемые в секундах, $D = \frac{d}{dt}$ – оператор дифференцирования, θ – угол крена, измеряемый в градусах, $\varphi(t)$ – возмущающее воздействие, угол волнового склона, измеряемый в градусах. Постоянная времени T_1 зависит от момента инерции корпуса судна относительно его продольной оси, постоянная T_2 отражает демпфирующее воздействие воды.

Постоянные T_1 и T_2 легко определяются экспериментально, обработкой записей свободных колебаний судна на тихой воде, т.е. при $\varphi(t) = 0$. Если при $\varphi(t) = 0$ придать судну начальный крен θ_0 , то изменение крена в функции времени будет иметь вид

$$\theta(t) = e^{-\frac{T_2}{2T_1^2}t} \left[c_1 \cos \frac{\sqrt{1 - \left(\frac{T_2}{4T_1}\right)^2} t}{T_1} + c_2 \sin \frac{\sqrt{1 - \left(\frac{T_2}{4T_1}\right)^2} t}{T_1} \right], \quad (3.3)$$

где c_1 и c_2 – постоянные интегрирования, определяемые из начальных условий. Если при $t = 0$ будет $\theta = \theta_0$, а $\dot{\theta} = 0$, то $c_1 = \theta_0$, $c_2 = 0$. По-

скольку обычно $T_2 \ll T_1$, то $\sqrt{1 - \left(\frac{T_2}{4T_1}\right)^2} \approx 1$ и частота свободных колебаний судна близка к величине $1/T_1$. Отношение $\frac{T_2}{2T_1}$ характеризует быстроту затухания свободных колебаний.

Если возмущающее воздействие $\varphi(t)$ известно, то решить уравнение (2) и найти $\theta(t)$ несложно. Если, например, возмущающее воздействие является гармоническим $\varphi(t) = A \sin \beta t$, то после быстрого затухания свободных колебаний останутся вынужденные колебания, которые будут гармонической функцией с частотой β и амплитудой, зависящей от β и от параметров судна, от постоянных T_1 и T_2 по следующему простому закону

$$\theta(t) = \frac{A}{\sqrt{(1 - T_1^2 \beta^2)^2 + T_2^2 \beta^2}} \sin(\beta t + \psi_0). \quad (3.4)$$

Однако реальные возмущающие воздействия на суда и судовые системы не являются гармоническими колебаниями, они являются

случайными, не полностью предсказуемыми функциями, и в этих условиях прямое решение уравнения (3.2) ничего не дает: решение $\theta(t)$ тоже будет случайной, не полностью предсказуемой функцией, бесполезной для расчета (типичный пример изменения крена судна θ в функции времени при качке на нерегулярном морском волнении будет показан позже на рис. 3.8). Решая дифференциальное уравнение (3.2) при случайной правой части $\varphi(t)$, мы не решим главной задачи расчета качки, не сможем предсказать, перевернется корабль или нет, превысит ли угол крена предельный, при достижении которого корабль опрокидывается, или не превысит.

В этих условиях нужно изменить сам подход к задаче; надо искать не само решение $\theta(t)$, а его средний квадрат $\langle \theta^2 \rangle$:

$$\langle \theta^2 \rangle = \lim_{T \rightarrow \infty} \int_0^T \theta^2 dt, \quad (3.5)$$

и среднеквадратичное значение σ_θ :

$$\sigma_\theta = \sqrt{\langle \theta^2 \rangle}. \quad (3.6)$$

Поскольку возмущающие воздействия распределены почти всегда по нормальному закону, то согласно известному правилу «трех среднеквадратичных значений», максимальное значение переменной $\theta(t)$ равно трем ее среднеквадратичным значениям.

Если, например, опасным для судна, ведущим к его опрокидыванию, значением угла крена является $\theta_{кр} = 40^\circ$, а $\sigma_\theta = 12^\circ$, то мы можем быть уверены, что $\theta_{max} \leq 3 \cdot 12^\circ = 36^\circ$, и судно не опрокинется.

Отсюда сразу следует вопрос, который долгое время, до второй половины XX века, не получал ответа: какой характеристикой случайных возмущающих воздействий нужно располагать для того, чтобы правильно вычислить средние квадраты решений линейных дифференциальных уравнений (в том числе и уравнения (3.2)), в правых частях которых стоят случайные функции, случайные процессы $\varphi(t)$? Такой характеристикой оказались спектральная плотность мощности (или коротко спектр) случайной функции $\varphi(t)$.

Имеет место следующее важное соотношение (его краткий вывод мы дадим позже): если переменные $x(t)$ и $\varphi(t)$ связаны дифференциальным уравнением

$$A(D)x = B(D)\varphi(t), \quad (3.7)$$

где $A(D)$ и $B(D)$ – полиномы от оператора дифференцирования произвольной степени, т.е.

$$A(D) = a_n D^n + a_{n-1} D^{n-1} + \dots + a_0, \quad (3.8)$$

$$B(D) = b_m D^m + b_{m-1} D^{m-1} + \dots + b_0, \quad (3.9)$$

то спектры переменных $x(t)$ и $\varphi(t)$ (спектры обозначают буквой S с соответствующими индексами, т.е. S_φ , S_x и т.п.) связаны между собой простым соотношением:

$$S_x = \left| \frac{B(j\omega)}{A(j\omega)} \right|^2 S_\varphi. \quad (3.10)$$

Если учесть, что средний квадрат переменной легко вычислить через ее спектр:

$$\langle x^2 \rangle = \int_0^\infty S_x(\omega) d\omega, \quad (3.11)$$

то получается простое правило вычисления средних квадратов решений дифференциальных уравнений (3.7): нужно в операторных полиномах (3.8) и (3.9) заменить оператор дифференцирования D на число $j\omega$, вычислить функцию

$$\frac{B(j\omega)}{A(j\omega)}, \quad (3.12)$$

квадрат ее модуля умножить на S_φ и вычислить интеграл (3.11). Таким образом, если известен спектр S_φ , то дальнейшие вычисления затруднений не представляют, а сам спектр S_φ является, как известно, косинус преобразованием Фурье от корреляционной функции $K_\varphi(\tau)$ процесса $\varphi(t)$:

$$S_\varphi(\omega) = \frac{2}{\pi} \int_0^\infty K_\varphi(\tau) \cos \omega \tau dt. \quad (3.13)$$

Таким образом, спектр S_φ является функцией переменной ω (частоты) имеющей размерность 1/время и вычисляется путем интегрального преобразования корреляционной функции по формуле (3.13).

В свою очередь корреляционная функция легко вычисляется на основе своего определения:

$$K_\varphi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \varphi(t) \varphi(t + \tau) dt. \quad (3.14)$$

Формула (3.14) показывает, что для вычисления $K_\varphi(\tau)$ достаточно перемножить значения процесса $\varphi(t)$, разделенные сдвигом во време-

ни τ и усреднить полученные произведения.

Формулы (3.10), (3.11), (3.13), (3.14) лежат в основе расчетов судов и судовых систем, для которых возмущающим воздействием являются случайные процессы, порожденные ветром и морским волнением, лежат в основе расчетов авиационных систем, испытывающих возмущающие воздействия случайного характера от турбулентности атмосферы, а также лежат в основе расчетов многих других систем – всех систем, где возмущающие воздействия случайны.

Реально, конечно, интеграл (3.14) вычисляется не на бесконечно большом интервале $-T \leq t \leq T$, где $T \rightarrow \infty$, а на конечном интервале, сами значения процесса $\varphi(t)$ измеряются с неизбежной погрешностью, и поэтому неизбежны малые отличия истинной корреляционной функции от той, которая вводится в интегральное преобразование (3.13).

Возникает законный вопрос: в какой мере малые погрешности корреляционной функции влияют на точность интегрального преобразования, на спектр S_φ ? Забегая вперед, сразу скажем: малые, и даже сколько угодно малые погрешности $K_\varphi(\tau)$ могут сильно влиять на $S_\varphi(\omega)$. Задача вычисления $S_\varphi(\omega)$ по заданному $K_\varphi(\tau)$ на основе формулы (3.13) в общем случае некорректна. Для успешного подхода к этой задаче необходим учет дополнительной информации о свойствах корреляционных функций и спектров, не полагаясь только на формулы (3.1), (3.13), (3.14), приводимые в справочниках. В последующих разделах мы проведем анализ свойств корреляционных функций и спектров, необходимых для успешного подхода к некорректной задаче определения спектра возмущающего воздействия на основе интегрального преобразования (3.13). Мы проанализируем также часто встречающиеся ошибки.

3.2. Свойства корреляционных функций

Отметим, прежде всего, что корреляционную функцию вычисляют для процессов, $\varphi(t)$, среднее значение которых, т.е. интеграл

$$\langle \varphi(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \varphi(t) dt, \quad (3.15)$$

равно нулю. Если $\langle \varphi \rangle \neq 0$, то корреляционную функцию вычисляют для отклонения процесса $\varphi(t)$ от его среднего значения, т.е. для процесса $\varphi_1(t) = \varphi(t) - \langle \varphi \rangle$.

Из формулы (3.14) непосредственно вытекают следующие свойства.

1. Величина $K_\varphi(t=0)$, т.е. значение корреляционной функции при $\tau=0$ равно среднему квадрату процесса $\varphi(t)$, т.е.

$$K_\varphi(0) = \langle \varphi^2 \rangle. \quad (3.16)$$

Поэтому вместо корреляционной функции (3.14) рассматривают часто нормированную корреляционную функцию k_φ (для отличия от обычной ее обозначают не прописной, а строчной буквой), равную частному от деления $K_\varphi(\tau)$ на $K_\varphi(0)$, т.е.

$$k_\varphi(\tau) = \frac{K_\varphi(\tau)}{K_\varphi(0)}. \quad (3.17)$$

Нормированная корреляционная функция при $\tau=0$ всегда равна единице. Обычная корреляционная функция равна нормированной, умноженной на средний квадрат, т.е. $K_\varphi(\tau) = \langle \varphi^2 \rangle k_\varphi(\tau)$.

2. Корреляционная функция является функцией четной, т.е.

$$K_\varphi(\tau) = K_\varphi(-\tau). \quad (3.18)$$

Действительно,

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \varphi(t) \varphi(t + \tau) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \varphi(t - \tau) \varphi(t) dt,$$

откуда и следует формула (3.18).

С учетом формулы (3.18) корреляционную функцию часто пишут лишь для положительных значений аргумента, для $\tau \geq 0$. Нужно всегда учитывать, что на самом деле она существует и для $\tau < 0$ и распространяется на значения $\tau < 0$ симметрично, согласно формуле (3.18).

3. Из очевидного неравенства

$$[\varphi(t) - \varphi(t + \tau)]^2 \geq 0$$

следует

$$\varphi^2(t) + \varphi^2(t + \tau) \geq 2\varphi(t)\varphi(t + \tau). \quad (3.19)$$

Беря среднее по времени от обеих частей неравенства (3.19), мы получим:

$$K_\varphi(0) \geq K_\varphi(\tau). \quad (3.20)$$

Таким образом, значения корреляционной функции для любого τ не могут быть больше, чем для $\tau=0$.

4. Четвертое важное свойство – это формула

$$K_\varphi(\tau) = -\frac{d^2}{d\tau^2} K_\varphi(\tau) \quad (3.21)$$

для корреляционной функции производной $\dot{\varphi}(t)$ процесса $\varphi(t)$.

Ее получают из основной формулы (3.14) дифференцированием под знаком интеграла с последующим интегрированием по частям. Корреляционная функция производной $\dot{\varphi}$ равна второй производной (с обратным знаком) обычной корреляционной функции $K_{\varphi}(\tau)$ самого процесса $\varphi(t)$.

Примеры.

а) Если $\varphi(t) = A \sin(\beta t + \psi_0)$, то заменяя в формуле

$$K_{\varphi}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} A^2 \sin(\beta t + \psi_0) \sin(\beta t + \beta \tau + \psi_0) dt$$

произведение синусов на косинусы полуразности и полусуммы, согласно известным формулам тригонометрии, и учитывая, что интеграл от полусуммы косинусов при $\tau \rightarrow \infty$ обратится в нуль, получаем:

$$K_{\varphi}(\tau) = \frac{A^2}{2} \cos \beta \tau. \quad (3.22)$$

Мы убеждаемся, что корреляционная функция гармонического колебания $A \sin(\beta t + \psi_0)$ зависит от его амплитуды и частоты и не зависит от фазы.

б) Если $\varphi(t)$ является периодическим процессом, разложимым в ряд Фурье:

$$\varphi(t) = \sum_{n=1}^{\infty} A_n \sin(n\beta t + \psi_n), \quad (3.23)$$

то, учитывая, что интегралы от произведений гармоник с различными номерами обращаются в нуль, и, используя те же тригонометрические преобразования, что и при выводе формулы (3.22), получим

$$K_{\varphi}(\tau) = \frac{1}{2} \sum_{n=1}^{\infty} A_n^2 \cos n\beta \tau. \quad (3.24)$$

Формула (3.24) показывает, что корреляционная функция не зависит от сдвига по фазе отдельных гармоник периодического процесса относительно друг друга. Поэтому различные процессы, имеющие одинаковые амплитуды, но разные фазы отдельных гармоник будут иметь одну и ту же корреляционную функцию. Это означает, что корреляционная функция не является однозначной характеристикой процесса $\varphi(t)$. Различные процессы могут иметь одинаковые корреляционные функции.

в) Рассмотрим теперь случайный процесс, когда $\varphi(t)$ случайно и с равной вероятностью принимает значения $+a$ и $-a$ (рис. 3.1).

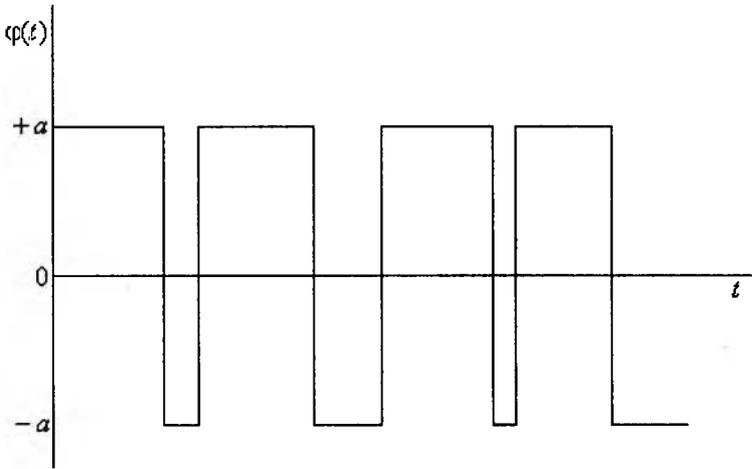


Рис. 3.1

Такой вид имеет напряжение на выходе телеграфного аппарата, поэтому данный процесс называют «телеграфным сигналом». Среднее значение «телеграфного сигнала» равно нулю. Найдем теперь корреляционную функцию. Вероятность перемены знака подчинена закону Пуассона с параметром μ , т.е. вероятность того, что за время τ не произойдет ни одной перемены знака, равна $P_0 = e^{-\mu\tau}$, а вероятность того, что за это же время произойдет ровно n перемен знака, равна

$$P_n = \frac{(\mu\tau)^n}{n!} e^{-\mu|\tau|}.$$

Произведение $\varphi(t)\varphi(t + \tau)$, входящее в корреляционную функцию, будет равно либо a^2 , либо $-a^2$, в зависимости от того, четное или нечетное число раз переменит $\varphi(t)$ свой знак за время τ . Но это означает, что

$$K_\varphi(\tau) = a^2(P_{\text{ч}} - P_{\text{н}}),$$

где $P_{\text{ч}}$ – вероятность четного числа перемен знака за время τ , а $P_{\text{н}}$ – вероятность нечетного числа перемен.

Поскольку

$$P_{\text{ч}} = \left[\sum_{k=0}^{\infty} \frac{(\mu\tau)^{2k}}{(2k)!} \right] e^{-\mu|\tau|}, \quad P_{\text{н}} = \left[\sum_{k=0}^{\infty} \frac{(\mu\tau)^{2k+1}}{(2k+1)!} \right] e^{-\mu|\tau|},$$

то, просуммировав ряды, получаем:

$$K_{\varphi}(t) = a^2 e^{-2\mu|t|}, \quad (3.25)$$

т.е. корреляционная функция затухает экспоненциально, стремясь к нулю с ростом τ . Чем больше μ , т.е. чем чаще $\varphi(t)$ меняет знак, тем быстрее затухает корреляционная функция.

г) Рассмотрим теперь кусочно-постоянный случайный процесс, который переходит от одного значения φ_i к другому в случайные моменты времени, подчиненные закону Пуассона с параметром μ , т.е. вероятность того, что $\varphi(t)$ останется постоянным на интервале времени τ , равна $e^{-\mu\tau}$. График такого процесса показан на рис. 3.2.

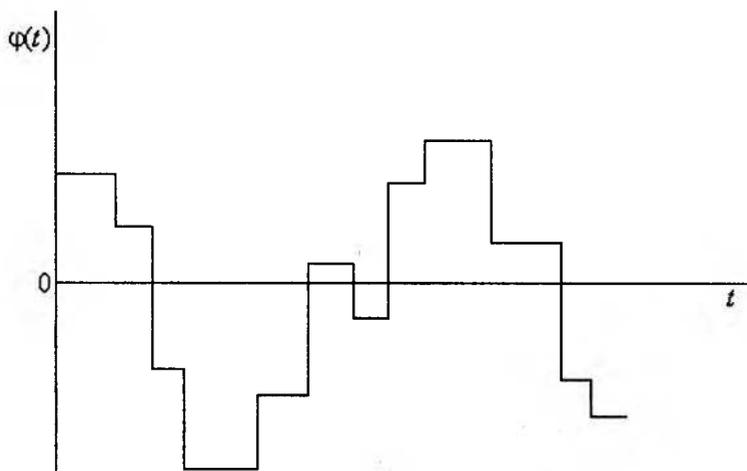


Рис. 3.2

Произведение $\varphi(t)\varphi(t + \tau)$ будет различным в зависимости от того, происходит или нет за время τ переход процесса $\varphi(t)$ от одного значения φ_k к другому значению φ_i , или не происходит.

В первом случае

$$\varphi(t)\varphi(t + \tau) = \varphi_k\varphi_k = \varphi_k^2,$$

во втором случае

$$\varphi(t)\varphi(t + \tau) = \varphi_k\varphi_i.$$

Вероятность первого случая равна $e^{-\mu\tau}$, вероятность второго $1 - e^{-\mu\tau}$. Следовательно,

$$K_{\varphi}(\tau) = \langle \varphi_k \rangle e^{-\mu\tau} + \langle \varphi_k\varphi_i \rangle (1 - e^{-\mu|\tau|}).$$

Но если φ_k и φ_i независимы и подчинены одному и тому же закону распределения, то $\langle \varphi_k \varphi_i \rangle = 0$ и окончательно имеем:

$$K_\varphi(\tau) = \langle \varphi^2 \rangle e^{-\mu|\tau|}. \quad (3.26)$$

Снова убеждаемся, что не очень похожие друг на друга случайные процессы имеют схожие корреляционные функции. Это еще раз подтверждает, что различные процессы могут иметь схожие и даже тождественные корреляционные функции. Процессы, имеющие одну и ту же корреляционную функцию, например, функцию (3.26), называют «реализациями», точнее – реализациями процесса с данной корреляционной функцией.

3.3. Свойства спектров

Спектральная плотность мощности процесса $\varphi(t)$ (называемая коротко спектром) является известным из курса математического анализа интегральным преобразованием Фурье от его корреляционной функции. В анализе интегральным преобразованием Фурье называют преобразование:

$$\int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} f(t) (\cos \omega t - j \sin \omega t) dt = F(j\omega), \quad (3.27)$$

которое переводит функцию времени $f(t)$ в функцию $F(j\omega)$ от частоты ω , в общем случае – комплексную. Поскольку корреляционная функция является функцией четной, то нас будет интересовать преобразование Фурье для четных функций, для которых

$$\int_{-\infty}^{+\infty} f(t) \sin \omega t dt = 0$$

и поэтому общее преобразование Фурье (3.27) переходит в косинусное преобразование:

$$\int_{-\infty}^{+\infty} f(t) \cos \omega t dt = F_c(\omega), \quad (3.28)$$

которое переводит четную функцию времени $f(t)$ в вещественную (и тоже четную) функцию частоты $F_c(\omega)$. Как известно из курса анализа, обратное преобразование Фурье:

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} F_c(\omega) \cos \omega t d\omega = f(t) \quad (3.29)$$

восстанавливает исходную функцию времени $f(t)$. Учитывая четность

функций $f(t)$ и $F_c(\omega)$, интегрирование в пределах от $-\infty$ до $+\infty$ можно заменить интегрированием на меньшем промежутке, от 0 до $+\infty$, и тогда получается известная из математического анализа следующая пара формул для прямого и обратного косинус преобразования Фурье:

$$F_c(\omega) = \int_0^{+\infty} f(t) \cos \omega t dt, \quad (3.30)$$

$$f(t) = \frac{2}{\pi} \int_0^{+\infty} F_c(\omega) \cos \omega t d\omega. \quad (3.31)$$

Обратное косинус преобразование (3.31) отличается от прямого косинус преобразования (3.3) множителем $\frac{2}{\pi}$. Однако можно условиться писать этот множитель перед прямым косинус преобразованием и тогда этого множителя не будет в обратном преобразовании. Именно в этой форме принято записывать используемые при расчете судов и судовых систем косинус преобразования для корреляционной функции. Имеют место следующие важные формулы:

$$S_\varphi(\omega) = \frac{2}{\pi} \int_0^{+\infty} K_\varphi(\tau) \cos \omega t d\tau, \quad (3.32)$$

$$K_\varphi(\tau) = \int_0^{+\infty} S_\varphi(\omega) \cos \omega t d\omega. \quad (3.33)$$

Функцию $S_\varphi(\omega)$, определяемую равенством (3.32), называют «спектральной плотностью мощности» или коротко «спектром» процесса $\varphi(t)$.

Пример: для корреляционной функции $K_\varphi = e^{-\alpha\tau}$, вычисляя интеграл

$$\frac{2}{\pi} \int_0^{\infty} e^{-\alpha\tau} \cos \omega\tau = \frac{2}{\pi} \frac{\alpha}{\alpha^2 + \omega^2}, \quad (3.34)$$

убеждаемся, что в этом случае спектр $S_\varphi(\omega) = \frac{2}{\pi} \frac{\alpha}{\alpha^2 + \omega^2}$.

Вычислив интеграл

$$\int_0^{\infty} \frac{2}{\pi} \frac{\alpha}{\alpha^2 + \omega^2} \cos \omega t d\omega = e^{-\alpha t}, \quad (3.35)$$

убеждаемся, что обратное преобразование (3.33) действительно восстановило исходную корреляционную функцию.

Представляет интерес вычисление спектра корреляционной функции $K_\varphi = e^{-\alpha\tau}$ в предельном случае при $\alpha \rightarrow 0$, когда $K_\varphi(\tau) \rightarrow 1$, а сам процесс $\varphi(t)$ тоже превращается в постоянную величину, $\varphi(t) = 1$.

Предел спектра (3.34) при $\alpha \rightarrow 0$ превращается в очень своеобразную функцию, равную нулю для всех $\omega \neq 0$, а при $\omega = 0$ эта функция бесконечно велика.

В то же время, поскольку формула (3.35) справедлива для всех α , в том числе и при $\alpha \rightarrow 0$, то интеграл от этой функции, имеющей одно единственное значение, неравное нулю, равен единице.

Такую своеобразную функцию называют δ -функцией Дирака и обозначают через $\delta(\omega)$. Таким образом, спектр постоянной величины $\varphi(t) = A$ является дельта функцией:

$$S_{\varphi=A} = A^2\delta(\omega).$$

Дельта-функции, введенные впервые английским физиком Дираком в 1928 году, широко используются в приложениях. Их можно рассматривать как пределы многих обычных функций. Так, функция $\varphi(t)$, равная нулю для $-\infty < t < 0$ и равная $\varphi(t) = \alpha e^{-\alpha t}$ для $0 \leq t < \infty$ в пределе, при $\alpha \rightarrow 0$, переходит в дельта функцию: $\lim_{\alpha \rightarrow 0} \alpha e^{-\alpha t} = \delta(t)$.

Рассмотрим основные свойства функции $S_\varphi(\omega)$, определяемой формулами (3.32) и (3.33).

1. Формула (3.33) справедлива для всех τ , в том числе и для $\tau = 0$. Подставляя $\tau = 0$ в (3.33), получаем:

$$K_\varphi(0) = \int_0^\infty S_\varphi(\omega) d\omega. \quad (3.36)$$

Отсюда, с учетом формулы (3.16), имеем:

$$\langle \varphi^2 \rangle = \int_0^\infty S_\varphi(\omega) d\omega. \quad (3.37)$$

Таким образом, интеграл от функции $S_\varphi(\omega)$ равен среднему квадрату процесса $\varphi(t)$. А учитывая, что средний квадрат называют часто «мощностью» процесса $\varphi(t)$, становится понятным полное название функции $S_\varphi(\omega)$ – «спектральная плотность мощности» процесса $\varphi(t)$. Она отражает распределение мощности по частотам. Вместо полного названия функцию $S_\varphi(\omega)$ называют обычно просто «спектр».

2. Из основной формулы (3.32) сразу следует, что если $K_\varphi(\tau) = K_1(\tau) + K_2(\tau)$, то $S_\varphi(\omega) = S_1(\omega) + S_2(\omega)$, т.е. спектр суммы двух корреляционных функций равен сумме спектров слагаемых.

3. Из той же формулы следует, что преобразование однозначно. Каждой корреляционной функции соответствует свой спектр. Зная спектр, мы можем вычислить корреляционную функцию, и наоборот.

4. Как известно из математического анализа, для общего преобразования Фурье (3.27) справедлива формула:

$$\int_{-\infty}^{+\infty} \frac{df}{dt} e^{-j\omega t} dt = j\omega F(j\omega), \quad (3.38)$$

т.е. преобразование Фурье производной $\frac{df}{dt}$ равно преобразованию Фурье функции $f(t)$, умноженному на число $j\omega$. С учетом формулы (3.22) для производной корреляционной функции $\dot{\varphi}(t)$ процесса $\varphi(t)$, получаем:

$$S_{\dot{\varphi}}(\omega) = |j\omega|^2 S_\varphi(\omega) = \omega^2 S_\varphi(\omega). \quad (3.39)$$

Таким образом, спектр производной $\dot{\varphi}(t)$ равен спектру процесса $\varphi(t)$, умноженному на квадрат модуля числа $j\omega$, или, что то же самое – на число ω^2 .

Пользуясь формулой (3.39), нетрудно найти связь между спектром $S_x(\omega)$ решения $x(t)$ дифференциального уравнения (3.7) и спектром $S_\varphi(\omega)$ его правой части. Эта связь выражается уже приводимой нами ранее формулой (3.10). Теперь мы привели ее вывод.

С учетом формулы (3.37) нетрудно найти и средний квадрат решения $x(t)$ дифференциального уравнения (3.7):

$$\langle x^2 \rangle = \int_0^\infty \frac{|B(j\omega)|^2}{|A(j\omega)|^2} S_\varphi(\omega) d\omega. \quad (3.40)$$

Формула (3.40) открывает простой путь вычисления средних квадратов для всех тех многочисленных технических систем, возмущающими воздействиями для которых являются случайные, не полностью предсказуемые процессы. Для этого достаточно, используя измерения текущих значений возмущающих воздействий $\varphi(t)$ и опираясь на формулу (3.14), вычислить корреляционную функцию $K_\varphi(\tau)$, преобразовать ее в спектр $S_\varphi(\omega)$ на основе формулы (3.32), а потом для любого дифференциального уравнения (3.7) достаточно заменить в полино-

маж $A(D)$ и $B(D)$ оператор дифференцирования $D = \frac{d}{dt}$ на число $j\omega$, найти квадраты модулей $|B(j\omega)|^2$, $|A(j\omega)|^2$ и вычислить интеграл (3.40).

Спектры возмущающих воздействий и формула (3.40) лежат сейчас в основе всех расчетов технических объектов морского флота (поскольку возмущающее воздействие на суда и судовые системы от ветра и морского волнения являются случайными, не полностью предсказуемыми процессами), лежат в основе расчетов авиационных систем (поскольку возмущающие воздействия от турбулентности атмосферы тоже случайны) и многих других систем и технических устройств.

Однако многие важные аспекты интегральных преобразований освещены недостаточно. В следующем разделе мы рассмотрим более подробно важный вопрос о корректности преобразований.

3.4. Корректность интегральных преобразований

При вычислении корреляционной функции процесса $\varphi(t)$ на основе формулы (3.14) мы должны учитывать, что неизбежные малые погрешности в измерении значений $\varphi(t)$ и $\varphi(t + \tau)$, а также замена теоретического бесконечного интервала интегрирования в формуле (3.14) на реальный конечный интервал приводят к неизбежным отклонениям вычислений корреляционной функции от истинной. Возникает законный вопрос, в какой мере эти отклонения скажутся на конечном результате, на вычислении среднего квадрата решения $x(t)$ дифференциального уравнения

$$A(D)x = B(D)\varphi. \quad (3.41)$$

Является ли корректной задача вычисления среднего квадрата решения через интегральное косинус преобразование Фурье процесса $\varphi(t)$?

В качестве примера рассмотрим простейшее уравнение

$$x = D\varphi \quad (3.42)$$

для случая $K_\varphi(\tau) = e^{-\alpha|\tau|}$. Поскольку в этом случае

$$S_\varphi(\omega) = \frac{2}{\pi} \int_0^\infty e^{-\alpha\tau} \cos \omega\tau d\tau = \frac{2}{\pi} \frac{\alpha}{\alpha^2 + \omega^2}, \quad (3.43)$$

то спектр $x(t)$ равен

$$S_x(\omega) = \frac{2}{\pi} \frac{\alpha\omega^2}{\alpha^2 + \omega^2}. \quad (3.44)$$

Вычисляя средний квадрат $x(t)$ по формуле (3.37), убеждаемся, что в нашем случае средний квадрат бесконечно велик, поскольку несобственный интеграл

$$\langle x^2 \rangle = \int_0^{\infty} \frac{2}{\pi} \frac{\alpha \omega^2}{\alpha^2 + \omega^2} d\omega$$

расходится.

Теперь рассмотрим ту же задачу вычисления среднего квадрата $x(t)$ для процесса $\varphi(t)$ с корреляционной функцией

$$K_{\varphi}(\tau) = e^{-\alpha|\tau|} + \alpha\tau e^{-\beta|\tau|} \quad (3.45)$$

Вычислив интеграл

$$S_{\varphi}(\omega) = \frac{2}{\pi} \int_0^{\infty} (e^{-\alpha\tau} + \alpha\tau e^{-\beta\tau}) \cos \omega\tau d\tau = \frac{2\alpha}{\pi} \left[\frac{\beta^4 + \alpha^2\beta^2 + (3\beta^2 - \alpha^2)\omega^2}{(\alpha^2 + \omega^2)(\beta^2 + \omega^2)^2} \right], \quad (3.46)$$

мы можем теперь вычислить и средний квадрат $x(t)$ по формуле:

$$\langle x^2 \rangle = \int_0^{\infty} \frac{2\alpha\omega^2}{\pi} \left[\frac{\beta^4 + \alpha^2\beta^2 + (3\beta^2 - \alpha^2)\omega^2}{(\alpha^2 + \omega^2)(\beta^2 + \omega^2)^2} \right] d\omega. \quad (3.47)$$

Сразу замечаем, что интеграл (3.47) конечен для любых β , поскольку для больших ω он убывает, как $\frac{1}{\omega^2}$, а такие интегралы сходятся.

В то же время при малых значениях дроби $\frac{\alpha}{\beta}$, (т.е. при больших β) корреляционная функция (3.45) будет сколь угодно мало отличаться от $K_{\varphi}(\tau) = e^{-\alpha|\tau|}$.

Действительно, максимум разности этих корреляционных функций:

$$(e^{-\alpha|\tau|} + \alpha\tau e^{-\beta|\tau|}) - e^{-\alpha\tau} = \alpha\tau e^{-\beta\tau} \quad (3.48)$$

будет достигаться при $\tau = \frac{1}{\beta}$, и равен этот максимум

$$\Delta_{\max} = \frac{1}{e} \frac{\alpha}{\beta} = 0.368 \frac{\alpha}{\beta}. \quad (3.49)$$

При $\frac{\alpha}{\beta} \rightarrow 0$ различие между корреляционными функциями

$K_\varphi(\tau) = e^{-\alpha|\tau|}$ и (3.45) может быть сколь угодно малым (за счет большого β), в то время как различие между средними квадратами решения $x(t)$ дифференциального уравнения (3.42) принципиально: при $K_\varphi(\tau) = e^{-\alpha|\tau|}$ средний квадрат бесконечен, а при $K_\varphi(\tau) = e^{-\alpha|\tau|} + \alpha\tau e^{-\beta|\tau|}$ он конечен для любых α и β . (Заметим, что разность корреляционных функций, равная $\alpha\tau e^{-\beta\tau}$, при $\frac{\alpha}{\beta} \rightarrow 0$ будет малой и в среднеквадратичной метрике, поскольку интеграл

$$\int_0^{\infty} (\alpha\tau e^{-\beta\tau})^2 d\tau \quad (3.50)$$

будет сколь угодно малым при $\frac{\alpha}{\beta} \rightarrow 0$).

Таким образом, мы приходим к важному выводу: задача вычисления средних квадратов решений дифференциального уравнения через интегральное косинус преобразование Фурье может быть некорректной. Неизбежные малые погрешности в вычислении корреляционной функции (и даже сколь угодно малые погрешности) могут привести к коренным и принципиальным различиям в средних квадратах.

На этом примере удобно разъяснить одну из методик подхода к некорректным задачам. Увеличение точности вычисления корреляционной функции ничего не дает: коренные расхождения в величине среднего квадрата сохраняются при сколь угодно малых неточностях. Однако можно использовать дополнительную информацию о реальных процессах, спектры которых мы вычисляем, и эта информация позволит нам правильно подойти к некорректной задаче, позволит правильно выбрать аналитическую аппроксимацию корреляционной функции по экспериментальным данным, содержащим неизбежные погрешности измерения.

Реальные процессы имеют конечную скорость изменения, и поэтому средний квадрат скорости, средний квадрат производной $\dot{\varphi}(t)$ реального процесса $\varphi(t)$, т.е. интеграл

$$\langle \dot{\varphi} \rangle = \int_0^{\infty} S_{\dot{\varphi}}(\omega) d\omega = \int_0^{\infty} \omega^2 S_{\varphi} d\omega \quad (3.51)$$

должен быть конечным. Однако мы уже установили, что этот интеграл равен значению второй производной от корреляционной функции в точке $\tau = 0$, т.е.

$$\int_0^{\infty} \omega^2 S_{\varphi}(\omega) d\omega = -\frac{d^2}{d\tau^2} K_{\varphi}(0). \quad (3.52)$$

Но вторая производная при $\tau=0$ будет конечна только тогда, когда первая производная при $\tau=0$ непрерывна. Обращаясь к очень широко используемой корреляционной функции

$$K_{\varphi}(\tau) = e^{-\alpha|\tau|}, \quad (3.53)$$

сразу видим, что слева от точки $\tau=0$ производная $\frac{dK_{\varphi}}{d\tau}$ положительна и равна $\alpha e^{\alpha\tau}$ и при $\tau \rightarrow 0$ стремится к значению $+\alpha$, а справа от точки $\tau=0$ производная $\frac{dK_{\varphi}}{d\tau} = -\alpha e^{-\alpha\tau}$ отрицательна и при $\tau \rightarrow 0$ стремится к значению $-\alpha$. Непрерывности значения первой производной при $\tau=0$ нет. Отсюда следует, что при $\tau=0$ вторая производная функции (3.53) бесконечно велика, и поэтому не удивительно, что средний квадрат производной $\dot{\varphi}(t)$ процесса $\varphi(t)$, имеющего корреляционную функцию (3.53), оказывается бесконечно большим.

Однако, поскольку конечность или бесконечность среднего квадрата производной $\dot{\varphi}(t)$ зависит только от поведения второй производной корреляционной функции в единственной точке $\tau=0$, то восстановить конечность среднего квадрата можно сколь угодно малой добавкой к функции (3.53), достаточно лишь обеспечить, чтобы производная $\frac{dK_{\varphi}}{d\tau}$ справа и слева от точки $\tau=0$ были равны друг другу.

Мы уже убедились, что корреляционная функция (3.45), которая при малых значениях отношения $\frac{\alpha}{\beta}$ сколь угодно мало отличается от корреляционной функции (3.53) соответствует процессу $\varphi(t)$ с конечным средним квадратом скорости $\dot{\varphi}(t)$, поскольку производная функции (3.45):

$$\frac{d}{d\tau} (e^{-\alpha|\tau|} + \alpha\tau e^{-\beta|\tau|}) = -\alpha\beta|\tau| e^{-\beta|\tau|} \quad (3.54)$$

стремится при $|\tau| \rightarrow 0$ к одному и тому же не нулевому значению и слева и справа от точки $\tau=0$.

Таким образом, аналитическая аппроксимация корреляционной функции в форме (3.45) правильно отражает физический смысл реального процесса, скорость изменения которого не может быть бесконечной. Эта дополнительная информация (о конечности скорости) позво-

ляет правильно выбрать аналитическую аппроксимацию корреляционной функции в некорректной задаче ее построения по измерениям процесса $\varphi(t)$ с неизбежными малыми погрешностями измерения.

Что касается корреляционной функции (3.53), то она не может правильно отразить такую важную характеристику почти любого реального процесса как конечность производной, конечность скорости изменения процесса. (Напомним, что корреляционную функцию типа (3.53) имели процессы, показанные на рисунках 3.1 и 3.2, т.е. процессы, имеющие «скачки», изображенные на графиках вертикальными отрезками; в точках «скачка» производная $\varphi(t)$ бесконечна. В реальных процессах производная может быть большой, даже очень большой, но не бесконечной. Поэтому процессы, показанные на рисунках 3.1 и 3.2 и другие процессы, имеющие корреляционную функцию (3.53) являются лишь математической идеализацией.)

Однако, несмотря на это, корреляционная функция (3.53) очень широко и успешно используется на практике. О причинах этого расскажем в следующем разделе.

3.5. Задачи, мало чувствительные к погрешностям спектров

Большинство практических задач на вычисление средних квадратов тех или иных характеристик динамических систем, в которых возмущающими воздействиями являются случайные процессы, мало чувствительны к тем погрешностям спектров возмущающих воздействий, которые лежат в области высоких частот, больших значений переменной ω в спектрах $S_{\varphi}(\omega)$.

Мы уже упоминали, что если переменные $x(t)$ и $\varphi(t)$ связаны между собой дифференциальным уравнением (3.7), в котором полиномы от оператора дифференцирования $D = \frac{d}{dt}$ имеют вид (3.8) и (3.9), то спектры S_x и S_{φ} связаны между собой формулой (3.10), в которой величину $\frac{B(j\omega)}{A(j\omega)}$ называют «частотной характеристикой» динамической системы, или, что, то же самое – частотной характеристикой математической модели (3.7). Поскольку для подавляющего большинства систем степень полинома $A(D)$ больше степени полинома $B(D)$, то квадрат модуля частотной характеристики, т.е.

$$\left| \frac{B(j\omega)}{A(j\omega)} \right|^2 \quad (3.55)$$

для больших ω делается очень малой величиной и поэтому даже большое расхождение между спектрами S_1 и S_2 процессов $\varphi_1(t)$ и $\varphi_2(t)$ в области больших частот может привести к совсем малым расхождениям в средних квадратах $\langle x_1^2 \rangle$ и $\langle x_2^2 \rangle$ решений $x_1(t)$ и $x_2(t)$.

Мы уже пользовались этим обстоятельством в разделе 1.4 главы 1, когда для обеспечения корректности задачи оптимизации управления, для обеспечения сохранения устойчивости замкнутой системы при неизбежных малых отклонениях реальных параметров объекта управления от расчетных значений мы проводили замену спектра возмущающего воздействия

$$S_1(\omega) = \frac{2}{\pi} \frac{1}{1 + \omega^2} \quad (3.56)$$

(в наибольшей мере отвечающего экспериментальным данным) на спектр

$$S_2(\omega) = \frac{2}{\pi} \frac{1 + 0.01\omega^2}{1 + \omega^2}, \quad (3.57)$$

при котором выполняется «критерий Ю. Петрова», который является критерием корректности рассматриваемой задачи оптимизации. Несмотря на то, что при больших ω различие между спектрами (3.56) и (3.57) велико (так при $\omega = 10$ будет $S_2(10) = 2S_1(10)$, а при $\omega = 20$ будет даже $S_2(20) = 5S_1(20)$), различие между средними квадратами управляемого процесса составило всего 0,88% (глава первая, формула (40)).

Именно поэтому вместо корреляционной функции (3.45), которой соответствует спектр (3.46), гораздо чаще используют функцию (3.53), которой соответствует спектр (3.44). Различие между спектрами (3.44) и (3.46) велико только для больших частот, а спектр (3.44) много проще.

Более того, спектр (3.44) часто заменяют на еще более простой: $S_\varphi(\omega) = \text{const}$, т.е. заменяют на спектр, постоянный для всех частот. Процесс, имеющий такой спектр, называют «белым шумом». «Белый шум», безусловно, является чисто математической идеализацией, никакой реальный процесс не может быть «белым шумом» уже потому, что средний квадрат «белого шума» бесконечен. Тем не менее, замена реального спектра на простую постоянную величину часто используется. Причину допустимости такой замены удобно пояснить на примере. В качестве примера рассмотрим движение судна типа «Казбек» (водо-

измещение 16 тысяч тонн, скорость – 14 узлов) под действием руля и возмущающих воздействий от ветра и морского волнения. Математической моделью движения этого судна является уравнение:

$$(690D^2 + 60.8D + 2.5)\theta = \varphi(t), \quad (3.58)$$

где θ – угол отклонения судна от заданного курса, а $\varphi(t)$ – момент возмущающих сил ветра и морского волнения. Спектр возмущающих сил, как известно, хорошо аппроксимируется известной формулой Рахманина–Фирсова:

$$S_{\varphi}(\omega) = \langle \varphi^2 \rangle \frac{4\alpha}{\pi} \frac{\alpha^2 + \beta^2}{(\alpha^2 + \beta^2 + \omega^2)^2 - 4\beta^2\omega^2}, \quad (3.59)$$

где параметры α и β зависят от интенсивности волнения. Для волнения средней интенсивности часто встречающимися значениями являются $\beta = \frac{1}{\text{сек}}$, $\alpha = 0.21\beta$.

Спектр (3.59) для $\beta = \frac{1}{\text{сек}}$, $\alpha = 0.21\beta$ показан на рис.3.3 непрерывной линией и там же пунктирной линией показан квадрат модуля частотной характеристики судна «Казбек», равный

$$\frac{1}{|690(j\omega)^2 + 60.8j\omega + 2.5|^2}. \quad (3.60)$$

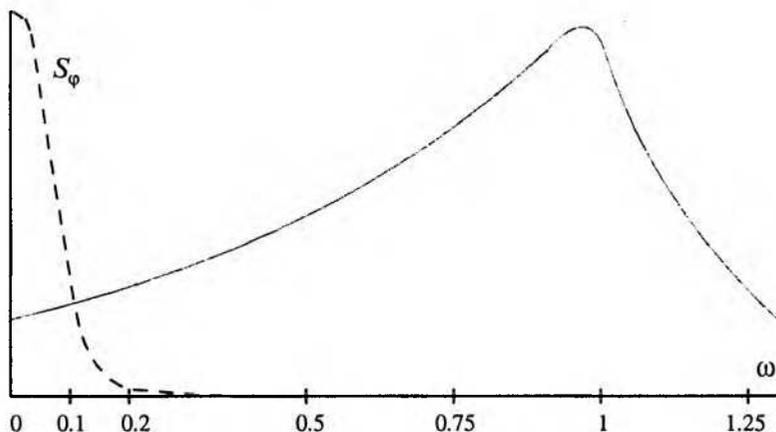


Рис. 3.3

Мы убеждаемся, что квадрат модуля частотной характеристики судна «Казбек» сколько-нибудь существенно отличается от нуля лишь

в полосе частот от $\omega = 0$ до $\omega = 0.15$, а в этой полосе спектр (3.59) отличается от $S_\varphi(\omega) = \text{const}$ не более чем на 4%. Поэтому различие в среднеквадратичном значении угла отклонения «Казбека» от заданного курса при замене спектра (3.59) на более простой спектр $S_\varphi(\omega) = \text{const}$ не превысит 1–1,5% (а для судов, чье водоизмещение больше «Казбека», т.е. больше чем 16 тысяч тонн, различие будет еще меньшим). Поэтому при расчете закона управления рулевой установкой, при расчетах отклонения судна от курса и других подобных расчетах заменяют реальное возмущающее воздействие на «белый шум». Такая замена упрощает и вычисления и реализацию рулевой установки, и в то же время дает вполне достаточную точность. Подробные расчеты оптимальных систем управления рулем при различных аналитических аппроксимациях спектров возмущающих воздействий приведены в монографии [12] на стр. 132–147. Эти расчеты еще раз подтвердили правомерность замены спектра (3.59) на $S_\varphi(\omega) = \text{const}$. Это связано с тем, что такая динамическая система, как судно, управляемое рулевой установкой, очень мало чувствительно к изменениям спектра за пределами полосы частот, существенных для судна, а эта полоса очень узка, не выходит обычно за пределы полосы $0 \leq \omega \leq 0.1$ или $0 \leq \omega \leq 0.15$.

При расчетах качки судов все уже обстоит по-другому. Для бортовой качки частота свободных колебаний судна и частота максимума спектра возмущающих сил соизмеримы. Поэтому для расчетов качки наиболее существенными оказываются участки спектра, лежащие вблизи максимума, а участки вблизи $\omega = 0$ несущественны. Поэтому при расчетах качки с равным успехом используются как спектр (3.59), которому соответствует корреляционная функция

$$K_\varphi(\tau) = \langle \varphi^2 \rangle e^{-\alpha|\tau|} \left(\cos \beta\tau + \frac{\alpha}{\beta} \sin \beta|\tau| \right), \quad (3.61)$$

так и модификация спектра (3.59):

$$S_\varphi(\omega) = \langle \varphi^2 \rangle \frac{2\alpha}{\pi} \frac{\alpha^2 + \beta^2 + \omega^2}{(\alpha^2 + \beta^2 + \omega^2)^2 - 4\beta^2\omega^2}, \quad (3.62)$$

которой соответствует корреляционная функция

$$K_\varphi(\tau) = \langle \varphi^2 \rangle e^{-\alpha|\tau|} \cos \beta\tau. \quad (3.63)$$

Легко заметить, что спектр (3.59) и корреляционная функция (3.61) соответствуют случайному процессу $\varphi(t)$ с конечным средним квадратом производной $\dot{\varphi}(t)$, а спектр (3.62) и корреляционная функция (3.63) соответствуют более идеализированному процессу $\varphi(t)$ – такому процессу, для которого средний квадрат производной $\dot{\varphi}(t)$ не имеет

конечного значения, поскольку интеграл

$$\langle \dot{\varphi} \rangle = \int_0^{\infty} \omega^2 S_{\varphi}(\omega) d\omega$$

для спектра (3.62) расходится (для спектра (3.59) этот интеграл конечен, поскольку при $\omega \rightarrow \infty$ спектр (3.59) убывает быстрее).

И спектр (3.59), и спектр (3.62) называются спектрами Рахманина–Фирсова и с равным успехом используются при расчетах качки несмотря на то, что при частотах, близких к $\omega = 0$, ординаты спектра (3.59) при равных $\langle \varphi^2 \rangle$ в два раза больше, чем у спектра (3.62). Но при расчетах качки малые частоты несут существенны и оба спектра применимы в равной мере.

Все изменилось, когда для буровых и разведывательных работ потребовались суда, которые несмотря на возмущающее воздействие от ветра и морского волнения точно удерживаются (с очень малыми отклонениями) в заданной точке моря с помощью систем динамической стабилизации. Расчет подобных систем показал, что для них существенны и малые, и большие частоты. Быстро выяснилось, что используемые до 80-х годов двадцатого века спектры Рахманина–Фирсова (3.59) и (3.62) оба давали неверные соотношения между максимумами спектра и его значениями при $\omega = 0$, а это приводило к неоправданному завышению требуемой мощности систем динамической стабилизации. Еще более худший результат получался при использовании широко известных и приводимых в большинстве справочников спектров Дербишайра, Неймана, Бретшнайдера, Ролла и Фишера. Эти спектры объединяет общая формула (формула Барлинга):

$$S_{\varphi}(\omega) = A\omega^{-k} e^{-B\omega^{-n}}, \quad (3.64)$$

где коэффициенты A, B, k, n каждый из авторов перечисленных спектров выбирает с небольшими отклонением один от другого. Все спектры, основанные на формуле (3.64), объединяет общая черта: при $\omega \rightarrow 0$ ординаты спектра также стремятся к нулю, и поэтому $S_{\varphi}(\omega = 0) = 0$. При расчетах качки судов эти спектры применялись успешно, но когда их стали применять к расчетам систем динамической стабилизации, где малые частоты важны, сразу проявилось резкое различие между рассчитанным поведением системы и реальным.

Причина заключается в том, что неизбежные погрешности в измерении значения процесса $\varphi(t)$, а также погрешности в их обработке, вычислении корреляционной функции и спектра $S_{\varphi}(\omega)$ особенно сильно сказывается для малых значений переменной ω , близких к

$\omega = 0$. Задача вычисления величины $S_\varphi(\omega)$ для значения $\omega = 0$ и близких к нему оказывается задачей некорректной (или близкой к некорректной), поэтому неудивительно, что в разных экспериментах, проводимых разными исследователями, получались разные значения для $S_\varphi(0)$ и поэтому было предложено применить типичный для некорректных задач подход, а именно – использовать дополнительную информацию о спектре $S_\varphi(\omega)$ сверх той, которую дают измерения, и, прежде всего, использовать информацию, основанную на физических соображениях.

Рассуждали тогда следующим образом: для систем динамической стабилизации судов возмущающим воздействием является угол волнового склона $\varphi(t)$, а реальное нерегулярное морское волнение – это сложная комбинация разных волн, самых различных длин и частот. Считали, что значениям $S_\varphi(\omega)$ для малых ω , малых частот, соответствуют очень длинные «элементарные волны», но такие волны не могут получить энергию от ветра, поэтому они должны быстро затухнуть и для малых ω должно быть $S_\varphi(\omega) = 0$. Даже если измерения приводили к значениям $S_\varphi(0) \neq 0$, их считали происходящими от погрешностей измерения и заменяли на $S_\varphi(0) = 0$. К сожалению, все эти рассуждения изначально были неверны. Дело в том, что ординаты спектра $S_\varphi(\omega)$, соответствующие малым частотам, совсем не обязательно должны отражать реальные «длинные волны». На самом деле эти ординаты отражают неравномерность значений $\varphi(t_i)$ на сколь угодно длинных интервалах времени. Если мы рассмотрим конкретную реализацию углов волнового склона на интервалах времени $t_i \leq t \leq t_{i+1}$, то на одних интервалах будет, например, преобладание положительных значений, на других – отрицательных, и эта неравномерность, пусть и небольшая, сохранится на интервалах времени любой длины, какой большой не была разность $t_{i+1} - t_i$ (именно эта неравномерность приводит, в первую очередь, к сносу судна, компенсировать который должны системы динамической стабилизации). А в спектре $S_\varphi(\omega)$ эта неравномерность отразится тем, что при малых частотах, в том числе и при $\omega = 0$, будет $S_\varphi(\omega) \neq 0$. Заметим, что если $S_\varphi(0) \neq 0$, то это совсем не означает (как иногда думают), что в процессе $\varphi(t)$ присутствует постоянная составляющая; «телеграфный сигнал» со средним значением, равным нулю и корреляционной функцией $K_\varphi(\tau) = a^2 e^{-2\mu\tau}$ имеет спектр

$$S_{\varphi}(\omega) = a^2 \frac{8\mu}{\pi(4\mu^2 + \omega^2)}, \quad S_{\varphi}(0) = \frac{a^2}{\pi} \frac{2}{\mu}.$$

Мы убеждаемся, что хотя использование дополнительной информации остается одним из самых сильных методов подхода к некорректным (или близким к некорректным) задачам, но выбирать эту дополнительную информацию, нужно осторожно и обосновано. Ошибка в выборе дополнительной информации приводит к ошибке в решении основной задачи, что и происходило первоначально с расчетами систем динамической стабилизации. Перелом произошел после 1985 года, когда было предложено ввести для описания возмущающих воздействий на буровые и разведующие нефть суда корреляционную функцию уже не с двумя, а с тремя параметрами:

$$K_{\varphi}(\tau) = \langle \varphi^2 \rangle e^{-\alpha|\tau|} (\cos \beta \tau + \gamma \sin \beta |\tau|), \quad (3.65)$$

которой соответствует спектр

$$S_{\varphi}(\omega) = \langle \varphi^2 \rangle \frac{2\alpha}{\pi} \frac{\alpha(\alpha^2 + \beta^2 + \omega^2) + \gamma\beta(\alpha^2 + \beta^2 + \omega^2)}{(\alpha^2 + \beta^2 + \omega^2)^2 - 4\beta^2\omega^2}. \quad (3.66)$$

Значение спектра при $\omega = 0$:

$$S_{\varphi}(0) = \langle \varphi^2 \rangle \frac{2}{\pi} \frac{\alpha + \gamma\beta}{(\alpha^2 + \beta^2)} \quad (3.67)$$

зависит от γ . Практика расчетов для различных интенсивностей волнения показывает, что обычно коэффициент γ отрицателен, по модулю не превосходит 0,1, и поэтому почти не влияет на спектр вблизи точки максимума, что сразу облегчает расчеты. Раньше старались избегать спектров с тремя параметрами, поскольку трудно подобрать сразу три параметра, удовлетворяющих экспериментальным данным. Особенности спектра (3.66) позволяют подобрать их не одновременно, а последовательно: параметр β равен частоте максимума спектра, отношение

$\frac{\alpha}{\beta}$ для морского волнения заключено в узких пределах: $0.1 < \frac{\alpha}{\beta} < 0.25$

и определяется по скорости затухания корреляционной функции традиционными методами, как и ранее для спектра Рахманина–Фирсова. После того, как α и β определены, нужно по возможности точно определить $S_{\varphi}(0)$ и затем вычислить γ по формуле (3.67).

Трехпараметрический спектр (3.66), предложенный и обоснованный в монографии: Петров Ю.П., Червяков В.В. «Системы стабилизации буровых судов», первое издание которой вышло в 1985 году, а второе, дополнительное, в 1997 году, в издательстве С-Петербургского технического университета, обеспечил достоверность расчетов систем

динамической стабилизации для судов, ведущих разведку и добычу полезных ископаемых с морского дна.

3.6. Дифференцирование функций, отягощенных помехами

Поскольку идеальных измерительных приборов, приборов с нулевой погрешностью, не существует, то любая непрерывно измеряемая переменная величина всегда отягощена неизбежной помехой.

Мы хотим измерить функцию $f_1(t)$, а на выходе измерительного прибора получаем

$$f_2(t) = f_1(t) + \varphi(t), \quad (3.68)$$

где $\varphi(t)$ – неизвестный нам, и, как правило, случайный процесс. Его невозможно обратить в нуль. Однако необходимо (и возможно) добиваться малой величины погрешности $\varphi(t)$, а точнее – малого отношения погрешности (помехи) $\varphi(t)$ к функции $f_2(t)$, т.е. следует добиваться малости дроби $\varphi(t)/f_2(t)$, что уже автоматически будет означать хорошее приближение измеренной величины $f_2(t)$ к неизвестной нам истинной функции $f_1(t)$.

Новые трудности возникают в задаче дифференцирования, когда нам нужна не только сама функция $f_1(t)$, но и ее производная $\frac{df_1}{dt}$, а измерительный прибор может дать нам лишь сумму $f_1(t) + \varphi(t)$.

Задача дифференцирования, задача восстановления истинной производной функции $f_1(t)$ по измерениям, отягощенным помехами, часто возникает в теории и в практике автоматического управления и очень важна для них.

Поскольку даже малая по абсолютной величине функция может иметь большую производную, то задача дифференцирования функции, отягощенной помехами, часто может быть некорректной. Сколь бы малой ни была разность между измеренной величиной $f_2(t)$ и истинной функцией $f_1(t)$, разность их производных

$$\frac{df_2(t)}{dt} - \frac{df_1(t)}{dt} = \frac{d\varphi(t)}{dt} \quad (3.69)$$

может быть велика. Еще чаще эта задача оказывается плохо обусловленной, когда малой (но конечной) разности функций $f_2(t) - f_1(t)$ соответствует гораздо более существенная разница их производных за счет того, что погрешность (помеха) $\varphi(t)$ очень часто, как общее правило, имеет более высокую частоту изменения.

Поставим теперь задачу о поиске оптимального оператора, который наилучшим образом решает задачу дифференцирования функции, отягощенной помехами. Поскольку задача о поиске оптимального оператора в необъятном море всех возможных операторов слишком сложна и на сегодняшний день не изобретено методов ее решения, мы сузим задачу и будем искать оптимальный линейный оператор, составленный из всевозможных комбинаций оператора дифференцирования

$D = \frac{d}{dt}$ и оператора интегрирования $\frac{1}{D} = \int_0^t dt$, т.е. мы, будем искать

оператор вида $\frac{B(D)}{A(D)}$, где $A(D)$, $B(D)$ – полиномы от оператора дифференцирования $D = \frac{d}{dt}$.

Обозначим, искомый оператор через L . Из его линейности следует, что

$$L[x_1(t) + x_2(t)] = L[x_1(t)] + L[x_2(t)], \quad (3.70)$$

т.е. результат применения оператора к сумме функций $x_1 + x_2$ равен сумме результатов применения операторов к каждому слагаемому в отдельности. Заметим, что примером линейного оператора является как раз оператор дифференцирования $D = \frac{d}{dt}$. Формуле (3.70) соответствует известное правило: производная суммы двух функций равна сумме их производных.

Итак, пусть на вход искомого линейного оператора (и реализующего этот оператор технического устройства) поступает сумма измеряемой функции $x(t)$ и помехи $\varphi(t)$ и будем искать оператор L , который преобразовывал бы сумму $x + \varphi$ так, чтобы разность между преобразованной суммой $L(x + \varphi)$ и производной \dot{x} , т.е. разность

$$L(x + \varphi) - \frac{d}{dt} x = y, \quad (3.71)$$

была бы минимальной.

Из линейности оператора L следует

$$y = L(x) - L(\varphi) - \frac{d}{dt} x = \left(L - \frac{d}{dt} \right) (x) + L(\varphi). \quad (3.72)$$

Теперь применим ко всем членам равенства (3.72) интегральное косинус преобразование Фурье. После преобразования место функций y, x, φ займут их спектры $S_y(\omega)$, $S_x(\omega)$, $S_\varphi(\omega)$, а место операторов займут квадраты модулей их Фурье преобразований, (т.е. квадраты

модулей частотных характеристик этих операторов). Получим:

$$S_y(\omega) = |L_\omega - j\omega|^2 S_x(\omega) + |L_\omega|^2 S_\varphi(\omega). \quad (3.73)$$

В равенстве (3.73) через L_ω обозначено косинус преобразование Фурье оператора L , являющееся функцией от частоты ω (напомним, что косинус преобразование оператора дифференцирования $\frac{d}{dt}$ равно $j\omega$). Из (3.73) следует

$$\langle y^2 \rangle = \int_0^\infty S_y(\omega) d\omega = \int_0^\infty |L_\omega - j\omega|^2 S_x + |L_\omega|^2 S_\varphi d\omega. \quad (3.74)$$

Минимум $\langle y^2 \rangle$ обеспечит та функция L_ω , которая доставит минимум подынтегральному выражению (в этом, кстати, и заключается польза перехода к преобразованию Фурье: непосредственно найти оператор L , доставляющий минимум разности (3.71), затруднительно, поскольку методы поиска оптимальных операторов мало разработаны. В то же время найти функцию L_ω , доставляющую минимум интегралу (3.74), совсем нетрудно: достаточно воспользоваться известными методами вариационного исчисления). Применив их, получим

$$L_{\omega \text{ опт}} = j\omega \frac{S_x(\omega)}{S_x(\omega) + S_\varphi(\omega)}, \quad (3.75)$$

квадрат модуля оптимального оператора равен:

$$|L_{\omega \text{ опт}}|^2 = \omega^2 \frac{(S_x)^2}{(S_x + S_\varphi)^2}.$$

Подставив (3.75) в (3.74), получаем наименьший из возможных средний квадрат погрешности дифференцирования:

$$\langle y^2 \rangle = \int_0^\infty \omega^2 \frac{S_x S_\varphi}{S_x + S_\varphi} d\omega. \quad (3.76)$$

Предельный случай при $S_\varphi \rightarrow 0$ соответствует отсутствию помехи и $S_\varphi(\omega) = 0$. В этом предельном случае, как и следовало ожидать,

$$L_{\omega \text{ опт}} = j\omega, \quad L_{\text{опт}} = \frac{d}{dt}, \quad (3.77)$$

т.е. если помехи нет, можно просто обычным образом дифференцировать функцию $x(t)$, затруднений не возникает. Если помеха есть, то, естественно, все сложнее. Кроме того, точная реализация оператора (3.75) не всегда возможна и приходится ограничиваться приближением к нему, исключая нереализуемые множители, но в целом использова-

ние информации о спектрах дифференцируемой функции и помехи позволяет существенно снизить ошибку дифференцирования.

Пример. Рассмотрим задачу дифференцирования функции $x(t)$, имеющей средний квадрат $\langle x^2 \rangle = 1$, корреляционную функцию $K_\varphi(\tau) = e^{-\alpha\tau}(1 + \alpha\tau)$ и спектр

$$S_x = \frac{4}{\pi} \frac{\alpha^3}{(\alpha^2 + \omega^2)^2}. \quad (3.78)$$

В измерениях функции $x(t)$ присутствует погрешность $\varphi(t)$ со средним квадратом $\langle \varphi^2 \rangle$ и спектром

$$S_\varphi = \langle \varphi^2 \rangle \frac{2}{\pi} \frac{\beta}{\beta^2 + \omega^2}, \quad (3.79)$$

т.е. реально нам приходится дифференцировать сумму $x(t) + \varphi(t)$. Как мы уже указывали, спектр (3.79) соответствует недифференцируемому процессу $\varphi(t)$. Отметим, что если реальные физические процессы имеют конечную скорость изменения и поэтому дифференцируемы, то погрешность измерения может изменяться скачками, может иметь вид, показанный ранее на рис. 3.2, и поэтому вполне может иметь спектр, соответствующий формуле (3.79). Наличие недифференцируемой помехи делает некорректной задачу дифференцирования сигнала $x(t) + \varphi(t)$, отягощенного помехой $\varphi(t)$.

Действительно, вычисляя средний квадрат производной сигнала $y = x + \varphi$ (т.е. вычисляя результат преобразования сигнала $y = x + \varphi$ простым оператором дифференцирования $D = \frac{d}{dt}$), получаем:

$$\langle y^2 \rangle = \int_0^\infty \frac{4}{\pi} \frac{\alpha^3 \omega^2}{(\alpha^2 + \omega^2)^2} d\omega + \int_0^\infty \langle \varphi^2 \rangle \frac{2}{\pi} \frac{\beta \omega^2}{\beta^2 + \omega^2} d\omega. \quad (3.80)$$

Из формулы (3.80) сразу видно, что если первый несобственный интеграл сходится и равен α^2 , то второй интеграл расходится при любом, даже очень малом, среднем квадрате $\langle \varphi^2 \rangle$ помехи $\varphi(t)$. Погрешность дифференцирования при любой малой помехе $\varphi(t)$ со спектром (3.79) оказывается бесконечно большой.

Для подхода к этой некорректной задаче можно использовать дополнительную информацию, заключенную в спектрах полезного сигнала $x(t)$ и помехи $\varphi(t)$, и вместо оператора дифференцирования $\frac{d}{dt}$ использовать оператор, преобразование Фурье которого имеет вид

(3.75), или близкий к нему. Нетрудно проверить (пользуясь формулой (3.76)), что погрешность дифференцирования будет в этом случае чаще всего небольшой величиной.

Если затруднительно получить точные спектры полезного сигнала $x(t)$ и помехи $\varphi(t)$, то можно пользоваться приближенным правилом: если в спектре помехи значительную роль играют более высокие частоты, чем в полезном сигнале, то вместо оператора дифференцирования $\frac{d}{dt}$ следует использовать оператор, для которого квадрат модуля преобразования Фурье убывает с ростом частоты. Чем меньше квадрат модуля оператора будет отличаться от оптимального, определяемого равенством (75), тем меньше будет погрешность дифференцирования.

Для иллюстрации на рис. 3.4 показаны спектры полезного сигнала (3.78) при $\alpha = 1$, спектр помехи (3.79) при $\langle \varphi^2 \rangle = 1$ и $\beta = 10$, а пунктирной линией показано отношение квадрата модуля оптимального оператора (3.75) к квадрату модуля частотной характеристики оператора идеального дифференцирования $\frac{d}{dt}$. Помеха имеет спектр, медленнее спадающий с ростом частоты ω , чем спектр полезного сигнала, поэтому квадрат модуля частотной характеристики оптимального фильтра с ростом частоты ω быстро убывает.

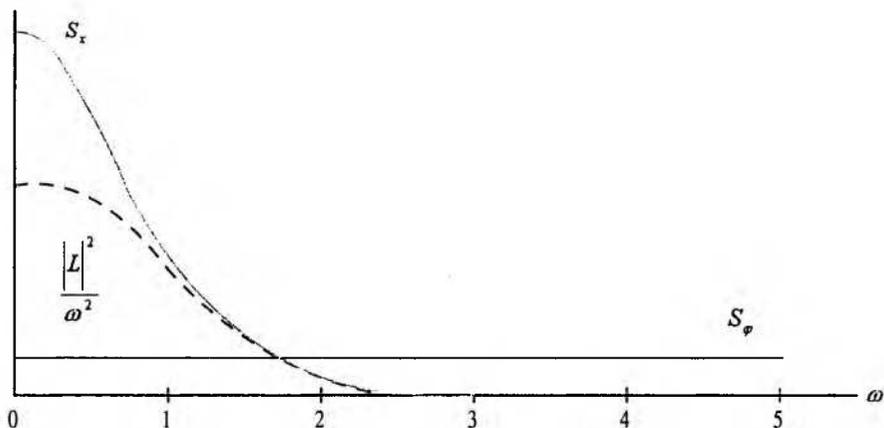


Рис. 3.4

Аналогичные соображения позволяют решить и более простую задачу оптимальной фильтрации, т.е. задачу наилучшего выделения сигнала $x(t)$ из показаний измерительного прибора $x(t) + \varphi(t)$, отягощенных погрешностью $\varphi(t)$. Будем предполагать, что нам известны сред-

ние квадраты $\langle x^2 \rangle$ и $\langle \varphi^2 \rangle$ сигнала $x(t)$ и погрешности $\varphi(t)$, известны их спектры $S_x(\omega)$, $S_\varphi(\omega)$ и поставим задачу поиска линейного оператора L , доставляющего минимум разности между выходом оператора $L(x + \varphi)$ и искомым сигналом $x(t)$:

$$L(x + \varphi) - x = y. \quad (3.81)$$

Из линейности оператора L следует:

$$y = L(x) - L(\varphi) - x = (L - 1)(x) + L(\varphi). \quad (3.82)$$

Теперь применим ко всем членам равенства (3.82) интегральное косинус преобразование Фурье. В результате преобразования функции $y(t)$, $x(t)$, $\varphi(t)$ перейдут в свои спектры $S_y(\omega)$, $S_x(\omega)$, $S_\varphi(\omega)$, а место операторов L и $L - 1$ займут квадраты модулей их Фурье преобразований, которые являются функциями от ω и обозначаются L_ω и $L_\omega - 1$:

$$S_y(\omega) = |L_\omega - 1|^2 S_x(\omega) - |L_\omega|^2 S_\varphi(\omega), \quad (3.83)$$

из (3.83) следует:

$$\langle y^2 \rangle = \int_0^\infty S_y(\omega) d\omega = \int_0^\infty [|L_\omega - j\omega|^2 S_x + |L_\omega|^2 S_\varphi] d\omega. \quad (3.84)$$

Для отыскания оптимального оператора, доставляющего минимум среднему квадрату $y(t)$, достаточно найти функцию L_ω , доставляющую минимум интегралу (3.84). Взяв производную по L_ω и приравняв ее к нулю, находим:

$$L_{\omega \text{ опт}} = \frac{S_x}{S_x + S_\varphi}, \quad (3.85)$$

а подставив (3.85) в (3.84), получим наименьшую возможную ошибку фильтрации

$$\langle y^2 \rangle_{\min} = \int_0^\infty \frac{S_x S_\varphi}{S_x + S_\varphi} d\omega. \quad (3.86)$$

Пример. Пусть на вход системы управления поступает полезный сигнал $x(t)$ с корреляционной функцией $K_x = e^{-\tau}$ и спектром

$$S_x = \frac{1}{1 + \omega^2}, \quad (3.87)$$

смешанный с погрешностями измерения, помехами и т.п., являющимися «белым шумом» со спектром

$$S_\varphi = 1 \quad (3.88)$$

(т.е. с бесконечным средним квадратом; реально это означает, что сигнал сильно засорен, мощность помехи велика, и ее спектр можно ап-

проксимировать простой зависимостью (3.88)).

Оптимальный оператор L_ω в этом случае равен

$$L_{\omega \text{ опт}} = \frac{1}{2 + \omega^2}, \quad (3.89)$$

квадрат его модуля равен

$$|L_\omega|^2 = \frac{1}{4 + 4\omega^2 + \omega^4}. \quad (3.90)$$

Средний квадрат погрешности

$$\langle y^2 \rangle = \int_0^\infty \frac{(1 + \omega^2)^{-1}}{(1 + \omega^2)^{-1} + 1} d\omega = \int_0^\infty \frac{1}{2 + \omega^2} d\omega = \frac{\sqrt{2}\pi}{4}. \quad (3.91)$$

Фильтрация может быть реализована динамической системой, описываемой дифференциальным уравнением:

$$(D^2 + \sqrt{8}D + 2)y = x + \varphi. \quad (3.92)$$

Поскольку спектр помехи постоянен для всех частот ω , а спектр полезного сигнала убывает с ростом частоты, то отсюда, естественно, следует, что квадрат модуля частотной характеристики фильтра должен убывать с ростом ω , что и отражает формула (3.90). Если точная реализация динамической системы с квадратом модуля частотной характеристики (3.75) или (3.85) затруднительна, то можно использовать более простую динамическую систему, стараясь, чтобы ее частотная характеристика примерно соответствовала формулам (3.75) или (3.85).

Так, вместо динамической системы (3.92) можно использовать более простую:

$$(\sqrt{2} + D)y = x + \varphi, \quad (3.93)$$

для которой

$$\langle y^2 \rangle = \int_0^\infty S_y(\omega) d\omega = \int_0^\infty \frac{S_x + S_\varphi}{2 + \omega^2} d\omega = \frac{\pi}{2}. \quad (3.94)$$

Погрешность фильтрации динамической системы (3.93) больше, чем у (3.92), но остается вполне приемлемой.

3.7. Предсказание будущего

В настоящем разделе мы рассмотрим задачу, которая принципиально не может быть решена точно; мы рассмотрим задачу предсказания будущих значений случайного процесса $\varphi(t)$ по наблюдениям его прошлого и настоящего.

Условие задачи: некоторое время мы наблюдаем процесс $\varphi(t)$, за это время мы убедились, что процесс стационарен и вычислили его

корреляционную функцию $K_\varphi(\tau)$ и спектр $S_\varphi(\omega)$. Отметим при этом, что корреляционную функцию можно вычислить с достаточной точностью за весьма умеренное время: если она, например, имеет вид (3.26), где $\mu = \frac{1}{\text{сек}}$, то уже через 7 секунд наблюдения и измерения можно прекратить, поскольку при $\tau \geq 7$ секунд значения корреляционной функции (3.26) будут менее одной тысячной от ее максимального значения при $\tau = 0$; поэтому наблюдения за пределами $\tau > 7$ секунд для обычных технических задач смысла уже не имеют. Если $\mu \neq \frac{1}{\text{сек}}$, то наблюдения достаточно вести на интервале времени $0 \leq t \leq 7 \frac{1}{\mu}$.

Итак, пусть к моменту времени $t = 0$ мы знаем корреляционную функцию и спектр $S_\varphi(\omega)$ процесса $\varphi(t)$, можем измерить $\varphi(0)$ и $\dot{\varphi}(0)$, но другой информации о процессе не имеем. Как наилучшим образом предсказать будущие значения процесса $\varphi(t)$, его значения при $t > 0$?

Эта задача была впервые поставлена и решена выдающимся американским математиком Норбертом Винером (1894–1964) в 1943 году. Шла вторая мировая война и Винер предложил метод предсказания положения неприятельского самолета для улучшения точности огня зенитной артиллерии.

Метод предсказания, предложенный Н. Винером, был сложен. Он приведен в его известной книге «Кибернетика» (стр. 81–123 русского издания 1958 года), а когда этот метод впервые был опубликован в 1943 году в небольшой секретной книге в желтой обложке, то эту книгу работники проектно-конструкторских организаций дружно называли «желтой опасностью». Они чувствовали, что метод Винера может помочь значительно увеличить эффективность артиллерийского огня, но понять метод Винера и использовать его было практически невозможно. Многие сотни тысяч часов столь дорогого в военных условиях времени инженеров были потрачены напрасно на попытки понять «желтую опасность». Было даже предложение сбросить книгу Винера над Германией, чтобы и немецкие инженеры потратили дорогое время на попытки понять Н. Винера.

Мы дадим гораздо более простой вывод, не претендующий на строгость, основанный на вычислении спектра последовательности импульсных функций, т.е. таких функций, для которых средний квадрат (мощность) равен нулю, поскольку

$$\langle \varphi^2 \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \varphi^2 dt \rightarrow 0, \quad (3.95)$$

но интеграл

$$\mathcal{E}_\varphi = \lim_{T \rightarrow \infty} \int_{-T}^T \varphi^2 dt \quad (3.96)$$

существует и конечен. Названия «мощность» и «энергия» связаны с тем, что если $\varphi(t)$ – это электрический ток, то его квадрат $\varphi^2(t)$ – это мгновенная мощность тока, выделяемая в сопротивлении $r=1$ Ом, средний квадрат (3.95) – это средняя мощность тока, а интеграл (3.96) – это энергия тока, которая выделяется в сопротивлении 1 Ом за все время $-\infty < t < \infty$ (отсюда, кстати, и происходит полное название спектра $S_\varphi(\omega)$ – спектральная плотность мощности; в последнее время предпочитают более короткий термин «спектр»). По аналогии с корреляционной функцией введем функцию корреляции $r_\varphi(\tau)$, определив ее равенством

$$r_\varphi = \int_0^\infty \varphi(t) \varphi(t + \tau) dt = \int_{-\infty}^0 \varphi(t) \varphi(t - \tau) dt, \quad (3.97)$$

и введем спектральную плотность энергии $\mathcal{E}_\varphi(\omega)$, определив ее как косинус преобразование Фурье функции корреляции:

$$\mathcal{E}_\varphi(\omega) = \frac{2}{\pi} \int_0^\infty r_\varphi(\tau) \cos \omega \tau d\tau. \quad (3.98)$$

Обратное косинус преобразование

$$\int_0^\infty \mathcal{E}_\varphi(\omega) \cos \omega \tau d\omega = r_\varphi(\tau) \quad (3.99)$$

восстанавливает функцию корреляции $r_\varphi(\tau)$, а положив в формуле (3.99) $\tau = 0$, получим:

$$r_\varphi(0) = \int_0^\infty \varphi^2 dt = \int_0^\infty \mathcal{E}_\varphi(\omega) d\omega = \mathcal{E}_\varphi, \quad (3.100)$$

т.е. энергия импульсной функции $\varphi(t)$ равна интегралу от спектральной плотности энергии.

Пример. Рассмотрим экспоненциальный импульс, т.е.

$$\varphi(t) = \begin{cases} \alpha e^{-\alpha t}, & 0 \leq t < \infty, \\ 0, & -\infty < t < 0. \end{cases} \quad (3.101)$$

Для нее

$$r_{\varphi} = \alpha^2 \int_0^{\infty} e^{-\alpha t} e^{-\alpha(t+\tau)} dt = \frac{\alpha}{2} e^{-\alpha|\tau|}, \quad (3.102)$$

а применив косинус преобразование, получим:

$$\mathfrak{E}_{\varphi}(\omega) = \frac{2}{\pi} \int_0^{\infty} \frac{1}{2\alpha} e^{-\alpha\tau} \cos \omega\tau d\tau = \frac{1}{\pi} \frac{1}{\alpha^2 + \omega^2}. \quad (3.103)$$

Мы убеждаемся, что спектральная плотность энергии для экспоненциальной функции отличается от спектральной плотности мощности экспоненциальной корреляционной функции только постоянным множителем.

Рассмотрим теперь случайную последовательность экспоненциальных импульсов разного знака $\pm \alpha e^{-\omega t}$, расположенных во времени случайно по закону Пуассона. За единицу времени приходит в среднем μ импульсов. Такая последовательность показана на рис. 3.5.

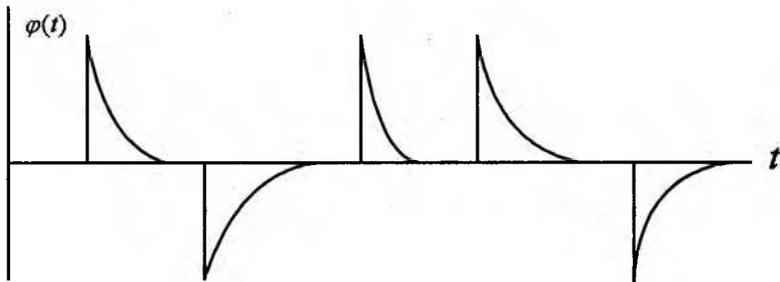


Рис. 3.5

Пусть чередование знаков таково, что среднее значение последовательности импульсов равно нулю, а расположены они так редко, что один импульс успевает практически полностью затухнуть к приходу следующего. Поскольку спектральная плотность энергии импульса $\pm \alpha e^{-\alpha t}$ равна

$$\mathfrak{E}_{\varphi}(\omega) = \frac{1}{\pi} \frac{1}{\alpha^2 + \omega^2} \quad (3.104)$$

и за единицу времени приходит в среднем μ импульсов, то спектральная плотность мощности (спектр) последовательности импульсов равна

$$S_{\varphi}(\omega) = \frac{\mu\alpha}{2} \frac{1}{\pi \alpha^2 + \omega^2}, \quad (3.105)$$

т.е. всего лишь постоянным множителем отличается от уже неоднократно рассмотренных нами процессов со спектром (3.34). Мы уже упоминали, что корреляционная функция и спектр не являются исчерпывающими характеристиками процесса $\varphi(t)$. Могут существовать различные процессы, имеющие один и тот же спектр (их называют реализациями процесса со спектром $S_\varphi(\omega)$). Мы уже убедились, что экспоненциальную корреляционную функцию $K_\varphi = e^{-\alpha t}$ и спектр (3.34) имеют реализации в виде «телеграфного сигнала», (рис. 3.1), а также реализации в виде функции, показанной на рис. 3.2.

Рис. 3.5 изображает еще одну реализацию случайного процесса с экспоненциальной корреляционной функцией, реализацию в виде случайной последовательности экспоненциальных импульсов.

В пределе, при $\alpha \rightarrow \infty$, каждый импульс $\alpha e^{-\alpha t}$ переходит в δ -функцию, случайная последовательность импульсов переходит в случайную последовательность δ -функций, а спектр (3.105) при $\alpha \rightarrow \infty$ переходит в

$$S_\varphi(\omega) = \frac{\mu}{\pi}, \quad (3.106)$$

т.е. в постоянную величину. Процесс с постоянным для всех частот спектром называют, как мы уже упоминали, «белым шумом». Мы убеждаемся, что одной из реализаций «белого шума» является случайная последовательность δ -функций разного знака. Эта реализация изображена (условно) на рис. 3.6. Поскольку амплитуда δ -функции бесконечна, ее условно изображают «стрелкой с наконечником», символизирующей бесконечность амплитуды.

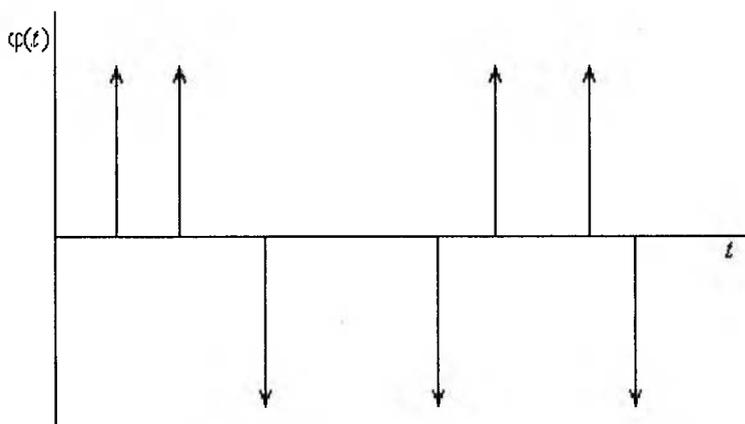


Рис. 3.6

Теперь подготовлены все предпосылки для ответа на главный вопрос: для случайного процесса $\varphi(t)$ по наблюдениям за ним до момента времени $t = 0$ вычислен спектр процесса и корреляционная функция $K_\varphi = \langle \varphi^2 \rangle = e^{-\alpha t}$. В момент $t = 0$ измерено $\varphi(t) = \varphi(0)$. Какое наиболее вероятное суждение о значениях $\varphi(t)$ для $t > 0$ можем мы вынести на основании информации о спектре? Ответ ясен: наилучшее предсказание будущих значений $\varphi(t)$ – это функция $\varphi(0)e^{-\alpha t}$.

По крайней мере, для одной реализации процесса $\varphi(t)$ с $K_\varphi = e^{-\alpha t}$, для реализации в виде случайной последовательности импульсов, показанной на рис. 3.5, это предсказание будет точным предсказанием, для всех других реализаций – приближенным.

Особенно отчетливо приближенность предсказания $\varphi_{\text{предск}} = \varphi(0)e^{-\alpha t}$ видна для реализации $\varphi(t)$ в виде «телеграфного сигнала», также имеющего экспоненциальную корреляционную функцию. На рис. 3.7 пунктирной кривой показано наилучшее предсказание значений $\varphi(t)$ для $t \geq 2$ (т.е. значения процесса $\varphi(t)$ в последний раз измерены для $t = 2$, дальше для $t \geq 2$ идет предсказание).

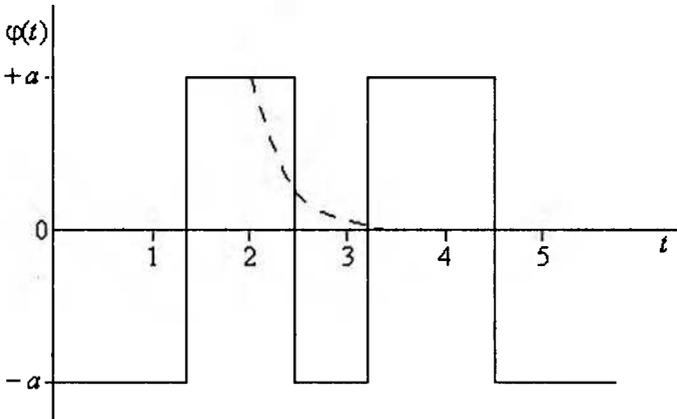


Рис. 3.7

Точность предсказания быстро падает с ростом времени t (а точнее, с ростом произведения αt). Для $t = 0.7 + 0.8$ от $\frac{1}{\alpha}$ точность предсказания еще вполне приемлемая, затем падает ниже и ниже. Это и понятно, поскольку корреляционная функция и спектр дают лишь небольшую информацию о процессе $\varphi(t)$.

Если мы хотим более точного предсказания, то нужны глубокие исследования механизмов, порождающих процесс $\varphi(t)$. Однако корреляционная функция вычисляется очень просто и для малых интервалов предсказания дает хороший прогноз.

Теперь перейдем к более сложной задаче предсказания будущих значений процесса $\varphi(t)$, имеющего спектр

$$S_{\varphi}(\omega) = \frac{A}{\omega^4 + M\omega^2 + N}, \quad (3.107)$$

где $N \geq \frac{M^2}{4}$ (если $N < \frac{M^2}{4}$, то спектр (3.107) можно разложить на сумму спектров вида (3.34), а для них задача уже решена). Если $N \geq \frac{M^2}{4}$, то спектр (3.107) можно преобразовать к виду

$$S_{\varphi} = \frac{A}{(\omega^2 + \alpha^2 + \beta^2)^2 - 4\beta^2\omega^2}, \quad (3.108)$$

где

$$\alpha^2 = \frac{M}{4} + \frac{\sqrt{N}}{2},$$

$$\beta^2 = \frac{\sqrt{N}}{2} - \frac{M}{4}.$$

Пусть теперь на вход динамической системы, описываемой дифференциальным уравнением

$$(D^2 + 2\alpha D + \beta^2 - \alpha^2)\varphi = x(t), \quad (3.109)$$

приходит реализация процесса $x(t)$ в виде случайной последовательности δ -функций. Такая реализация (как мы уже показали) имеет спектр постоянный для всех частот, поэтому спектр процесса $\varphi(t)$ на выходе системы (3.109) будет иметь вид (3.108), а реализацией процесса со спектром (3.108) будет в этом случае случайная последовательность затухающих импульсов вида

$$\varphi_i(t) = \pm A_i e^{-\alpha t} \sin \beta t. \quad (3.110)$$

Каждый импульс порождается очередной δ -функцией последовательности. (Мы рассматриваем такую редкую последовательность δ -функций, что процессы, порожденные одной δ -функцией, успевают затухнуть к приходу следующей). Коэффициенты A_i , стоящие перед импульсами, зависят от величины и знака коэффициентов перед δ -функциями последовательности.

Рассмотрим теперь спектр вида

$$S_{\varphi}(\omega) = \frac{B\omega^2}{\omega^4 + M\omega^2 + N}. \quad (3.111)$$

Процесс со спектром (3.111) является производной процесса со спектром (3.108). Следовательно, одной из реализаций такого процесса может быть случайная последовательность импульсов вида

$$\varphi_i(t) = \pm A_i e^{-\alpha t} (\beta \cos \beta t - \alpha \sin \beta t) \quad (3.112)$$

(каждый импульс является производной от функции (3.110)).

Располагая формулами (3.110) и (3.112), можно найти одну из реализаций процесса, имеющего спектр вида

$$S_{\varphi}(\omega) = \frac{A + B\omega^2}{\omega^4 + M\omega^2 + N}. \quad (3.113)$$

Такой реализацией будет случайная последовательность затухающих импульсов

$$\varphi_i(t) = e^{-\alpha t} (c_1 \cos \beta t + c_2 \sin \beta t), \quad (3.114)$$

первый из которых начинается в точке $t = 0$; постоянные c_1, c_2 зависят от A и B в спектре (3.113).

Теперь сразу делается ясным решение задачи наилучшего предсказания процесса $\varphi(t)$, имеющего спектр (3.113), по значениям $\varphi(t)$ и $\dot{\varphi}(t)$ при $t = 0$. Наилучшим предсказанием будет функция (3.114), в которой постоянные c_1, c_2 определяются на основе измеренных $\varphi(0)$ и $\dot{\varphi}(0)$. Из (3.114) следует, что

$$\begin{cases} c_1 = \varphi(0), \\ c_2 = \frac{\dot{\varphi}(0) + \alpha\varphi(0)}{\beta}. \end{cases} \quad (3.115)$$

С учетом (3.115) оптимальное предсказание будущих (при $t \geq 0$) значений случайного процесса со спектром (3.113) примет вид:

$$\varphi(t \geq 0) = e^{-\alpha t} \left[\varphi(0) \cos \beta t + \frac{\dot{\varphi}(0) + \alpha\varphi(0)}{\beta} \sin \beta t \right]. \quad (3.116)$$

Для реализации (3.114) предсказание (3.116) будет точным (точным до прихода следующего импульса), для всех других реализаций случайного процесса со спектром (3.113) предсказание (3.116) с самого начала будет предсказанием приближенным, но дающим приемлемую точность для времени $t = 0.2 + 0.3$ от $\frac{1}{\alpha}$.

Для морских судов, у которых спектр бортовой качки хорошо описывается спектром (3.108), а типичными значениями α и β являются

значения $\beta = 0.4 + 1 \frac{1}{\text{сек}}$, $\alpha = 0.08 + 0.2 \frac{1}{\text{сек}}$, можно ожидать хорошего предсказания углов крена на время $t = 3 + 5 \text{ сек}$.

Даже такое предсказание нужно признать хорошим, поскольку реализации случайного процесса со спектром (3.108) могут быть очень разнообразными, даже реализации вида (3.110). Простейший случай – это когда импульсы расположены настолько редко, что не перекрывают друг друга. Если же импульсы налагаются один на другой, то реализация процесса $\varphi(t)$ может иметь самый причудливый вид, и в этих условиях предсказание (3.116), дающее для всех разнообразнейших реализаций хорошую точность при $t = 0.2 + 0.3$ от $\frac{1}{\alpha}$ следует признать большим достижением теории случайных процессов.

На рис. 3.8 сплошной линией показана типичная реализация случайного процесса со спектром вида (3.113) – бортовой качки судна на нерегулярном морском волнении. Показан крен судна θ в градусах в функции времени t , а пунктирной кривой показано оптимальное предсказание для $t \geq 20$. При $t = 20$ сек в последний раз измеряется $\theta(20)$ и производная $\dot{\theta}(20)$, а дальше идет предсказание.

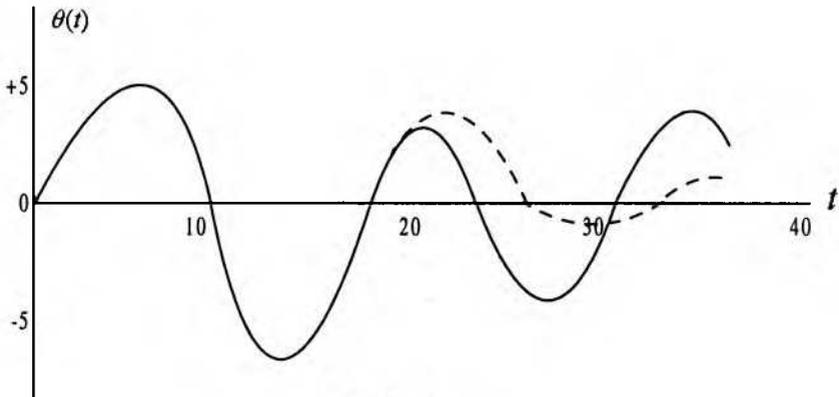


Рис. 3.8

Отметим, что решив задачу предсказания для процессов со спектрами (3.105) и (3.113), мы фактически, решили ее в общем виде, для любого дробно-рационального спектра

$$S_{\varphi}(\omega) = \frac{a_p \omega^{2p} + a_{p-1} \omega^{2p-2} + \dots + a_0}{b_q \omega^{2q} + b_{q-1} \omega^{2q-2} + \dots + b_0}, \quad (3.117)$$

где $p < q$. Поскольку любой спектр является четной функцией, а

спектр (3.117) является функцией от ω^2 , то обозначив $\omega^2 = x$, запишем знаменатель в виде

$$b_q x^q + b_{q-1} x^{q-1} + \dots + b_0 \quad (3.118)$$

и тогда, как известно, дробно-рациональный спектр (3.117) может быть разложен на сумму спектров:

$$S_\varphi(\omega) = \sum_{i=1}^n \frac{c_i}{\alpha_i^2 + \omega^2} + \sum_{i=1}^n \frac{A_i + B_i \omega^2}{N_i + M_i \omega^2 + \omega^4}. \quad (3.119)$$

Каждый из спектров вида

$$S_i = \frac{c_i}{\alpha_i^2 + \omega^2} \quad (3.120)$$

соответствует одному из вещественных корней $x_i = \omega_i^2$ полинома (3.118), а каждый из спектров

$$S_i = \frac{A_i + B_i \omega^2}{\omega^4 + M_i \omega^2 + N_i} \quad (3.121)$$

соответствует каждой паре комплексных сопряженных корней. Постоянные c_i, A_i, B_i в разложении (3.119) определяются по обычным правилам методом неопределенных коэффициентов.

Задача предсказания будущего на основе информации о корреляционной функции (или, что то же самое, о спектре случайного процесса) является интересным примером задачи, в которой погрешность решения зависит не только от погрешности измерения и вычислений, но и от принципиальных причин. Даже если предположить, что мы идеально точно вычислили корреляционную функцию процесса $\varphi(t)$, это вычисление не позволит дать точное предсказание будущих значений $\varphi(t)$ просто потому, что информации, заключенной в корреляционной функции, для точного предсказания недостаточно.

Для более точного предсказания следует использовать подход, уже доказавший свою пользу при решении некорректных задач – нужно привлечь дополнительную информацию об исследуемом процессе; она поможет дать более точное предсказание будущего.

В качестве примера рассмотрим предсказание погоды, точнее – предсказание средней температуры дня на завтра и несколько дней вперед. Температура, а точнее – отклонение температуры каждого текущего дня от хорошо известной, плавно меняющейся в течение года средней многолетней температуры каждого дня, является стационарным случайным процессом, для которого нетрудно вычислить корреляционную функцию (точнее – значения корреляционной функции для $\tau_1 = 1$ день, $\tau_2 = 2$ дня и т.д.). Эти значения, как нетрудно проверить,

чаще всего ложатся примерно на экспоненциальную кривую.

Я вычислял для Московской области и получалось, что $K_{\phi}(\tau = 1 \text{ сутки}) = 0.65$; $K_{\phi}(\tau = 2 \text{ суток}) = 0.42$ и т.п.

Отсюда сразу следует простейшее правило предсказания: если температура сегодняшнего дня, например, превышает среднюю многолетнюю температуру этого дня на 4 градуса, то весьма вероятно, что завтра она будет превышать среднюю многолетнюю на 2.6 градуса, послезавтра – на 1.7 и т.п. Можете сами проверить, что такое элементарное правило прогноза дает несколько лучший результат, чем простая, ничем не подкрепленная, догадка о том, какая температура будет завтра. Но этот элементарный прогноз далеко уступает по точности современной метеорологической службе, которая использует многообразную информацию об атмосферных процессах, движениях циклонов и т.п.

Общее правило: если вы встретились с задачей некорректной или близкой к некорректной, то попробуйте найти дополнительную информацию. Она может помочь и очень часто помогает.

ЛИТЕРАТУРА к Части I

Основная

1. *Абдуллаев Н.Д., Петров Ю.П.* Теория и методы проектирования оптимальных регуляторов. – Л.: Энергоатомиздат, 1985. 240 с.
2. *Андронов А.А., Витт А.А., Хайкин С.Э.* Теория колебаний. – М.: Наука, 1981. 568 с.
3. *Гайдук А.Р.* К исследованию устойчивости линейных систем // Автоматика и телемеханика, 1997, № 3, с. 153–160.
4. *Данилевич Я.Б., Петров Ю.П.* О необходимости расширения понятия эквивалентности математических моделей // Докл. РАН, 2000, т. 371, № 4, с. 473–475.
5. *Зубов В.И.* Математические методы исследований систем автоматического регулирования. – Л.: Машиностроение, 1974. 335 с.
6. *Лаврентьев М.М.* Некорректные задачи для дифференциальных уравнений. – Новосибирск: Изд-во НГУ, 1981. 74 с.
7. *Ларин В.Б. Науменко К.И., Сунцев В.Н.* Спектральные методы синтеза линейных систем с обратной связью. – Киев: Наук. думка, 1971. 137 с.
8. *Летов А.М.* Аналитическое конструирование регуляторов // Автоматика и телемеханика, 1960, № 4, 5, 6; 1961, № 4.
9. *Летов А.М.* Динамика полета и управление. – М.: Наука, 1969. 359 с.
10. *Иванов В.К., Васин В.В., Танана В.П.* Теория линейных некорректных задач и ее приложения. – М.: Наука, 1978. 206 с.
11. *Надеждин В.П.* О потере грубости при элементарных преобразованиях дифференциальных уравнений управляемых систем // Автоматика и телемеханика, 1973, № 1, с. 185–187.
12. *Петров Ю.П.* Оптимизация управляемых систем, испытывающих воздействие ветра и морского волнения. – Л.: Судостроение, 1973. 216 с.
13. *Петров Ю.П.* Вариационные методы теории оптимального управления. – Л.: Энергия, 1977. 280 с.
14. *Петров Ю.П.* Синтез оптимальных систем управления при неполностью известных возмущающих силах. – Л.: Изд-во ЛГУ, 1987. 289 с.
15. *Петров Ю.П.* Устойчивость линейных систем управления при вариациях параметров // Автоматика и телемеханика, 1994, № 11, с. 186–189.
16. *Петров Ю.П.* Третий класс задач физики и техники – промежуточ-

ных между корректными и некорректными. – СПб: Изд-во СПбГУ, 1998. 30 с.

17. *Петров Ю.П. Петров Л.Ю.* Неожиданное в математике и его связь с авариями и катастрофами последних лет. – СПб: Изд-во СПбГУ; 1-е изд., 1999, 108 с.; 3-е изд., 2002, 141 с.
18. *Петров Ю.П.* Лекции по истории прикладной математики. – СПб: Изд-во СПбГУ, 2001. 337 с.
19. *Сергеев В.О.* Некорректно поставленные задачи и методы их решения. – СПб: Изд-во СПбГУ, 1999. 30 с.
20. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. – М.: Наука; 1-е изд., 1974; 2-е изд., 1979; 3-е изд. 1986. 288 с.
21. *Тихонов А.Н., Леонов А.С., Ягола А.Г.* Нелинейные некорректные задачи. – М.: Наука, 1995. 550 с.
22. *Харитонов В.Л.* Об асимптотической устойчивости положения равновесия семейства линейных дифференциальных уравнений // Дифф. уравнения, 1978, № 11.

Дополнительная

23. *Верлань А.Ф., Сизиков В.С.* Интегральные уравнения: методы, алгоритмы, программы. – Киев: Наук. думка, 1986. 544 с.
24. *Сизиков В.С.* Математические методы обработки результатов измерений. – СПб.: Политехника, 2001. 240 с.
25. *Стеклов В.А.* Основы теории интегрирования обыкновенных дифференциальных уравнений. – М.-Л.: ГИЗ, 1927.
26. *Уилкинсон Дж. Х.* Алгебраическая проблема собственных значений. – М.: Наука, 1970. 564 с.
27. *Лаврентьев М.М., Романов В.Г., Шишатский С.П.* Некорректные задачи математической физики и анализа. – М.: Наука, 1980. 287 с.
28. *Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г.* Численные методы решения некорректных задач. – М.: Наука, 1990. 232 с.
29. *Морозов В.А.* Регулярные методы решения некорректно поставленных задач. – М.: Наука, 1987. 240 с.
30. *Бакушинский А.Б., Гончарский А.В.* Некорректные задачи. Численные методы и приложения. – М.: Изд-во МГУ, 1989. 199 с.
31. *Петров Ю.П.* Новые главы теории управления. – СПб: Изд-во СПбГУ, 2000. 156 с.
32. *Петров Ю.П.* Управление, устойчивость, оптимизация (научно-популярные очерки). – СПб: Изд-во СПбГУ, 2002. 94 с.
33. *Петров Ю.П., Фроленков Д.Б.* Изменение корректности при преобразованиях уравнений // Вестник СПбГУ, 2000, сер. 1, вып. 1, с. 52–57.
34. *Галактионов М.А.* О скрытом характере свойства «грубости» ли-

- нейных оптимальных систем управления // Изв. вузов. Электромеханика, 1999, № 4, с. 48–50.
35. *Галактионов М.А.* Структурная неустойчивость в линейных оптимальных системах управления. – СПб: Изд-во СПбГУ, 2001. 34 с.
 36. *Тарарыкин С.В. Тютиков В.В.* Робастное модальное управление динамическими системами // Автоматика и телемеханика, 2002, № 5, с. 41–55.
 37. *Engl H.* Analyse und numerik schlecht gesteller Probleme. – J. Kepler Univ. Linz, 1980.
 38. *Groetsch C.W.* The theory of Tikhonov regularization for Fredholm equations of the first kind. – Boston etc.: Pitman, 1984.
 39. *Engl H.W., Hanke M., Neubauer A.* Regularization of inverse problems. – Dordrecht: Kluwer, 1996.

Часть II

УСТОЙЧИВЫЕ МЕТОДЫ РЕШЕНИЯ ОБРАТНЫХ ЗАДАЧ

В данной части излагаются некоторые *устойчивые методы* решения тех обратных прикладных задач и описывающих их уравнений, которые приведены в главах 1–3, а также будут приведены в главах 4 и 5.

В главе 4 изложены элементы *функционального анализа*, даны некоторые сведения из *линейной алгебры*, приведены основные *типы* рассматриваемых *уравнений* и используемых *преобразований*.

В главе 4 снова сформулировано *определение корректности и некорректности по Адамару* (а также по Тихонову и по Фикера) решения различных уравнений, а именно, операторных уравнений, систем линейных алгебраических уравнений (СЛАУ), систем обыкновенных дифференциальных уравнений (ОДУ), дифференциальных уравнений в частных производных и интегральных уравнений. Корректность и некорректность рассматриваются в различных функциональных пространствах и показывается, что уравнение и задача могут являться корректными или некорректными в зависимости от типа пространств.

В главе 4 изложена *предыстория устойчивых методов* (методы квадратур, преобразования Фурье, наименьших квадратов Гаусса, псевдообратной матрицы Мура–Пенроуза).

Далее изложен *метод регуляризации Тихонова* в общем виде (применительно к операторному уравнению), а также показано, как применять этот метод для решения некорректных интегральных уравнений, СЛАУ, систем ОДУ и дифференциальных уравнений в частных производных. Кратко изложены такие *неустойчивые задачи*, как задача численного дифференцирования, задача вертикального движения ракеты (как задача оптимального управления, описываемая системой ОДУ), задача теории тяготения, описываемая уравнением Лапласа, задача теплопроводности, описываемая уравнением теплопроводности, задача реконструкции смазанного изображения, описываемая или неклассическим интегральным уравнением Вольтерра I рода, или интегральным уравнением Фредгольма I рода типа свертки, и показано, как решать эти задачи методом регуляризации Тихонова.

В главе 4 изложен также *метод решения на компакте*, например, на множестве монотонных ограниченных функций. Этот метод дает

устойчивое решение, однако может использоваться лишь в случае, если априори (из физических соображений) известно, что на искомое решение можно наложить жесткие ограничения, помещающие его в компакт.

В главе 5 изложены: задачи реконструкции смазанных и дефокусированных изображений (описываемые интегральными уравнениями I рода типа свертки), задача рентгеновской томографии (описываемая интегральным уравнением Радона или двухмерным интегральным уравнением Фредгольма I рода типа свертки) и задача синтеза магнитного поля в ЯМР-томографе (описываемая интегральным уравнением Фредгольма I рода с ограничениями на искомое решение).

Главы 4 и 5 содержат ряд иллюстративных численных примеров.

Часть II отличается от части I более строгим и последовательным изложением материала. Поэтому рекомендуется изучать сначала материал части I, а затем – материал части II. Кроме того, в части II изложен более широкий круг задач, уравнений и методов. При этом, однако, имеются повторы по сравнению с частью I в изложении некоторых вопросов (корректность и некорректность, системы ОДУ, СЛАУ, задачи управления и т.д.), но эти вопросы изложены в разном ключе. Поэтому части I и II следует рассматривать как взаимодополняющие.

Глава 4

РЕГУЛЯРНЫЕ МЕТОДЫ РЕШЕНИЯ НЕКОРРЕКТНЫХ ЗАДАЧ

Прежде чем изучать вопросы корректности и некорректности задач и уравнений, а также методы их решения, нужно усвоить основные понятия функционального анализа, а также вспомнить некоторые сведения из линейной алгебры.

4.1. Элементы функционального анализа

Изложим основные понятия функционального анализа [4, 25, 26].

Некоторые обозначения и определения теории множеств: $a \in A$ или $a \notin A$ – элемент a принадлежит или не принадлежит множеству A ; $A \subset B$ – все элементы множества A входят в множество B , другими словами, A есть *подмножество* множества B ; $A \subseteq B$ – то же, но, сверх того, может $A = B$; \emptyset – символ *пустого множества*; $A \cup B$ – *сумма множеств* – множество, состоящее из элементов, принадлежащих хотя бы одному из множеств A или B ; $A \cap B$ – *произведение множеств* – множество, состоящее из элементов, принадлежащих как A , так и B ; *счетным множеством* называется множество, элементы которого можно сопоставить натуральным числам (т.е. занумеровать).

Топологические пространства. Система множеств τ множества X называется *топологией* в X , если: 1) само множество X и пустое множество \emptyset принадлежат τ , 2) сумма $\bigcup_k G_k$ любого числа и произведение

$\bigcap_{k=1}^n G_k$ конечного числа множеств из τ принадлежит τ . Множество X с

заданной в нем топологией τ называется *топологическим пространством* (X, τ) или кратко: T .

Множества, принадлежащие системе τ , называются *открытыми*. Элементы топологического пространства называются *точками*. *Окрестностью* точки $x \in T$ называется всякое открытое множество $G < T$, содержащее точку x ; точка $x \in T$ называется *точкой прикосновения множества* $M \subset T$, если любая окрестность точки x содержит хотя бы одну точку из M ; x называется *предельной точкой множества* M , если каждая окрестность точки x содержит хотя бы одну точку из M , отличную от x . Совокупность всех точек прикосновения множества M назы-

вается замыканием множества M и обозначается \overline{M} .

Множество M топологического пространства T называется *всюду плотным*, если его замыкание есть T , т.е. $\overline{M} = T$. Топологическое пространство со счетным всюду плотным множеством называется *сепарабельным*.

Метрические пространства. Частный случай топологических пространств – метрические пространства. *Метрическим пространством* $R = (X, \rho)$, или кратко X называется множество X элементов (чисел, функций и т.д.) такое, что каждой паре элементов $x, y \in X$ ставится в соответствие неотрицательное вещественное число $\rho(x, y)$ – *расстояние* между x и y , удовлетворяющее аксиомам:

- 1) $\rho(x, y) \geq 0$, причем $\rho(x, y) = 0$ тогда и только тогда, когда $x = y$;
- 2) $\rho(x, y) = \rho(y, x)$ – аксиома симметрии;
- 3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ – аксиома треугольника.

Пусть x_n , $n = 1, 2, 3, \dots$ – некоторая бесконечная последовательность элементов. Элемент x_0 называется *пределом последовательности* x_n , если $\rho(x_n, x_0) \rightarrow 0$ при $n \rightarrow \infty$. Обозначают: $\lim_{n \rightarrow \infty} x_n = x_0$. В этом случае при любом $\varepsilon > 0$ существует такое N , что

$$\rho(x_m, x_n) \leq \varepsilon \text{ при } m, n \geq N, \quad (4.1)$$

а последовательность x_n называется *фундаментальной*.

Но из (4.1), вообще говоря, не следует обратное, т.е. что x_n имеет предел. Если же введено дополнительное условие, что из (4.1) следует существование предела последовательности x_n (т.е. всякая фундаментальная последовательность сходится), то такое метрическое пространство называется *полным*.

Множество U элементов X называется *компактным*, если любая бесконечная последовательность элементов x_n из U содержит сходящуюся последовательность. Замыкание компактного множества называется *компактом*.

Пусть даны два метрических пространства X и Y . Отображение $y = Ax$, ставящее в соответствие элементам $x \in X$ определенные элементы $y \in Y$, называется *оператором* A , действующим из X в Y . Это отображение обозначается как $A: X \rightarrow Y$. Частный случай операторов – *функционалы*. Это – операторы в том случае, когда Y есть пространство чисел с расстоянием $\rho(y_1, y_2) = |y_1 - y_2|$, где $y_1, y_2 \in Y$.

Линейные пространства. *Линейным* (или *векторным*) *пространством* L называется непустое множество элементов x, y, z, \dots , удовле-

творяющее условиям:

I. Каждым двум элементам $x, y \in L$ однозначно поставлен в соответствие третий элемент $z = x + y \in L$, называемый их *суммой*, причем

- 1) $x + y = y + x$ (*коммутативность*),
- 2) $x + (y + z) = (x + y) + z$ (*ассоциативность*),
- 3) существует элемент $0 \in L$ такой, что $x + 0 = x$ (существование нуля),
- 4) для каждого $x \in L$ существует такой элемент $-x$, что $x + (-x) = 0$ (существование *противоположного элемента*).

II. Каждому элементу $x \in L$ и каждому числу α поставлен в соответствие элемент $\alpha x \in L$ (*произведение* элемента x на скаляр α), причем

- 5) $\alpha(\beta x) = (\alpha\beta)x$,
- 6) $1 \cdot x = x$,
- 7) $(\alpha + \beta)x = \alpha x + \beta x$,
- 8) $\alpha(x + y) = \alpha x + \alpha y$.

Элементы x, y, \dots, w линейного пространства L называются *линейно зависимыми*, если существуют такие числа $\alpha, \beta, \dots, \lambda$, не все равные нулю, что $\alpha x + \beta y + \dots + \lambda w = 0$. В противном случае эти элементы называются *линейно независимыми*.

Нормированные пространства. *Нормированным пространством* называется линейное пространство L , в котором задано понятие нормы. Норма элемента $x \in L$ обозначается символом $\|x\|$. *Нормой* $\|x\|$ в L называется вещественное неотрицательное число такое, что

- 1) $\|x\| \geq 0$, причем $\|x\| = 0$ лишь при $x = 0$ (*невыврожденность*),
- 2) $\|\alpha x\| = |\alpha| \cdot \|x\|$ (*однородность*),
- 3) $\|x + y\| \leq \|x\| + \|y\|$ (*неравенство треугольника*).

Если в нормированном пространстве ввести понятие *расстояния* $\rho(x, y) = \|x - y\|$, то оно становится *метрическим пространством*.

Пусть дано множество A , состоящее из элементов x, y, \dots, w линейного нормированного пространства L . Множество S называется *линейной оболочкой* множества A , если его элементами являются линейные комбинации $\alpha x + \beta y + \dots + \lambda w$, где $\alpha, \beta, \dots, \lambda$ – произвольные числа.

Банахово пространство. Полное нормированное пространство называется *банаховым пространством* (B -пространством, пространством типа B , пространством B).

Примеры линейных нормированных пространств, являющихся

банаховыми, и запись нормы в них.

1. Пространство R^1 вещественных чисел:

$$\|x\| = |x|.$$

2. Вещественное n -мерное пространство R^n чисел $x = (x_1, x_2, \dots, x_n)$:

$$\|x\| = \left\{ \sum_{k=1}^n x_k^2 \right\}^{1/2}. \quad (4.2)$$

3. Пространство l_2 суммируемых с квадратом последовательностей чисел $x = (x_1, x_2, \dots, x_k, \dots)$:

$$\|x\| = \left\{ \sum_{k=1}^{\infty} x_k^2 \right\}^{1/2}.$$

4. Пространство $L_2[a, b]$ функций $y(x)$, интегрируемых с квадратом:

$$\|y\| = \left\{ \int_a^b |y(x)|^2 dx \right\}^{1/2} < \infty. \quad (4.3)$$

5. Пространство $L_p[a, b]$ ($p \geq 1$):

$$\|y\| = \left\{ \int_a^b |y(x)|^p dx \right\}^{1/p} < \infty.$$

6. Пространство $L_1[a, b]$ функций $y(x)$, интегрируемых с модулем:

$$\|y\| = \int_a^b |y(x)| dx < \infty. \quad (4.4)$$

7. Пространство $C[a, b]$ непрерывных на $[a, b]$ функций:

$$\|y\| = \max_{a \leq x \leq b} |y(x)|. \quad (4.5)$$

8. Пространство $C^{(n)}[a, b]$ ($n \geq 1$):

$$\|y\| = \max_{a \leq x \leq b} \sum_{k=0}^n \left| \frac{d^k y(x)}{dx^k} \right|.$$

9. Пространство Соболева $W_p^l[a, b]$ функций $y(x)$, l раз непрерывно дифференцируемых:

$$\|y\| = \left\{ \sum_{k=0}^l \int_a^b \left| \frac{d^k y(x)}{dx^k} \right|^p dx \right\}^{1/p}.$$

10. Пространство Соболева $W_2^l[a, b]$ – частный случай пространства 9 – пространство функций $y(x)$, имеющих производную, интегрируемую с квадратом:

$$\|y\| = \left\{ \int_a^b y^2(x) dx + \int_a^b y'^2(x) dx \right\}^{1/2}.$$

Пространства со скалярным произведением (евклидовы, гильбертовы и др. пространства).

Евклидовы пространства. Вещественное линейное пространство E называется *евклидовым*, если каждой паре его элементов x и y поставлено в соответствие вещественное число (x, y) – *скалярное произведение* – такое, что выполняются следующие аксиомы:

- 1) $(x, y) = (y, x)$ (*симметрия*),
- 2) $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ (*аддитивность*),
- 3) $(\lambda x, y) = \lambda(x, y)$ (*однородность*),
- 4) $(x, x) \geq 0$, причем $(x, x) = 0$ только при $x = 0$.

Евклидово пространство превращается в *нормированное пространство*, если в нем определить норму по формуле:

$$\|y\| = \sqrt{(y, y)}.$$

Пусть E – пространство со скалярным произведением. Система ненулевых элементов $x_1, x_2, \dots, x_m \in E$ называется *ортogonalной*, если $(x_k, x_l) = 0$ при $k \neq l$. Если система x_1, x_2, \dots, x_m такова, что

$$(x_k, x_l) = \begin{cases} 0, & k \neq l, \\ 1, & k = l, \end{cases}$$

то она называется *ортонормированной системой*.

Гильбертово пространство. Пространство H со скалярным произведением называется *гильбертовым*, если оно полно в норме, порожденной скалярным произведением. Еще одно определение: *гильбертово пространство* – это банахово пространство, в котором норма определена с помощью скалярного произведения $\|y\| = \sqrt{(y, y)}$ или $\|x - y\| = \sqrt{(x - y, x - y)}$.

Пространства 1–4, 10 являются гильбертовыми пространствами.

Множество $A \subseteq H$ называется *линейно независимым*, если любой конечный набор его элементов является линейно независимым. Мно-

жество $A \subseteq H$ называется *базисом* гильбертова пространства H , если оно линейно независимо и замыкание линейной оболочки A совпадает с H . Базис называется *ортгоналим*, если система его элементов ортогональна, и *ортонормированным*, если система его элементов ортонормирована.

Операторы в гильбертовом пространстве. Пусть заданы некоторые множества Y и F и пусть в Y выделено подмножество $D \subseteq Y$. Если каждому элементу $y \in D$ ставится в соответствие определенный элемент $f \in F$, то говорят, что задан *оператор* $f = Ay$. При этом множество D называется *областью определения оператора* A и обозначается $D(A)$ или $\text{Im } A$. Множество $R \equiv R(A) = \{f \in F : f = Ay, y \in D\}$ называется *областью значений оператора* A . Действие оператора записывается в виде $A: Y \rightarrow F$.

Оператор A , действующий из банахова пространства B_1 в банахово пространство B_2 , называется *линейным*, если для $y_k \in D(A)$ и любых чисел $c_k, k=1, \dots, m$, имеет место равенство $A(c_1 y_1 + \dots + c_m y_m) = c_1 A y_1 + \dots + c_m A y_m$. Оператор A называется *ограниченным*, если $\|Ay\| \leq g \|y\|$ для любого $y \in D(A)$, где g – конечное число. Пример линейного ограниченного оператора – линейный интегральный оператор $A: L_2[a, b] \rightarrow L_2[a, b]$, порождающий функцию

$$Ay = \int_a^b K(x, s) y(s) ds \equiv f(x), \quad a \leq x \leq b.$$

Норма оператора определяется как

$$\|A\| = \sup_{y \in D(A)} \frac{\|Ay\|}{\|y\|} = \sup_{\|y\|=1} \|Ay\|. \quad (4.6)$$

Оператор B , действующий из гильбертова пространства H в H , называется *сопряженным* (обозначается A^*) по отношению к оператору $A: H \rightarrow H$, если $(Ax, y)_H = (x, By)_H$ для любых $x, y \in H$. Например, по отношению к интегральному оператору A сопряженный оператор есть оператор A^* , порождающий функцию

$$A^* y = \int_a^b K^*(x, s) y(s) ds = \int_a^b \overline{K(s, x)} y(s) ds,$$

где черта означает комплексное сопряжение.

Самосопряженный (эрмитов) оператор – это такой линейный ограниченный оператор $A: H \rightarrow H$, что для любых $x, y \in H$ имеет место равенство $(Ax, y)_H = (x, Ay)_H$. *Единичный оператор* – это такой опера-

тор E (или I), что $Ey = y$ для любого $y \in D(A)$. Унитарный оператор – такой оператор U , что $U^*U = E$. Самосопряженный оператор A называется *положительно определенным* (положительным), если $(Ay, y) > 0$ для любого элемента $y \in D(A)$. Пример положительно определенного оператора – оператор A^*A . Обратный по отношению к A оператор (обозначается A^{-1}) – это оператор, определенный на $R(A)$ и сопоставляющий каждому элементу $f \in R(A)$ определенный элемент $y \in D(A)$. Это записывается в виде $y = A^{-1}f$. Из линейности A следует линейность A^{-1} . Но из ограниченности A , вообще говоря, не следует ограниченность A^{-1} (это характерно, в первую очередь, для некорректных задач).

Операторы A и B называются *дистрибутивными* (перестановочными), если $AB = BA$.

Линейный оператор $A: Y \rightarrow F$ называется *непрерывным* в точке $y_0 \in Y$, если $Ay \rightarrow Ay_0$ при $y \rightarrow y_0$. Для линейных операторов непрерывность эквивалентна ограниченности.

Линейный оператор A называется *вполне непрерывным*, если он преобразует всякое ограниченное по норме множество в компактное множество.

Пусть X – линейное пространство и A – линейный оператор, действующий из X в X , с областью определения $D(A)$. Число λ называется *собственным значением оператора* A , если существует вектор $x \neq 0$, $x \in D(A)$ такой, что $Ax = \lambda x$. При этом x называется *собственным вектором* (собственной функцией) оператора A , соответствующим данному собственному значению λ . Если A – линейный самосопряженный оператор, то все λ вещественны, при этом $\|A\| = |\lambda(A)|_{\max}$, $\|A^{-1}\| = 1/|\lambda(A)|_{\min}$. Если A положителен, то все λ вещественны и неотрицательны, при этом $\|A\| = \lambda(A)_{\max}$, $\|A^{-1}\| = 1/\lambda(A)_{\min}$. Если же A – произвольный (линейный) оператор, то все λ – вообще говоря, комплексные, при этом $\|A\| = \sqrt{\lambda(A^*A)_{\max}}$, $\|A^{-1}\| = \sqrt{1/\lambda(A^*A)_{\min}}$. Совокупность всех собственных значений оператора называется его *спектром*.

4.2. Некоторые сведения из линейной алгебры

Напомним некоторые сведения из линейной алгебры [3, с. 156–161], [4, с. 504–509], [7, с. 142–147], [18, 21, 28, 47].

Векторы и матрицы. Прямоугольной $m \times n$ -матрицей называется совокупность чисел (вообще говоря, комплексных), расположенных в виде прямоугольной таблицы из m строк и n столбцов:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \equiv (a_{ij}), \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Если $m > 1, n = 1$, то A есть *вектор-столбец*, а если $m = 1, n > 1$, то A – *вектор-строка*. Если $m = n$, то матрица A называется *квадратной*. Если $a_{ij} = 0$ при $i \neq j$, то квадратная матрица A называется *диагональной*. Если $a_{ij} = 0$ при $i \neq j$ и $a_{ij} = 1$ при $i = j$, то квадратная матрица A называется *единичной* и обозначается E или I . Если все $a_{ij} = 0$, то матрица A называется *нулевой*.

Матрица $A^T = (b_{ij}), i = 1, \dots, n, j = 1, \dots, m$, у которой $b_{ij} = a_{ji}$, называется *транспонированной* по отношению к $m \times n$ -матрице A . В частности, транспонирование переводит вектор-столбец в вектор-строку и обратно. Матрица $\bar{A} = (\bar{a}_{ij}), i = 1, \dots, m, j = 1, \dots, n$, называется *комплексно-сопряженной* с A . Матрица $A^* = \bar{A}^T$ называется *сопряженной* с A (или *эрмитово-сопряженной* или *сопряженной и транспонированной*). При этом $(A^*)^* = A$. Если A вещественна, то $A^* = A^T$.

Определитель (детерминант) квадратной матрицы A обозначается $|A|$ или $\det(A)$. *Минор* порядка k матрицы A есть определитель k -го порядка, составленный из любой части A с соблюдением расположения элементов a_{ij} . *Ранг* $r = \text{rang}(A)$ матрицы A – максимальный порядок k_{\max} отличных от нуля миноров.

Ленточная $m \times n$ -матрица $A = (a_{ij})$ – такая матрица, что $a_{ij} = 0$ при $|i - j| \geq \Delta$, причем $\Delta < \max(m, n)$. *Квазидиагональная* матрица A – такая квадратная $n \times n$ -матрица, у которой вдоль главной диагонали расположены квадратные клетки порядка $< n$, а остальные элементы равны нулю. Ленточным и квазидиагональным (т.е. *редким*) матрицам противостоят *плотные* матрицы – матрицы со сравнительно малым числом нулевых элементов.

Квадратная матрица A называется *вырожденной* (или *особенной*), если $|A| = 0$, в противном случае A называется *невырожденной*.

Квадратная $n \times n$ -матрица A^{-1} называется *обратной* по отношению

к квадратной $n \times n$ -матрице A , если $A^{-1}A = E$. Необходимое и достаточное условие существования A^{-1} – невырожденность A .

Квадратная $n \times n$ -матрица $A = (a_{ij})$ называется *левой*, или *нижней* (соответственно *правой*, или *верхней*) *треугольной* матрицей, если $a_{ij} = 0$ при $j > i$ (соответственно при $j < i$). Для треугольной $n \times n$ -матрицы A справедливо: $|A| = \prod_{i=1}^n a_{ii}$ (определитель равен произведению диагональных элементов).

Сумма диагональных элементов $a_{11} + \dots + a_{nn}$ называется *следом* квадратной $n \times n$ -матрицы $A = (a_{ij})$ и обозначается $\text{Sp } A$ или $\text{Tr } A$.

Квадратная $n \times n$ -матрица $A = (a_{ij})$ называется *симметричной*, если $a_{ij} = a_{ji}$, $i, j = 1, \dots, n$. Вещественная симметричная $n \times n$ -матрица $A = (a_{ij})$ называется *положительно определенной* (*положительной*), если $\sum_{i,j=1}^n a_{ij} x_i x_j > 0$ при любых вещественных x_i и $\sum_{i,j=1}^n a_{ij} x_i x_j = 0$ лишь при $x_1 = x_2 = \dots = x_n = 0$. Примеры положительно определенных матриц: $E, A^T A, A A^T, A^* A, A A^*$ (A прямоугольна).

Квадратная матрица A называется *нормальной*, если $A^* A = A A^*$. Квадратная комплексная матрица A называется *эрмитовой* (*самосопряженной*), если $A = A^*$. Вещественная эрмитова матрица является симметричной. *Унитарная* матрица – такая комплексная матрица A , что $A^* A = E$ или $A A^* = E$. Вещественная унитарная матрица называется *ортогональной* ($A^T A = E$ или $A A^T = E$).

Невырожденную квадратную произвольную матрицу A можно представить в виде $A = LU$, где L и U – соответственно нижняя и верхняя треугольные матрицы. Если при этом A – ленточная матрица, то L и U – треугольные ленточные матрицы.

Положительно определенную матрицу A можно представить в виде $A = LL^T$, где L – нижняя треугольная матрица, в частности, с положительными диагональными элементами (*схема Холецкого*), или в виде $A = LDL^T$, где L – нижняя треугольная матрица с единичной диагональю, а D – диагональная матрица с положительными элементами. Если при этом положительно определенная матрица A является ленточной, то ее можно представить в виде $A = LL^T$, где L – нижняя треугольная ленточная матрица.

Собственные значения и сингулярные числа матриц. Корни λ_i , $i = 1, \dots, n$, *характеристического уравнения*

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \quad (4.7)$$

называются *собственными значениями* или *собственными числами* квадратной $n \times n$ -матрицы $A = (a_{ij})$. Величины $1/\lambda_i$, $i = 1, \dots, n$, называются *характеристическими числами*. *Сингулярными числами* $m \times n$ -матрицы A называются вещественные неотрицательные числа $\mu_i(A) = \sqrt{\lambda_i(A^*A)}$, $i = 1, \dots, n$, обычно расположенные в порядке невозрастания (если A^*A вырождена и имеет ранг $r < n$): $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > \mu_{r+1} = \dots = \mu_n = 0$ или (если A^*A невырождена): $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n > 0$. Справедливо равенство: $|\det(A)| = \mu_1 \mu_2 \dots \mu_n$.

Собственные значения λ_i произвольной (квадратной) матрицы A , вообще говоря, комплексны. Если A симметрична, то все λ_i вещественны (но, в принципе, любого знака). Если A положительна и симметрична, то все λ_i вещественны и неотрицательны. Все сингулярные числа μ_i любой матрицы вещественны и неотрицательны. При этом во всех случаях λ_i и μ_i могут иметь кратность, доходящую до n . Множество собственных значений матрицы называется ее *спектром*.

Если A – прямоугольная $m \times n$ -матрица, то вместо собственных значений обычно используются сингулярные числа $\mu_i(A)$. Если ранг матрицы $r < n$, то матрица A является *вырожденной*, или *особенной*; ее определитель $\det(A) \equiv |A| = 0$ и обратная матрица A^{-1} (при $m = n$) или $(A^*A)^{-1}$ (в общем случае) не существует. Если же $r = n$, то A невырождена, ее определитель $|A| = \sqrt{|A^*A|} = \mu_1 \mu_2 \dots \mu_n$, а обратная матрица A^{-1} (при $m = n$) или $(A^*A)^{-1}$ (в общем случае) существует. При этом в случае $m = n$

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

где A_{ij} – алгебраические дополнения.

Нормы векторов и матриц. Обычно изначально вводится понятие *нормы вектора* $\|x\|$, а затем порожденное ею (подчиненное ей,

согласованное с ней) понятие *нормы матрицы* $\|A\|$, определяемое в общем случае как (ср. (4.6))

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|.$$

Ниже приведены более конкретные нормы вектора и матрицы.

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор-строка, а $x^T = (x_1, x_2, \dots, x_n)^T$ – вектор-столбец в вещественном n -мерном пространстве R^n . В этом случае *евклидова (или шуровская) норма вектора* x есть (ср. (4.2))

$$\|x\| = \left(\sum_{k=1}^n x_k^2 \right)^{1/2} = \sqrt{xx^T}. \quad (4.8)$$

Порожденная нормой (4.8) *евклидова норма* вещественной квадратной $n \times n$ -матрицы A есть

$$\|A\| = \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2} = \mu(A)_{\max}.$$

В случае комплексности x и A норма (4.8) записывается в виде:

$$\|x\| = \left(\sum_{k=1}^n |x_k|^2 \right)^{1/2} = \sqrt{xx^*},$$

при этом нормы $\|x\|$ и $\|A\|$ называются *эрмитовыми*.

Если A – положительно определенная и симметричная матрица, то

$$\|A\| = \mu(A)_{\max} = \lambda(A)_{\max}, \quad \|A^{-1}\| = 1/\mu(A)_{\min} = 1/\lambda(A)_{\min}.$$

Если A – симметричная матрица, то

$$\|A\| = \mu(A)_{\max} = |\lambda(A)|_{\max}, \quad \|A^{-1}\| = 1/\mu(A)_{\min} = 1/|\lambda(A)|_{\min}.$$

Если же A – произвольная (в том числе, прямоугольная), то

$$\|A\| = \mu(A)_{\max} = \sqrt{\lambda(A^*A)_{\max}}, \quad \|A^{-1}\| = 1/\mu(A)_{\min} = 1/\sqrt{\lambda(A^*A)_{\min}}.$$

Используются еще «октаэдрическая» норма вектора (ср. (4.4))

$$\|x\| = \sum_{k=1}^n |x_k|$$

и порожденная ею «октаэдрическая» норма $m \times n$ -матрицы

$$\|A\| = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|,$$

а также «кубическая» норма вектора (ср. (4.5))

$$\|x\| = \max_{1 \leq k \leq n} |x_k|$$

и порожденная ею «кубическая» норма $m \times n$ -матрицы

$$\|A\| = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|,$$

называемые также *максимум-нормами*.

Иногда изначально полагают понятие нормы матрицы, а производным – понятие нормы вектора.

Умножение матриц и векторов. Умножение двух прямоугольных матриц:

$$Z = X \times Y$$

$n \times l$ $m \times n$ $n \times l$

или подробнее:

$$z_{ik} = \sum_{j=1}^n x_{ij} y_{jk}, \quad i = 1, \dots, m, \quad k = 1, \dots, l. \quad (4.9)$$

Умножение матрицы на вектор (частный случай умножения матриц):

$$f = A \times y$$

$n \times l$ $m \times n$ $n \times l$

или подробнее:

$$f_i = \sum_{j=1}^n a_{ij} y_j, \quad i = 1, \dots, m. \quad (4.10)$$

СЛАУ, число обусловленности, МНК и псевдообратная матрица. Запишем *систему линейных алгебраических уравнений (СЛАУ)* в виде:

$$Ay = f \quad (4.11)$$

или

$$\sum_{j=1}^n a_{ij} y_j = f_i, \quad i = 1, \dots, m, \quad (4.12)$$

где A – $m \times n$ -матрица, y – n -вектор-столбец, f – m -вектор-столбец. Если $m > n$ (точнее, число линейно независимых строк в (4.11) больше n), то СЛАУ называется *переопределенной*. Если $m = n$ (и $r = m$), то СЛАУ называется *определенной*. Если же $m < n$, то СЛАУ называется *недоопределенной*.

Можно использовать также следующий, более строгий, критерий для анализа СЛАУ. Для этого наряду с *рангом* $r = \text{rang}(A)$ матрицы A (максимальным порядком отличных от нуля миноров) введем в рассмотрение $\rho = \text{rang}(A|f)$ – *ранг расширенной матрицы*. Тогда если $\rho > r$, то СЛАУ не имеет решений и является переопределенной. Если $\rho = r$, то при $\rho = n$ СЛАУ имеет единственное решение и является определенной, а при $\rho < n$ СЛАУ имеет множество решений и является

ся недоопределенной.

Решение определенной СЛАУ символически записывается в виде: $y = A^{-1} f$, где A^{-1} – обратная матрица, а на практике решение обычно осуществляется гауссовскими (или другими) методами с использованием разложения $A=LU$ – в этом случае задача сводится к решению двух СЛАУ с треугольными матрицами $Lx=f$ и $Uy=x$. Если же A положительна, то наиболее эффективны метод Холецкого и Краута («метод квадратного корня»).

Если вместо точных f и A заданы приближенные \tilde{f} и \tilde{A} такие, что $\|\tilde{f} - f\| \leq \delta$, $\|\tilde{A} - A\| \leq \xi$ (δ и ξ – погрешности задания правой части и матрицы или, более точно, их верхние оценки), то часто используемой оценкой относительной погрешности решения является следующая:

$$\frac{\|\delta y\|}{\|y\|} \leq \text{cond}(A) \left(\frac{\delta}{\|f\|} + \frac{\xi}{\|A\|} \right), \quad (4.13)$$

где $\text{cond}(A) = \text{cond}(A^{-1}) = \|A\| \cdot \|A^{-1}\| = \mu(A)_{\max} / \mu(A)_{\min} \geq 1$ – число обусловленности матрицы A . Если $\text{cond}(A)$ относительно мало (обычно $< 10^3$), то матрица A (и СЛАУ) называется *хорошо обусловленной*. Если же $\text{cond}(A)$ относительно велико (обычно $> 10^4$), то матрица A (и СЛАУ) называется *плохо обусловленной*. Заметим, что малость (по сравнению с единицей) определителя $|A|$, вообще говоря, не есть критерий плохой обусловленности.

Решение переопределенной СЛАУ (псевдорешение) обычно осуществляется *методом наименьших квадратов (МНК)* Гаусса, согласно которому вместо (4.11) решается так называемая *нормальная СЛАУ*

$$A^* A y = A^* f \quad (4.14)$$

с квадратной положительно определенной $n \times n$ -матрицей $A^* A$. СЛАУ (4.14) вытекает из условия

$$\|Ay - f\|^2 = \min_y. \quad (4.15)$$

Величина $\|Ay - f\|$ или $\|Ay - f\|^2$ называется *невязкой* между левой и правой частями СЛАУ (4.11). Решение y , удовлетворяющее условию (4.15), т.е. минимизирующее невязку, называется *псевдорешением*. Подробности – в п. 4.6.

Недоопределенная СЛАУ имеет множество решений. Ее единственное (нормальное) решение может быть получено *методом псевдообратной матрицы* Мура-Пенроуза:

$$y = A^+ f,$$

тельной погрешности решения порядка 10^{-2} и число обусловленности можно считать «небольшим», а СЛАУ – хорошо обусловленной.

Отметим также следующее обстоятельство. Если вычисления выполняются с конечной точностью, то часто невозможно установить, является ли заданная СЛАУ вырожденной (т.е. $\mu(A)_{\min} = 0$) или плохо обусловленной (т.е. $\mu(A)_{\min} > 0$). Поэтому обычно вырожденные и плохо обусловленные СЛАУ рассматривают с одинаковых позиций и применяют одинаковые (устойчивые) методы. Подробности о методах см. в п. 4.7.

Системы обыкновенных дифференциальных уравнений [3, с. 306], [37]. Рассмотрим систему обыкновенных дифференциальных уравнений, записанную в нормальной форме Коши:

$$\frac{dy_i(t)}{dt} = f_i(t, y_1, \dots, y_n), \quad i = 1, \dots, n, \quad (4.17)$$

с начальными условиями (задача Коши):

$$y_i(t^0) = y_i^0. \quad (4.18)$$

Для такой системы справедлива теорема Коши о существовании и единственности решения [3, с. 306].

Теорема Коши. Пусть выполнены следующие условия:

1) функции $f_i(t, y_1, \dots, y_n)$ непрерывны и ограничены (т.е. $|f_i| \leq A$) в замкнутой области $G = \{|t - t^0| \leq a, |y_i - y_i^0| \leq b, i = 1, \dots, n\}$;

2) в области G выполняется условие Липшица:

$$|f_i(t, \tilde{y}_1, \dots, \tilde{y}_n) - f_i(t, y_1, \dots, y_n)| \leq L \sum_{k=1}^n |\tilde{y}_k - y_k|, \quad i = 1, \dots, n,$$

где L – константа Липшица. Тогда система (4.17) с начальными условиями (4.18) имеет решение и притом единственное для $|t - t^0| \leq \alpha$, где $\alpha = \min(a, b/A)$.

Замечание. Условие Липшица выполняется всегда, когда f_i имеют в области G ограниченные частные производные по y_k , т.е.

$$\left| \frac{\partial}{\partial y_k} f_i(t, y_1, \dots, y_n) \right| \leq M, \quad i, k = 1, \dots, n. \quad (4.19)$$

Теперь рассмотрим вопрос об устойчивости решения системы уравнений (4.17) [3, с. 325–326]. Система (4.17) может быть представлена в виде:

$$\frac{dy_i(t)}{dt} = \sum_{j=1}^n a_{ij} y_j(t) + \varphi_i(t, y_1, \dots, y_n), \quad i = 1, \dots, n, \quad (4.20)$$

причем $a_{ij} = \frac{\partial f_i}{\partial y_j}(t; 0, \dots, 0) = \text{const}$, $i, j = 1, \dots, n$, т.е. a_{ij} не зависят от t .

Система

$$\frac{dy_i(t)}{dt} = \sum_{j=1}^n a_{ij} y_j(t), \quad i = 1, \dots, n, \quad (4.21)$$

называется системой, *линеаризованной* по отношению к (4.20).

Назовем $y_i = y_i(t; t^0, y_1^0, \dots, y_n^0)$, $i = 1, \dots, n$, *невозмущенным решением системы* (решением системы (4.17) с начальными данными (4.18)), а $\tilde{y}_i = y_i(t; t^0, \tilde{y}_1^0, \dots, \tilde{y}_n^0)$, $i = 1, \dots, n$, – *возмущенным решением системы* (решением системы (4.17) с возмущенными начальными данными $\tilde{y}_1^0, \dots, \tilde{y}_n^0$). Далее, назовем отклонение возмущенного решения от невозмущенного $\tilde{y}_i = \tilde{y}_i - y_i$ *тривиальным решением*.

Достаточное условие устойчивости тривиального решения дает следующая

Теорема. Пусть

1) *все корни характеристического уравнения линеаризованной системы (4.21), т.е. корни λ уравнения (4.7) или уравнения $\det(a_{ij} - \lambda \delta_{ij}) = 0$ имеют отрицательные действительные части: $\text{Re} \lambda_i < 0$, $i = 1, \dots, n$;*

2) *все функции $\Phi_i(t; y_1, \dots, y_n)$ удовлетворяют условию:*

$$|\Phi_i(t; y_1, \dots, y_n)| \leq M \left\{ \sum_{i=1}^n y_i^2 \right\}^{1/2+\alpha}, \quad i = 1, \dots, n,$$

где $M = \text{const}$ и $\alpha > 0$. Тогда *тривиальное решение системы (4.17) устойчиво (по Ляпунову). Если же хотя бы один корень λ характеристического уравнения имеет положительную действительную часть (и Φ_i удовлетворяет условию 2), то тривиальное решение системы (4.17) неустойчиво.*

Далее будем рассматривать также систему обыкновенных дифференциальных уравнений с управляющей функцией [8, с. 233, 241]:

$$\frac{dy(t)}{dt} = f(t, y, u), \quad (4.22)$$

где $y(t) = \{y_1(t), \dots, y_n(t)\}$ – *искомая вектор-функция на отрезке $t_0 \leq t \leq T$, $f(t) = \{f_1(t), \dots, f_n(t)\}$ – правая часть (вектор-функция), $u(t) = \{u_1(t), \dots, u_m(t)\}$ – управляющая вектор-функция. Начальные условия:*

$$y(t_0) = y_0, \quad (4.23)$$

где y_0 – заданный вектор.

Дальнейшее рассмотрение систем обыкновенных дифференциальных уравнений см. в п. 4.4, 4.7.

Дифференциальные уравнения в частных производных [3, с. 340], [5, 6, 8, 29, 45]. Из всего многообразия уравнений в частных производных рассмотрим наиболее характерные уравнения – уравнение Лапласа и уравнение теплопроводности.

Уравнение Лапласа (уравнение эллиптического типа) в двухмерном случае имеет вид:

$$\Delta u(x, y) \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad (4.24)$$

причем будем рассматривать следующий тип *граничных условий* (*задача Коши*):

$$u(x, 0) = f(x), \quad \left. \frac{\partial u}{\partial y} \right|_{y=0} = \varphi(x), \quad (4.25)$$

где $f(x)$ и $\varphi(x)$ – заданные функции.

Рассмотрим также одномерное *уравнение теплопроводности* (уравнение параболического типа). Будем рассматривать четыре варианта постановки задачи его решения.

1-й вариант (*задача с прямым ходом времени*). Решается *уравнение теплопроводности*

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2}, \quad 0 \leq x \leq l, \quad (4.26)$$

с *начальным условием*: $u(x, 0) = \varphi(x)$ и *граничными условиями*: $u(0, t) = \psi_1(t)$, $u(l, t) = \psi_2(t)$. Функция $u(x, t)$ означает распределение температуры вдоль стержня в момент времени t при условиях, что в начальный момент ($t = 0$) это распределение равнялось $\varphi(x)$ и концы стержня поддерживаются при температурных режимах $\psi_1(t)$ и $\psi_2(t)$. Уравнение решается от $t = 0$ в сторону увеличения t , поэтому данный вариант называется *прямой задачей* или *задачей с прямым ходом времени*.

2-й вариант задачи (с обратным ходом времени). В обратной задаче по распределению температуры вдоль стержня в некоторый момент времени $t_* > 0$, т.е. по заданной функции $u(x, t_*)$ необходимо, решая задачу в направлении уменьшения t , определить начальное распределение температуры $u(x, 0) = \varphi(x)$. При этом граничные условия $\psi_1(t)$ и $\psi_2(t)$ не заданы.

3-й вариант задачи. Это – формулировка задачи теплопроводности в виде задачи Коши: необходимо решить уравнение теплопроводности

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}, \quad -\infty < x < \infty, \quad 0 \leq t \leq T, \quad (4.27)$$

с начальными условиями: $u(x,0) = \varphi(x)$ и $u(x,T) = \chi(x)$, где $\varphi(x)$ и $\chi(x)$ – заданные функции.

4-й вариант задачи. Это, как и 3-й вариант – решение задачи Коши: необходимо решить уравнение теплопроводности (4.27), однако лишь с одним начальным условием: $u(x,0) = \varphi(x)$.

Анализ задач решения дифференциальных уравнений в частных производных дан в п. 4.4, а устойчивые методы решения изложены в п. 4.7.

Интегральные уравнения [4–10]. Мы рассмотрим лишь некоторые типы интегральных уравнений, а именно, уравнения I рода, задача решения которых обычно неустойчива.

Одномерное интегральное уравнение Фредгольма I рода:

$$\int_a^b K(x,s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (4.28)$$

где $K(x,s)$ – ядро (заданная функция), $f(x)$ – правая часть (измеренная функция), $y(s)$ – искомая функция. Типы пространств, к которым принадлежат функции K , f и y , будут в дальнейшем указываться. В уравнении (4.28) ядро $K(x,s)$ является фредгольмовским.

Определение. Ядро $K(x,s)$ называется *фредгольмовским*, если $K(x,s)$ является функцией, непрерывной в прямоугольнике $\Pi = \{a \leq s \leq b, c \leq x \leq d\}$, замкнутой, т.е. однородное уравнение (уравнение (4.28) при $f(x) \equiv 0$) имеет лишь нулевое решение, и удовлетворяющей условию Гильберта–Шмидта:

$$\iint_a^b \int_c^d |K(x,s)|^2 dx ds < \infty. \quad (4.29)$$

Обычно в прикладных задачах $y(s)$ – это входной процесс или сигнал, подлежащий восстановлению, $f(x)$ – измеренный выходной процесс, а $K(x,s)$ – аппаратная функция измерительного устройства.

Одномерное интегральное уравнение Вольтерра I рода:

$$\int_a^x K(x,s) y(s) ds = f(x), \quad a \leq x \leq b, \quad (4.30)$$

где смысл функций K , f и y такой же, как в уравнении (4.28).

Одномерное интегральное уравнение Фредгольма I рода типа свертки:

$$\int_{-\infty}^{\infty} K(x-s) y(s) ds = f(x), \quad -\infty < x < \infty. \quad (4.31)$$

Двухмерное интегральное уравнение Фредгольма I рода типа свертки:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x_1 - s_1, x_2 - s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2), \quad -\infty < x_1, x_2 < \infty. \quad (4.32)$$

Дальнейшее рассмотрение интегральных уравнений дано в п. 4.4–4.8 и в гл. 5.

Операторные уравнения. Все вышеприведенные уравнения или их системы (4.16), (4.17), (4.22), (4.24), (4.26)–(4.28), (4.30)–(4.32) можно записать в виде единого *операторного уравнения I рода*:

$$Ay = f, \quad y \in Y, \quad f \in F, \quad (4.33)$$

где A – заданный линейный оператор (алгебраический, дифференциальный, интегральный и т.д.), f – правая часть (элемент некоторого пространства F), y – искомое решение (элемент некоторого пространства Y), Y и F – некоторые гильбертовы пространства, например, L_2 , C (см. (4.3), (4.5)). В п. 4.4 и 4.7 мы будем активно использовать аппарат операторных уравнений.

Преобразования Фурье, Хартли и Лапласа [3, 4, 7, 15, 17, 38]. Мы рассмотрим также некоторые *интегральные преобразования*.

Непрерывное преобразование Фурье. Пусть задана некоторая кусочно-непрерывная функция (исходный процесс) $y(t)$, $-\infty < t < \infty$, где t – время, линейная координата, угловая координата и т.д. (если t – время, то $y(t)$ – временной процесс). Тогда интеграл

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{i\omega t} dt, \quad -\infty < \omega < \infty, \quad (4.34)$$

называется *одномерным прямым непрерывным преобразованием Фурье (НПФ)* или, более кратко, *преобразованием Фурье (ПФ)*, *Фурье-образом*, *изображением по Фурье*, *спектром* и т.д. А функция $y(t)$ в этом случае называется обратным преобразованием Фурье (ОПФ), или оригиналом. Переменная ω называется *частотой Фурье*. Функция $y(t)$ может быть вещественной или комплексной, а $Y(\omega)$ – вообще говоря, комплексная функция, что вытекает из следующих операций.

Воспользуемся формулой Эйлера:

$$e^{i\varphi} = \cos\varphi + i \sin\varphi. \quad (4.35)$$

Тогда

$$e^{i\omega t} = \cos\omega t + i \sin\omega t. \quad (4.36)$$

В результате выражение (4.34) в случае вещественной $y(t)$ можно записать в виде

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos\omega t dt + i \int_{-\infty}^{\infty} y(t) \sin\omega t dt$$

или

$$Y(\omega) = \operatorname{Re} Y(\omega) + i \operatorname{Im} Y(\omega), \quad (4.37)$$

где

$$\operatorname{Re} Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos\omega t dt, \quad (4.38)$$

$$\operatorname{Im} Y(\omega) = \int_{-\infty}^{\infty} y(t) \sin\omega t dt. \quad (4.39)$$

Соотношение (4.38) называется *косинус-преобразованием Фурье*, а (4.39) – *синус-преобразованием Фурье*. Используют также квадрат модуля ПФ

$$|Y(\omega)|^2 = \operatorname{Re}^2 Y(\omega) + \operatorname{Im}^2 Y(\omega),$$

который называют *спектром мощности*, и модуль ПФ

$$|Y(\omega)| = \sqrt{\operatorname{Re}^2 Y(\omega) + \operatorname{Im}^2 Y(\omega)},$$

который называют спектром интенсивности.

Из (4.34) может быть получено *обратное преобразование Фурье*

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega, \quad -\infty < t < \infty. \quad (4.40)$$

Если задана двухмерная функция (*исходный процесс*) $y(t_1, t_2)$, то над ней может быть выполнено *двухмерное прямое непрерывное преобразование Фурье*

$$Y(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(t_1, t_2) e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2, \quad -\infty < \omega_1, \omega_2 < \infty, \quad (4.41)$$

а формула обращения, аналогичная (4.40), позволяет получить

$$y(t_1, t_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Y(\omega_1, \omega_2) e^{-i(\omega_1 t_1 + \omega_2 t_2)} d\omega_1 d\omega_2, \quad -\infty < t_1, t_2 < \infty, \quad (4.42)$$

– *двухмерное обратное непрерывное преобразование Фурье*.

Непрерывное преобразование Хартли. Приведем также одномерное прямое непрерывное преобразование Хартли

$$Y_H(\omega) = \int_{-\infty}^{\infty} y(t) \text{cas}(\omega t) dt, \quad -\infty < \omega < \infty,$$

где (вещественная) функция cas определяется как

$$\text{cas } x = \cos x + \sin x$$

или

$$\text{cas}(\omega t) = \cos(\omega t) + \sin(\omega t).$$

Обратное непрерывное преобразование Хартли имеет вид

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_H(\omega) \text{cas}(\omega t) d\omega, \quad -\infty < t < \infty.$$

Отличительной особенностью преобразования Хартли (ПХ) является то, что как прямое ПХ $Y_H(\omega)$, так и обратное ПХ $y(t)$ являются вещественными.

Непрерывное преобразование Лапласа. Одностороннее прямое преобразование Лапласа есть интегральное преобразование вида

$$\Psi(p) = \int_0^{\infty} \varphi(x) e^{-px} dx, \quad (4.43)$$

где $p = \lambda + i\sigma$ – комплексная переменная; $\varphi(x)$ – функция действительной переменной x (обычно времени), называемая оригиналом; $\Psi(p)$ – изображение функции $\varphi(x)$, что часто записывается как $\varphi(x) \rightarrow \Psi(p)$ или $\Psi(p) = \mathcal{L}[\varphi(x)]$.

Двустороннее преобразование Лапласа отличается от (4.43) тем, что нижний предел интегрирования полагается равным $-\infty$. Однако под термином «преобразование Лапласа» обычно подразумевается одностороннее преобразование (4.43).

При этом оригинал $\varphi(x)$ должен удовлетворять следующим условиям:

- а) $\varphi(x)$ – кусочно-непрерывная функция,
- б) $\varphi(x) = 0$ при $x < 0$,
- в) $|\varphi(x)| < M e^{cx}$ при $x > 0$, где $M > 0$ и $c \geq 0$ – некоторые константы, причем если $|\varphi(x)| \leq |\varphi(0)|$, то $c = 0$.

Тогда имеет место обратное преобразование Лапласа, дающее выражение для оригинала:

$$\varphi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Psi(p) e^{px} dp. \quad (4.44)$$

На рис. 4.1 пунктиром отображена прямая, по которой идет интегрирование в (4.44), если $\Psi(p)$ не имеет особых точек, иначе путь интегрирования в (4.44) лежит справа от всех особых точек $\Psi(p)$.

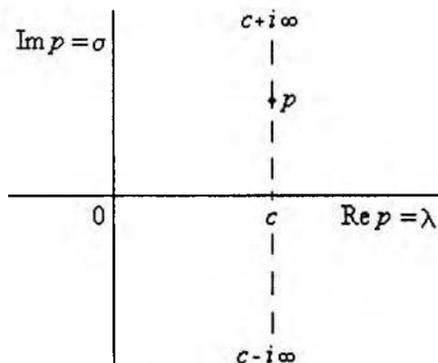


Рис. 4.1

Отметим, что интегральные преобразования Фурье, Хартли, Лапласа (а также Меллина, Гильберта, Ганкеля и др. [4, 27]), строго говоря, нельзя рассматривать как интегральные уравнения относительно оригинала. Например, соотношение (4.34), записанное в виде

$$\int_{-\infty}^{\infty} e^{i\omega t} y(t) dt = Y(\omega), \quad -\infty < \omega < \infty, \quad (4.45)$$

нельзя рассматривать как интегральное уравнение Фредгольма I рода относительно $y(t)$ с ядром $e^{i\omega t}$ и правой частью $Y(\omega)$. Это связано с тем, что в определение интегрального уравнения Фредгольма входит выполнение условия Гильберта–Шмидта (4.29) для ядра. А для ядра $e^{i\omega t}$, входящего в соотношение (4.45), это условие нарушается, так как

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |e^{i\omega t}|^2 d\omega dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega dt = \infty.$$

4.4. Корректность и некорректность по Адамару

Определения корректности и некорректности [4–8, 34]. Ж. Адамар в 1902 г. ввел понятия *корректности* и *некорректности*. Рассмотрим эти определения на примере операторного уравнения I рода:

$$Ay = f, \quad y \in Y, \quad f \in F, \quad (4.46)$$

где y – искомое решение, f – заданная правая часть, Y и F – некоторые гильбертовы пространства (например W_2^1 и L_2), A – заданный непрерывный оператор (линейный, нелинейный, алгебраический, дифференциальный, интегральный и т.д.).

Определение 1. Задача решения уравнения (4.46) называется *корректной* или *корректно поставленной* (*well-posed, correct*) по Адамару, если выполнены следующие три пункта:

- 1) решение существует,
- 2) решение единственно,
- 3) решение устойчиво.

Если хотя бы одно из этих условий (пунктов) не выполняется, то задача называется *некорректной* или *некорректно поставленной* (*ill-posed, incorrect*).

Более строго определение корректности по Адамару формулируется следующим образом [5, с. 9–10].

Определение 2. Задача решения уравнения (4.46) называется *корректной по Адамару*, если:

- 1) для любого $f \in F$ существует элемент $y \in Y$ такой, что $Ay = f$, т.е. область значений оператора равна $R(A) = F$ (*существование решения*),
- 2) элементом f решение y определяется однозначно, т.е. существует обратный оператор A^{-1} (*единственность решения*),
- 3) имеет место непрерывная зависимость y от f , т.е. обратный оператор A^{-1} непрерывен (*устойчивость решения*).

Определение корректности по Адамару называется также *классическим определением корректности*.

Существует еще определение корректности по Фикера [5, с. 31], развивающее определение корректности по Адамару и делающее акцент на свойствах оператора A .

Определение 3. Задача решения уравнения (4.46) называется *корректной по Фикера*, если оператор A нормально разрешим (имеет замкнутую область значений).

В п. 4.7 приведено также определение корректности по Тихонову.

Адамар высказал утверждение, что некорректные задачи не имеют физического смысла, другими словами, если уравнение, описывающее некоторую прикладную задачу, является некорректным, то или эта задача является искусственной (нереальной), или она описана математически неадекватно, в частности, не учтены ограничения на решение, при наличии которых задача станет корректной. Адамар привел при-

меры некорректных задач, в частности, задачу Коши для уравнения Лапласа (см. п. 4.3). Однако данная задача имеет широкое применение в астрономии, в геофизике, в космонавтике и т.д., т.е. имеет реальный физический смысл.

Более того, многие прикладные задачи (обработки сигналов и изображений, томографии, спектроскопии, теории управления и т.д.) являются некорректными, как это выяснилось в последние десятилетия. Поэтому следует заключить, что утверждение Адамара явилось необоснованным и привело к замедлению развития многих разделов как чистой, так и прикладной математики.

Примеры некорректных задач. Приведем примеры уравнений или их систем, задача решения которых некорректна.

Пример 1 [7, с. 179]. Переопределенная система линейных алгебраических уравнений (СЛАУ):

$$\left. \begin{aligned} 2y_1 - 3y_2 &= -4, \\ -y_1 + 2y_2 &= 3, \\ y_1 + 4y_2 &= 15. \end{aligned} \right\} \quad (4.47)$$

Данная СЛАУ действительно является переопределенной и не имеет решения, поскольку ранг расширенной матрицы $\rho = \text{rang}(A|f) = 3$, а ранг матрицы A равен $r = \text{rang}(A) = 2$, т.е. $\rho > r$ (см. п. 4.2). Кроме того, число независимых строк в (4.47) равно 3 и больше n , равного 2.

В том, что данная СЛАУ не имеет решения y_1, y_2 , можно убедиться непосредственно. Действительно, если рассматривать лишь первые два уравнения, то получим решение: $y_1 = 1, y_2 = 2$; если рассматривать второе и третье уравнения, то получим: $y_1 = y_2 = 3$; если же рассматривать первое и третье уравнения, то $y_1 = 2.635, y_2 = 3.09$, т.е. (единое) решение не существует – нарушен первый пункт корректности по Адамару.

Пример 2 [7, с. 179]. Недоопределенная СЛАУ:

$$2y_1 - 3y_2 = -4. \quad (4.48)$$

Для данной СЛАУ $\rho = r = 1 < n = 2$, поэтому она является недоопределенной и имеет множество решений (см. п. 4.2). Действительно, данная СЛАУ имеет не одно решение, например, 1) $y_1 = 1, y_2 = 2$; 2) $y_1 = 2, y_2 = 8/3$; 3) $y_1 = 0, y_2 = 4/3$ и т.д. Таким образом, решение СЛАУ неединственно – нарушен 2-й пункт корректности по Адамару.

Пример 3 [7, с. 179]. Определенная СЛАУ:

$$\left. \begin{aligned} 2y_1 - 3y_2 &= 3, \\ -1.33y_1 + 2y_2 &= -1.99. \end{aligned} \right\} \quad (4.49)$$

Для данной СЛАУ $\rho = r = n = 2$, кроме того, $m = n = 2$ (и $r = m = 2$), поэтому она действительно является определенной.

Решение СЛАУ (4.49) существует и единственно: $y_1 = 3$, $y_2 = 1$. Однако если правые части немного изменить, а именно, решать СЛАУ

$$\left. \begin{aligned} 2y_1 - 3y_2 &= 3.01, \\ -1.33y_1 + 2y_2 &= -2, \end{aligned} \right\} \quad (4.50)$$

т.е. внести относительные погрешности $\|\delta f\| / \|f\| < 0.5\%$, то получим новое, заметно отличное, решение: $y_1 = 2$ (относительная погрешность $|\delta y_1| / |y_1| \approx 33\%$), $y_2 = 0.33$ (относительная погрешность $\approx 67\%$), т.е. относительная погрешность решения на два порядка превосходит относительную погрешность правой части. Это можно оценить и через число обусловленности $\text{cond}(A)$ (см. п. 4.2). Для этого запишем характеристическое уравнение:

$$\begin{vmatrix} 2-\lambda & -3 \\ -1.33 & 2-\lambda \end{vmatrix} = 0, \quad (4.51)$$

откуда $\lambda_1(A) = 3.9975$, $\lambda_2(A) = 0.0025016$, т.е. собственные значения матрицы A вещественны и положительны. Значит, матрица A СЛАУ (4.49) является положительно определенной, но несимметричной. Поэтому число обусловленности $\text{cond}(A)$ следует вычислять с помощью матрицы A^*A , являющейся положительно определенной и симметричной. Имеем, используя (4.9):

$$A^*A = \begin{pmatrix} 5.7689 & -8.66 \\ -8.66 & 13 \end{pmatrix}.$$

Ее собственные значения равны

$$\lambda_1(A^*A) = 18.769, \quad \lambda_2(A^*A) = 0.53996 \cdot 10^{-5}.$$

Сингулярные числа матрицы A равны

$$\mu_1(A) = \sqrt{\lambda_1(A^*A)} = 4.3323, \quad \mu_2(A) = \sqrt{\lambda_2(A^*A)} = 0.0023237.$$

Число обусловленности матрицы A равно

$$\text{cond}(A) = \frac{\mu(A)_{\max}}{\mu(A)_{\min}} = 1.8644 \cdot 10^3 \quad (4.52)$$

и, согласно (4.13) (при $\xi = 0$),

$$\frac{\|\delta y\|}{\|y\|} \leq \text{cond}(A) \frac{\|\delta f\|}{\|f\|}. \quad (4.53)$$

Видим, что решение СЛАУ (4.49) в умеренной степени неустойчиво, а именно, относительная погрешность решения на 2 порядка больше относительной погрешности правой части.

На практике же СЛАУ имеют гораздо большее число обусловленности, т.е. гораздо более неустойчивы. Для таких СЛАУ целесообразно использовать термин «плохо обусловленные СЛАУ». Если же $\text{cond}(A) = \infty$ (в этом случае определитель равен нулю), то такая СЛАУ является вырожденной и погрешность решения будет сколь угодно большой.

Пример 4. Система обыкновенных дифференциальных уравнений (ОДУ). В п. 4.3 приведена задача Коши для системы ОДУ. Запишем ее в виде (ср. (4.17)):

$$\frac{dy(t)}{dt} = f(t, y), \quad (4.54)$$

где $y(t) = \{y_1(t), \dots, y_n(t)\}$ – искомая вектор-функция на отрезке $t_0 \leq t \leq T$, $f(t) = \{f_1(t), \dots, f_n(t)\}$ – правая часть (вектор-функция). *Начальные условия:*

$$y(t_0) = y_0, \quad (4.55)$$

где y_0 – заданный вектор.

В п. 4.3 сформулированы условия существования, единственности и устойчивости решения системы (4.54). А именно, если правые части f_i системы уравнений (4.54) непрерывны и ограничены и удовлетворяют условию Липшица в некоторой замкнутой области, то система (4.54) имеет решение и притом единственное. Далее, если все корни характеристического уравнения имеют отрицательные действительные части, то система уравнений (4.54) устойчива в том смысле, что малые изменения начальных условий не могут вызвать больших и сколь угодно больших изменений решения.

Рассмотрим также (как и в п. 4.3) систему обыкновенных дифференциальных уравнений с управляющей функцией (см. (4.22)):

$$\frac{dy(t)}{dt} = f(t, y, u), \quad (4.56)$$

где $u(t) = \{u_1(t), \dots, u_m(t)\}$ – управляющая вектор-функция. *Начальные условия* – в виде (4.23).

Будем считать функцию $u(t)$, как и $y(t)$, искомой – в этом случае получаем задачу оптимального управления.

Рассмотрим конкретный пример задачи оптимального управления – вертикальное движение ракеты переменной массы с целью достижения максимальной высоты. Это движение описывается

следующей системой ОДУ [8, с. 242–243]:

$$\left. \begin{aligned} \frac{dv(t)}{dt} &= \frac{au(t) - cv^2(t)}{m(t)} - g, \\ \frac{dm(t)}{dt} &= -u(t) \end{aligned} \right\} \quad (4.57)$$

с начальными условиями: $m(0) = m_0$, $v(0) = 0$. Здесь $m(t)$ – переменная масса ракеты, $v(t)$ – скорость подъема ракеты, $u(t)$ – управляющая функция, равная расходу горючего в функции времени, a , c , g – некоторые постоянные.

Ставится задача отыскания оптимальной управляющей функции $u_{\text{opt}}(t)$ – такой функции, при которой высота подъема $H = \max$.

Однако данная задача некорректна [8, с. 244]: малым вариациям H будут соответствовать сколь угодно большие вариации $u(t)$, т.е. нарушается 3-й пункт корректности по Адамару. В п. 4.7 изложен метод регуляризации решения систем ОДУ, делающий задачу устойчивой.

Пример 5. Задача Коши для уравнения Лапласа. В п. 4.3 приведено уравнение Лапласа (дифференциальное уравнение в частных производных эллиптического типа) в двухмерном случае:

$$\Delta u(x, y) \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad (4.58)$$

с граничными условиями типа (задача Коши):

$$u(x, 0) = f(x), \quad \left. \frac{\partial u}{\partial y} \right|_{y=0} = \varphi(x), \quad (4.59)$$

где $f(x)$ и $\varphi(x)$ – заданные функции.

В ряде работ ([6] и др.) показано, что при минимальных требованиях, налагаемых на класс функций, в котором ищется решение $u(x, y)$, оно существует и является единственным, т.е. выполняются 1-й и 2-й пункты корректности по Адамару. Однако решение является неустойчивым. Покажем это.

Если положить $f(x) = f_1(x) = 0$, $\varphi(x) = \varphi_1(x) = a \sin \omega x$, то решением задачи Коши будет функция

$$u(x, y) = u_1(x, y) = (a/\omega) \sin \omega x \operatorname{sh} \omega y,$$

где $\operatorname{sh} x = (e^x - e^{-x})/2$ – гиперболический синус.

А если положить $f(x) = f_2(x) = \varphi(x) = \varphi_2(x) = 0$, то решением задачи Коши будет функция $u(x, y) = u_2(x, y) = 0$. Для любых функциональных пространств и любых $\varepsilon > 0$, $c > 0$, $y > 0$ можно подобрать такие a и ω , чтобы имели место равенство

$$\|f_1 - f_2\| = 0$$

и неравенства

$\|\Phi_1 - \Phi_2\| = \|a \sin \omega x\| < \varepsilon$, $\|u_1 - u_2\| = \|(a/\omega) \sin \omega x \operatorname{sh} \omega y\| > c$, т.е. малым изменениям граничных условий $f(x)$ и $\varphi(x)$ будут соответствовать сколь угодно большие изменения решения $u(x, y)$.

Таким образом, решение задачи Коши для уравнения Лапласа является неустойчивым (нарушается 3-й пункт корректности по Адамару), а значит, эта задача является некорректной.

В п. 4.7 показано, как построить регулярное (устойчивое) решение задачи Коши для уравнения Лапласа.

Пример 6. Задача Коши для уравнения теплопроводности (дифференциального уравнения в частных производных параболического типа). В п. 4.3 приведено четыре варианта данной задачи [5, с. 14–15].

В 1-м варианте (прямой задаче, или задаче с прямым ходом времени) решается уравнение теплопроводности

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2}, \quad 0 \leq x \leq l, \quad t > 0, \quad (4.60)$$

с начальным условием: $u(x, 0) = \varphi(x)$ и граничными условиями: $u(0, t) = \psi_1(t)$, $u(l, t) = \psi_2(t)$.

Искомая функция $u(x, t)$ описывает распределение температуры вдоль стержня в момент времени t при условиях, что в начальный момент ($t = 0$) это распределение равнялось $\varphi(x)$ и концы стержня поддерживаются при температурных режимах $\psi_1(t)$ и $\psi_2(t)$.

Уравнение решается от $t = 0$ в сторону увеличения t . Это – классическая задача теплопроводности. Она является корректной: решение непрерывно зависит от начальных данных $\varphi(x)$.

Во 2-м варианте задачи (в обратной задаче, или задаче с обратным ходом времени) по распределению температуры вдоль стержня в некоторый момент времени $t_* > 0$, т.е. по заданной функции $u(x, t_*)$ необходимо, решая задачу в направлении уменьшения t , восстановить начальное распределение температуры $u(x, 0) = \varphi(x)$. При этом граничные условия $\psi_1(t)$ и $\psi_2(t)$ не заданы. Это – неустойчивая задача.

В 3-м варианте задачи необходимо решить уравнение теплопроводности

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2}, \quad -\infty < x < \infty, \quad 0 \leq t \leq T, \quad (4.61)$$

с начальными условиями: $u(x, 0) = \varphi(x)$ и $u(x, T) = \chi(x)$, где $\varphi(x)$ и

$\chi(x)$ – заданные функции. Такая задача, вообще говоря, не имеет решения (т.е. нарушается 1-й пункт корректности по Адамару). Однако можно, используя лишь начальное условие $u(x, T) = \chi(x)$ и идя от T в прошлое время ($t < T$), найти функцию $\varphi(x) = u(x, 0)$, а не задавать ее. Однако такая задача (аналогичная 2-му варианту) неустойчива (т.е. нарушается 3-й пункт корректности по Адамару).

В 4-м варианте задачи необходимо решить уравнение (4.60) или (4.61), однако лишь с одним начальным условием: $u(x, 0) = \varphi(x)$. Такая задача некорректна и некорректность обусловлена недостаточностью данных.

В п. 4.7 изложены некоторые устойчивые методы решения уравнения теплопроводности.

Корректность или некорректность в зависимости от типов пространств. Задача решения уравнения может быть корректной или некорректной в зависимости от типов функциональных пространств, в которых ищется решение и задаются исходные данные. Кроме того, различные типы уравнений имеют различную «степень некорректности» в одних и тех же пространствах. Поясним это на примерах.

Пример 1. Задача дифференцирования функции $f(x)$, известной приближенно [8, с. 18–19]. Пусть $y(x)$ есть производная функции $f(x)$, т.е. $y(x) = f'(x)$.

Искаженная функция, например, вида $\tilde{f}(x) = f(x) + A \sin \omega x$ в метрике C отличается от $f(x)$ на конечную величину

$$\|\tilde{f} - f\|_C = \max_{a \leq x \leq b} |\tilde{f}(x) - f(x)| = \max_{a \leq x \leq b} |A \sin \omega x| = |A|$$

при любых значениях ω .

Однако производная $\tilde{y}(x) = \tilde{f}'(x) = f'(x) + A \omega \cos \omega x$ отличается от $y(x) = f'(x)$ в метрике C на величину

$$\|\tilde{y} - y\|_C = \max_{a \leq x \leq b} |A \omega \cos \omega x| = |A \omega|,$$

которая может быть произвольно большой при достаточно больших значениях $|\omega|$. Таким образом, задача дифференцирования не обладает, вообще говоря, свойством устойчивости в пространстве C , т.е. является некорректной (нарушается 3-й пункт корректности по Адамару).

Теперь рассмотрим эту задачу на другой паре пространств, а именно, отклонение $\tilde{y}(x)$ от $y(x)$ будем по-прежнему рассматривать в пространстве C , т.е.

$$\|\tilde{y} - y\|_C = \max_{a \leq x \leq b} |\tilde{y}(x) - y(x)|,$$

а уклонение $\tilde{f}(x)$ от $f(x)$ – в пространстве $C^{(1)}$, а именно,

$$\|\tilde{f} - f\|_{C^{(1)}} = \max_{a \leq x \leq b} [|\tilde{f}(x) - f(x)| + |\tilde{f}'(x) - f'(x)|].$$

В этом случае $\|\tilde{y} - y\|_C = |A\omega|$ и

$$\|\tilde{f} - f\|_{C^{(1)}} = \max_{a \leq x \leq b} [|A \sin \omega x| + |A \omega \cos \omega x|] = |A| + |A\omega|.$$

Теперь уже нельзя сказать, что при конечной норме $\|\tilde{f} - f\|_{C^{(1)}}$ норма $\|\tilde{y} - y\|_C$ может быть произвольно большой, а можно говорить, что при конечных $|A|$ и $|\omega|$ уклонение $\|\tilde{y} - y\|_C$ также конечно. Другими словами, задача дифференцирования в паре пространств $(C, C^{(1)})$ (когда $y(x) \in C$, $f(x) \in C^{(1)}$) корректна. И это обусловлено тем, что используя пространство $C^{(1)}$, мы наложили более жесткие условия на функцию $f(x)$, чем при использовании пространства C .

Замечание. В данном примере погрешность функции $f(x)$ имитирована детерминированной функцией $A \sin \omega x$. В практических же задачах погрешность обычно является случайным процессом и требуется более глубокий анализ. Однако качественный вывод о том, что задача дифференцирования является корректной или некорректной в зависимости от пары пространств, сохраняется.

Пример 2. *Задача решения интегральных уравнений Фредгольма I рода*

$$\int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (4.62)$$

и Вольтерра I рода

$$\int_a^x K(x, s) y(s) ds = f(x), \quad a \leq x \leq b. \quad (4.63)$$

Аналогично примеру 1 можно показать [4, 12], что решение уравнения Фредгольма I рода неустойчиво в любой «разумной» паре функциональных пространств: (C, L_2) , (L_2, L_2) , (C, C) , $(C, C^{(1)})$, (W_2^1, L_2) и т.д., т.е. нарушается 3-й пункт (а также 1-й и 2-й пункты) корректности по Адамару.

Что касается уравнения Вольтерра I рода, то оно корректно на паре пространств $(C, C^{(1)})$ (когда $y(s) \in C$, $f(x) \in C^{(1)}$) и некорректно на паре (C, C) (когда $y(s) \in C$, $f(x) \in C$) [12]. В этом смысле уравнение Вольтерра I рода занимает промежуточное положение, например, между уравнением Фредгольма I рода и уравнением Вольтерра II рода

$$y(x) + \int_a^x K(x,s) y(s) ds = f(x), \quad a \leq x \leq b,$$

задача решения которого корректна при выполнении весьма нежестких требований к ядру $K(x,s)$ и правой части $f(x)$ [4, с. 20].

Замечание. Можно заметить, что условия корректности или некорректности для задачи дифференцирования и задачи решения уравнения Вольтерра I рода одинаковы. Это связано с тем, что задача определения производной n -го порядка $y(x)$ от функции $f(x)$ может быть сведена к решению следующего интегрального уравнения Вольтерра I рода [8, с. 19]

$$\int_a^x \frac{(x-s)^{n-1}}{(n-1)!} y(s) ds = f(x), \quad x \geq a,$$

причем $f(a) = f'(a) = \dots = f^{(n-1)}(a) = 0$.

Пример 3. Это – пример из задачи реконструкции смазанных изображений (см. п. 5.1). Исходное уравнение задачи – *интегральное уравнение Вольтерра I рода*

$$\int_x^{x+\Delta} \frac{1}{\Delta} w(\xi) d\xi = g(x), \quad -\infty < x < \infty, \quad (4.64)$$

где $g(x) \equiv g_y(x)$ – распределение интенсивности на смазанном изображении вдоль равномерного и прямолинейного смаза (сдвига), $w(\xi) \equiv w_y(\xi)$ – искомое распределение интенсивности на истинном (несмазанном) изображении, Δ – величина смаза, ось x направлена вдоль смаза, а ось y – перпендикулярно ему. В уравнении (4.64) y играет роль параметра (при каждом значении y нужно решать одномерное уравнение (4.64)).

По терминологии [12], это уравнение является *неклассическим уравнением Вольтерра I рода* (так как оба предела интегрирования переменны). Однако, как показано в [12] (см. также [4, с. 111]), задача решения неклассического уравнения, как и уравнения (4.63), корректна на паре пространств $(C, C^{(1)})$, т.е. когда $w(\xi) \in C$, $g(x) \in C^{(1)}$, и некорректна на паре (C, C) , когда $w(\xi) \in C$, $g(x) \in C$.

Далее, уравнение (4.64) может быть преобразовано в другой тип интегрального уравнения, а именно, в *интегральное уравнение Фредгольма I рода* (типа свертки):

$$\int_{-\infty}^{\infty} k(x-\xi) w(\xi) d\xi = g(x), \quad -\infty < x < \infty, \quad (4.65)$$

где

$$k(x) = \begin{cases} 1/\Delta, & -\Delta \leq x \leq 0, \\ 0, & \text{иначе.} \end{cases} \quad (4.66)$$

Уравнение Фредгольма I рода (4.65) некорректно в любой паре пространств, в том числе, в $(C, C^{(1)})$. При этом отметим, что переход от уравнения Вольтерра (4.64) к уравнению Фредгольма (4.65) выполнен лишь за счет эквивалентных преобразований (см. п. 2.2). Таким образом, данный пример подтверждает вывод, сделанный в гл. 1–2 о том, что эквивалентные преобразования могут переводить корректное уравнение в некорректное и наоборот. Однако при решении как уравнения (4.64), так и уравнения (4.65) методом регуляризации (см. п. 4.7) оба уравнения дают устойчивое решение, т.е. применение регуляризации «спасает» от возможных нежелательных эффектов, которые могут иметь место при эквивалентных преобразованиях [23, 36].

З а м е ч а н и е. Может быть высказано мнение, что можно подобрать подходящую пару функциональных пространств для решения того или иного уравнения и задача его решения станет корректной [6, с. 8]. Например, интегральное уравнение Вольтерра I рода решать не на паре пространств (C, C) , где задача его решения некорректна, а на паре $(C, C^{(1)})$, где задача корректна. Это же относится и к задаче дифференцирования, как показано выше. Аналогично можно подобрать соответствующие пары пространств для решения дифференциальных уравнений в частных производных и задача их решения станет корректной. Однако, во-первых, есть уравнения, задача решения которых некорректна в любой паре пространств (это – интегральные уравнения Фредгольма I рода). А во-вторых, выбор пары пространств не может быть произвольным, а должен согласовываться с особенностью практического измерения правой части $f(x)$ с учетом погрешности эксперимента $\delta f(x)$. Если полагать, что $f(x) \in C$, то это означает, что мы будем пользоваться оценкой максимальной погрешности измерений

$$\delta \equiv \|\delta f(x)\|_C = \max_{c \leq x \leq d} |\delta f(x)|,$$

а такая погрешность вполне может быть оценена из эксперимента. Аналогично может быть оценена погрешность δ в случае, когда $f(x) \in L_2$, – это будет среднеквадратическая погрешность

$$\delta \equiv \|\delta f(x)\|_{L_2} = \left(\int_c^d |f(x)|^2 dx \right)^{1/2}.$$

Однако практически невозможно оценить погрешность δ в случае,

когда $f(x) \in C^{(1)}$, – такая оценка будет равна

$$\delta \equiv \|\delta f(x)\|_{C^{(1)}} = \max_{c \leq x \leq d} \{|\delta f(x)| + |\delta f'(x)|\}.$$

Поэтому парой пространств $(C, C^{(1)})$ мы не сможем практически воспользоваться, а сможем использовать менее желательную пару (C, C) или (C, L_2) . Заметим, что значение δ используется в методах регуляризации решения некорректных задач (см. п. 4.7).

Далее будет показано, что если решение не существует (нарушен 1-й пункт корректности по Адамару), то нужно использовать метод типа метода наименьших квадратов Гаусса (будет получено псевдорешение – см. п. 4.6); если решение неединственно (нарушен 2-й пункт корректности по Адамару), то нужно использовать метод типа метода псевдообратной матрицы Мура–Пенроуза (будет получено нормальное решение – см. п. 4.6); если же решение неустойчиво (нарушен 3-й пункт корректности по Адамару), то нужно использовать устойчивые (регулярные) методы типа метода регуляризации Тихонова (будет получено регуляризованное решение – см. п. 4.7).

4.5. Классические методы решения интегральных уравнений Фредгольма I рода

Прежде чем излагать устойчивые (регулярные) методы решения некорректных задач, остановимся на некоторых классических методах, причем для большей убедительности сделаем это применительно к интегральному уравнению Фредгольма I рода как наиболее неустойчивому уравнению.

Метод квадратур [7, с. 180–182]. Рассмотрим *интегральное уравнение Фредгольма I рода*

$$\int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (4.67)$$

где $K(x, s)$ – ядро, $y(s)$ – искомая функция, $f(x)$ – правая часть, $[a, b]$ – область изменения s , $[c, d]$ – область изменения x .

Полагаем, что вместо $f(x)$ известна $\tilde{f}(x)$ – измеренная с погрешностями правая часть, вместо $K(x, s)$ известно $\tilde{K}(x, s)$ – неточное ядро, тогда под $y(s)$ нужно подразумевать $\tilde{y}(s)$ – решение с погрешностями, обусловленными погрешностями правой части и ядра (а также метода решения). Поэтому вместо (4.67) нужно писать:

$$\int_a^b \tilde{K}(x,s) \tilde{y}(s) ds = \tilde{f}(x), \quad c \leq x \leq d, \quad (4.68)$$

однако для упрощения записей мы будем далее использовать уравнение (4.67), подразумевая под ним уравнение (4.68).

Метод квадратур заключается в следующем [7, с. 180–182].

1) Вводим равномерные сетки узлов по s с шагом $\Delta s = h_1 = \text{const}$ и по x с шагом $\Delta x = h_2 = \text{const}$. Получим число узлов равным $n = (b-a)/h_1 + 1$ по s и $m = (d-c)/h_2 + 1$ по x .

2) Интеграл в (4.67) заменяем конечной суммой, расписывая его по некоторой квадратурной формуле, например, по формуле трапеций:

$$\int_a^b K(x,s) y(s) ds \approx \sum_{j=1}^n p_j K(x,s_j) y(s_j),$$

где

$$p_j = \begin{cases} 0.5h_1, & j=1 \text{ или } j=n, \\ h_1, & \text{иначе,} \\ s_j = a + (j-1)h_1. \end{cases}$$

3) Учитывая дискретизацию по x :

$$x_i = c + (i-1)h_2,$$

окончательно получим дискретный аналог уравнения (4.67):

$$\sum_{j=1}^n A_{ij} y_j = f_i, \quad i=1, \dots, m, \quad (4.69)$$

где $A_{ij} = p_j K(x_i, s_j)$ – элементы матрицы A размера $m \times n$; $y_j = y(s_j)$; $f_i = f(x_i)$.

Итак, получена система m линейных алгебраических уравнений (4.69) относительно n неизвестных y_j . Решая ее, можно получить решение интегрального уравнения (4.67) в дискретном численном виде.

Матрица A системы уравнений (4.69), вообще говоря, прямоугольна. Если $m = n$, то матрица A – квадратная и СЛАУ (4.69) можно решать по правилу Крамера, гауссовскими методами и т.п. Если $m > n$, то СЛАУ (4.69) нужно решать методом наименьших квадратов Гаусса (см. п. 4.6) – получим псевдорешение, а если $m < n$, то нужно использовать метод псевдообратной матрицы Мура–Пенроуза (см. п. 4.6) – получим нормальное решение. Таким образом, первые два пункта корректности по Адамару формально будут выполнены.

Однако все эти решения очень неустойчивы, т.е. нарушается 3-й

пункт корректности по Адамару. Эта неустойчивость обусловлена тем, что обратный оператор уравнения (4.67) неограничен. Это проявляется в том, что минимальное сингулярное число μ_{\min} интегрального оператора уравнения (4.67) равно нулю и число обусловленности $\text{cond} = \infty$. Если же интегральный оператор аппроксимируется алгебраическим оператором с конечными m и n , то μ_{\min} может стать несколько отличным от нуля, однако решение СЛАУ (4.69) будет по-прежнему очень неустойчивым.

На рис. 4.2 приведены результаты решения примера (приведенного далее в п. 4.7) методом квадратур согласно (4.69) при $m = n = 137$. Точное решение $y(s)$ является суммой пяти гауссиан. Пунктирными почти вертикальными линиями отображено приближенное (дискретное) решение y_j , $j = 1, \dots, n$, СЛАУ (4.69).

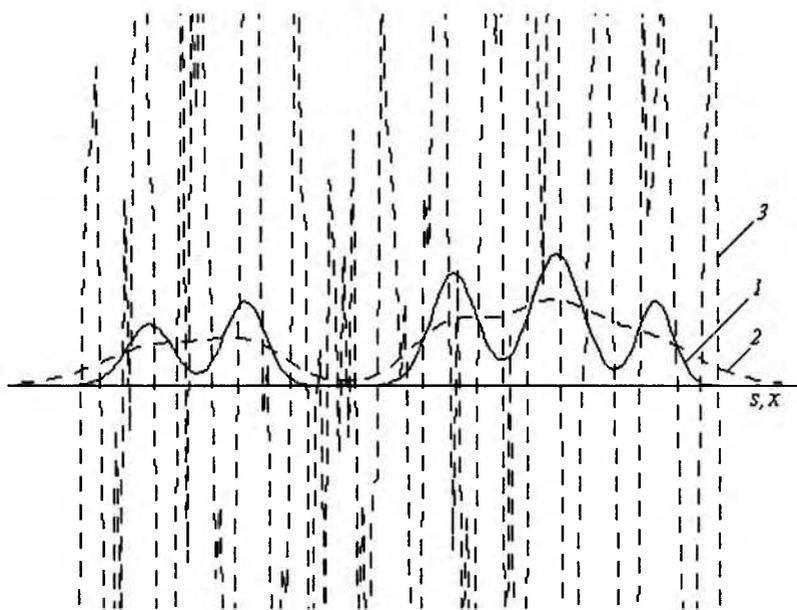


Рис. 4.2. Неустойчивое решение интегрального уравнения Фредгольма I рода методом квадратур
 1 — точное решение $y(s)$ 2 — правая часть $f(x)$
 3 — приближенное решение y_j

Видим, что решение y_j получилось в виде так называемой переменной «пилы» большой амплитуды (ср. [8, с. 162], [22, с. 91]),

ничего общего не имеющей с точным решением. Между тем, если для проверки подставить «пилу» в (4.69), то получим совпадение левой и правой частей (4.69) с точностью до 3–5 цифр в случае, если вычисления выполнять на компьютере с простой точностью (до 7 цифр), или до 6–10 цифр, если вычисления выполнять с двойной точностью (до 14 цифр). Заметим, что вид «пилы» зависит от метода решения СЛАУ, от компьютерной программы и т.д.

Из изложенного можно сделать следующие выводы:

1) Метод квадратур решения интегрального уравнения Фредгольма I рода является крайне неустойчивым (т.е. нарушается 3-й пункт корректности по Адамару). Применительно к интегральному уравнению Вольтерра I рода метод квадратур гораздо более устойчив (см. [4, с. 111–114, 120–123]).

2) Классическое определение точного решения \bar{y} как решения, при котором

$$\|A\bar{y} - f\| = 0, \quad (4.70)$$

т.е. невязка между левой и правой частями уравнения равна нулю, вообще говоря, для некорректных задач не подходит. Это связано с тем, что в случае, когда решение не существует (см. пример 1 в п. 4.4), отсутствует такое \bar{y} , для которого выполняется (4.70), а в случае неединственности решения (см. пример 2 в п. 4.4) существует множество \bar{y} , для которых выполняется (4.70), и наконец, в случае неустойчивости (см. пример 3 в п. 4.4 и рис. 4.2) критерий (4.70) дает неустойчивое решение.

Метод преобразования Фурье для уравнений типа свертки [4, с. 256–259], [7, с. 182–184]. Рассмотрим *интегральное уравнение Фредгольма I рода типа свертки*

$$\int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (4.71)$$

Такое уравнение имеет аналитическое решение и оно записывается в виде обратного преобразования Фурье (ОПФ):

$$y(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega)e^{-i\omega s} d\omega, \quad (4.72)$$

где спектр (или ПФ) решения равен

$$Y(\omega) = \frac{F(\omega)}{\lambda(\omega)}, \quad (4.73)$$

причем

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx, \quad (4.74)$$

$$\lambda(\omega) = \int_{-\infty}^{\infty} K(x) e^{i\omega x} dx \quad (4.75)$$

– спектры (или ПФ) правой части $f(x)$ и ядра $K(x)$.

Однако решение (4.72), как и решение методом квадратур, также крайне неустойчиво и это можно объяснить следующим образом. Ядро $K(x)$ обычно задается в виде гладкой аналитической формулы, поэтому ее спектр $\lambda(\omega)$ быстро распадается с ростом $|\omega|$ и $\lim_{\omega \rightarrow \infty} \lambda(\omega) = 0$.

Функция же $f(x)$ обычно задается в виде таблицы экспериментальных, а значит, зашумленных значений, т.е. вместо точной $f(x)$ мы имеем $\tilde{f}(x) = f(x) + \delta f(x)$, где $\delta f(x)$ – погрешность, обладающая тем свойством, что ее спектр при $|\omega| \rightarrow \infty$ обычно стремится к некоторой (пусть даже очень малой) константе – уровню «белого шума». Поэтому $\lim_{\omega \rightarrow \infty} F(\omega) / \lambda(\omega) = \lim_{\omega \rightarrow \infty} Y(\omega) = \infty$ и интеграл (4.72) расходится. Другими словами, неустойчивость метода ПФ обусловлена очень сильной реакцией высоких гармоник Фурье даже на сколь угодно малые погрешности измерений правой части $f(x)$.

Если же вычисления выполняются по конечным квадратурным формулам, то вместо непрерывных преобразований Фурье (НПФ) $F(\omega)$, $\lambda(\omega)$ и $Y(\omega)$ используются их дискретные преобразования Фурье (ДПФ) [7, 15, 17, 38]. В этом случае максимальная частота Фурье конечна: $\omega_{\max} = 2\pi/h$, где $h = \Delta x = \Delta s$ – шаг дискретизации. В результате неустойчивость решения снижается, хотя и остается большой.

Как показывает решение ряда примеров, метод ПФ дает более устойчивое решение, чем метод квадратур, и это обусловлено тем, что, во-первых, метод ПФ дает аналитическое решение (4.72) и во-вторых, при численной реализации метода ПФ автоматически отсекаются высокие частоты Фурье за счет конечности шага дискретизации h .

4.6. Методы наименьших квадратов Гаусса и псевдообратной матрицы Мура–Пенроуза

Изложим на примере решения системы линейных алгебраических уравнений (СЛАУ) и интегрального уравнения метод наименьших квадратов (МНК) Гаусса [3, с. 521], [7, с. 186–188], а также метод псевдообратной матрицы Мура–Пенроуза [4, с. 508], [7, с. 189], [18, 21].

Переопределенная СЛАУ. Рассмотрим систему m линейных алгебраических уравнений (СЛАУ) относительно n неизвестных, причем $m > n$ и $\text{rang}(A|f) > \text{rang}(A)$ или, что равнозначно, число линейно независимых строк больше n , т.е. рассмотрим переопределенную СЛАУ (например, (4.47)):

$$Ay = f, \quad (4.76)$$

где A – матрица $m \times n$, y – искомый вектор-столбец $n \times 1$, f – заданная правая часть – вектор-столбец $m \times 1$. Такая СЛАУ не имеет решения, т.е. не существует такого \bar{y} , для которого невязка равна нулю:

$$\|A\bar{y} - f\| = 0. \quad (4.77)$$

Псевдорешение и нормальная СЛАУ. В методе наименьших квадратов (МНК) Гаусса вместо (4.77) вводится условие

$$\|Ay - f\| = \min_y. \quad (4.78)$$

Определение. *Псевдорешением* СЛАУ (4.76) называется решение y , удовлетворяющее условию (4.78), т.е. минимизирующее невязку $\|Ay - f\|$.

Таким образом, в МНК условие равенства нулю невязки заменяется на условие ее минимума, а вместо точного решения \bar{y} рассматривается псевдорешение y . Обычно норма $\|Ay - f\|$ берется в пространстве R^n (см. (4.8)), поэтому метод называется *методом наименьших квадратов*.

Замечание. Если $\|Ay - f\| = 0$, то псевдорешение y совпадает с точным решением \bar{y} , т.е. псевдорешение обобщает понятие точного решения.

Запишем условие (4.78) в виде:

$$\|Ay - f\|^2 = \min_y. \quad (4.79)$$

Выведем новую СЛАУ из условия (4.79). Минимизация (4.79) означает равенство нулю вариации или производной по y (называемой производной Фреше [10, с. 85]): $2A^*(Ay - f) = 0$ или $A^*Ay - A^*f = 0$. В результате

$$A^*Ay = A^*f. \quad (4.80)$$

Итак, вместо переопределенной СЛАУ (4.76) получена новая СЛАУ (4.80), называемая *нормальной СЛАУ*.

Запишем (4.80) в виде:

$$By = u, \quad (4.81)$$

где

$$B = A^*A, \quad (4.82)$$

$$u = A^* f \quad (4.83)$$

или в случае вещественности A

$$B = A^T A, \quad (4.84)$$

$$u = A^T f. \quad (4.85)$$

Используя правило (4.9) умножения матриц и правило (4.10) умножения матрицы на вектор, запишем подробно выражения для элементов новой матрицы B и новой правой части u нормальной СЛАУ (4.81) в случае вещественности A :

$$B_{ij} = \sum_{k=1}^m A_{ik}^T A_{kj} = \sum_{k=1}^m A_{ki} A_{kj}, \quad (4.86)$$

$$u_i = \sum_{k=1}^m A_{ik}^T f_k = \sum_{k=1}^m A_{ki} f_k. \quad (4.87)$$

Решение СЛАУ (4.80) или (4.81) формально имеет вид:

$$y = (A^* A)^{-1} A^* f \quad (4.88)$$

или

$$y = B^{-1} u, \quad (4.89)$$

а практически может быть найдено методами Крамера, Гаусса и др.

МНК применительно к интегральному уравнению. Если применить МНК Гаусса к интегральному уравнению Фредгольма I рода (4.67), то получим новое интегральное уравнение (ср. (4.81)):

$$\int_a^b B(t,s) y(s) ds = u(t), \quad a \leq t \leq b, \quad (4.90)$$

где (ср. (4.86), (4.87))

$$B(t,s) = B(s,t) = \int_c^d K(x,t) K(x,s) dx, \quad (4.91)$$

$$u(t) = \int_c^d K(x,t) f(x) dx. \quad (4.92)$$

Сформулируем основные особенности МНК Гаусса.

1) Матрица B новой (нормальной) СЛАУ (4.81) является квадратной $n \times n$, т.е. решается новая система n уравнений относительно n неизвестных и в случае $\det(B) \neq 0$ решение СЛАУ (4.81) существует и является единственным.

2) Матрица B новой СЛАУ (4.81), а также ядро $B(t,s)$ нового интегрального уравнения (4.90) являются симметричными положительно определенными. Поэтому решать СЛАУ (4.81) целесообразно посред-

ством специальных методов (Холецкого, Краута и др.), а не общих методов (Крамера, Гаусса и т.п.). Это же относится и к решению СЛАУ, получающейся при решении нового интегрального уравнения (4.90) методом квадратур.

3) Если решения изначальных СЛАУ (4.76) и интегрального уравнения (4.67) неустойчивы, то решения новых СЛАУ (4.81) и интегрального уравнения (4.90) также неустойчивы, т.е. МНК проблему неустойчивости не решает.

Пример. Для иллюстрации МНК рассмотрим пример 1 из п. 4.4 – см. (4.47). Матрица СЛАУ (4.47) равна

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \\ 1 & 4 \end{pmatrix}. \quad (4.93)$$

Используя формулы (4.86), (4.87), получим

$$B = \begin{pmatrix} 6 & -4 \\ -4 & 29 \end{pmatrix}, \quad (4.94)$$

$$u = \begin{pmatrix} 4 \\ 78 \end{pmatrix}, \quad (4.95)$$

т.е. приходим к новой (нормальной) СЛАУ

$$\begin{cases} 6y_1 - 4y_2 = 4, \\ -4y_1 + 29y_2 = 78. \end{cases} \quad (4.96)$$

Ее решение: $y_1 = 2.71$, $y_2 = 3.06$. Этому соответствует невязка (минимально возможная) $\|Ay - f\| = 0.3993 \approx 0.4$ (согласно (4.8)). Пусть решение характеристического уравнения для матрицы $B = A^*A$

$$\begin{vmatrix} 6 - \lambda & -4 \\ -4 & 29 - \lambda \end{vmatrix} = 0$$

найдем собственные значения матрицы B : $\lambda_1(B) = 29.675$, $\lambda_2(B) = 5.325$, т.е. λ_1 и λ_2 вещественны и неотрицательны, как и должно быть для симметричной и положительно определенной матрицы $B = A^*A$.

Теперь изложим метод псевдообратной матрицы (МПОМ) Мура-Пенроуза, причем, как и вышеприведенный метод МНК Гаусса, сначала рассмотрим МПОМ на примере решения СЛАУ [4, 508], [7, с. 189–191], [18, 21].

Недоопределенная СЛАУ. Рассмотрим СЛАУ

$$Ay = f, \quad (4.97)$$

где A – $m \times n$ -матрица, y – искомый n -вектор, f – заданная правая часть – m -вектор, причем $m < n$ или $\rho < n$, где $\rho = \text{rang}(A|f)$ – ранг расширенной матрицы. Такая СЛАУ является *недоопределенной*.

Недоопределенная СЛАУ имеет множество решений \bar{y} , т.е. нарушается 2-й пункт корректности по Адамару. Например, СЛАУ (4.48) имеет следующие решения: 1) $\bar{y}_1 = \{1, 2\}^T$, 2) $\bar{y}_2 = \{2, 8/3\}^T$, 3) $\bar{y}_3 = \{0, 4/3\}^T$ и т.д. Для каждого из них выполняется равенство (4.77).

Нормальное решение и псевдообратная матрица. Выбор единственного решения из множества решений недоопределенной матрицы может быть осуществлен *методом псевдообратной матрицы* (МПОМ) Мура–Пенроуза (1930 г.).

Определение. *Нормальное решение* – это решение с минимальной нормой среди множества решений, т.е. удовлетворяющее условию

$$\|y\| = \min_y \quad (4.98)$$

или

$$\|y\|^2 = \min_y. \quad (4.99)$$

Нормальное решение – это наиболее гладкое из решений.

Согласно МПОМ, среди множества решений недоопределенной СЛАУ выбирается нормальное решение. Доказано [21], что нормальное решение существует и является единственным и находится по формуле:

$$y = A^+ f, \quad (4.100)$$

где A^+ – *псевдообратная* $n \times m$ -матрица Мура–Пенроуза. Матрица A^+ определяется соотношением:

$$AA^+A = A \quad (4.101)$$

или (асимптотическая формула)

$$A^+ = \lim_{\alpha \rightarrow 0} (\alpha E + A^*A)^{-1} A^*. \quad (4.102)$$

Существует одна и только одна псевдообратная матрица A^+ , определяемая соотношением (4.101) или (4.102).

Однако соотношения (4.101) и (4.102) неудобны для практического нахождения A^+ . Практически A^+ может быть найдена по формуле [4, с. 508], [21]:

$$A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*, \quad (4.103)$$

где $A = BC$ – *скелетное* (неоднозначное) *разложение матрицы* A , причем B – некоторая невырожденная $m \times r$ -матрица, C – $r \times n$ -

матрица, равная $C = B^{-1}A$, $r = r_A = r_B = r_C$, где r_A, r_B, r_C – ранги матриц A, B, C . При различных скелетных разложениях получается одна и та же псевдообратная матрица A^+ .

В случае квадратной невырожденной матрицы A имеем: $A^+ = A^{-1}$, а в случае переопределенной системы $A^+ = (A^* A)^{-1} A^*$ (ср. (4.88)), т.е. запись (4.100) является общей для переопределенной, определенной и недоопределенной СЛАУ. Кроме того, решение (4.100), которое уместно записать в виде $y^+ = A^+ f$ (где y^+ – нормальное решение), дает нулевую невязку $\|Ay^+ - f\| = 0$, т.е. оно является *псевдорешением* (ср. (4.78)). При этом среди всех псевдорешений (которых в случае недоопределенной СЛАУ множество) решение y^+ имеет, как нормальное решение, наименьшую норму. Другими словами, нормальное решение является одновременно и псевдорешением. Поэтому нормальное решение называют также *нормальным псевдорешением* (решением, наиболее гладким среди возможного множества псевдорешений). Нормальное псевдорешение существует и является единственным.

Пример. В качестве примера рассмотрим недоопределенную СЛАУ (4.48). Мы уже указали (см. п. 4.4) несколько ее решений: 1) $y_1 = 1, y_2 = 2$, норма решения (в соответствии с (4.8)) равна $\|y\| = \sqrt{1^2 + 2^2} = 2.24$; 2) $y_1 = 2, y_2 = 8/3$, $\|y\| = 3.34$; 3) $y_1 = 0, y_2 = 4/3$, $\|y\| = 1.33$.

Найдем нормальное решение СЛАУ (4.48), используя формулы (4.100) и (4.103). В данном примере $m = 1, n = 2$, матрица $A = \begin{pmatrix} 2 & -3 \end{pmatrix}$, ее ранг $r_A = 1$. Воспользуемся скелетным разложением: $A = BC$, где B – 1×1 -матрица (т.е. скаляр), C – 1×2 -матрица (т.е. вектор-строка), причем $r = r_A = r_B = r_C = 1$. Положим: $B = b$, где $b \neq 0$ – некоторое число. Тогда $C = B^{-1}A = (2/b \quad -3/b)$,

$$C^T = \begin{pmatrix} 2/b \\ -3/b \end{pmatrix}, \quad CC^T = (13/b^2), \quad (CC^T)^{-1} = (b^2/13),$$

$$C^T(CC^T)^{-1} = \begin{pmatrix} 2b/13 \\ -3b/13 \end{pmatrix}, \quad (B^T B)^{-1} B^T = (1/b).$$

Используя (4.103), получим:

$$A^+ = C^T(CC^T)^{-1}(B^T B)^{-1} B^T = \begin{pmatrix} 2/13 \\ -3/13 \end{pmatrix}$$

при любом b . Поскольку правая часть равна $f = (-4)$, то нормальное

решение согласно (4.100) равно

$$y = A^+ f = \begin{pmatrix} -8/13 \\ 12/13 \end{pmatrix}.$$

Итак, нормальное решение СЛАУ (4.48) равно: $y_1 = -8/13 \approx -0.615$, $y_2 = 12/13 \approx 0.923$, а его норма $\|y\| \approx 1.113 = \min$. Проверка решения: $y_2 = (2y_1 + 4)/3 = 12/13$.

Далее имеем:

$$A = \begin{pmatrix} 2 & -3 \end{pmatrix}, \quad A^* = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \quad A^*A = \begin{pmatrix} 4 & -6 \\ -6 & 9 \end{pmatrix}.$$

Характеристическое уравнение

$$\begin{vmatrix} 4 - \lambda & -6 \\ -6 & 9 - \lambda \end{vmatrix} = 0$$

дает корни: $\lambda_1(A^*A) = \lambda(A^*A)_{\max} = 13$, $\lambda_2(A^*A) = \lambda(A^*A)_{\min} = 0$. Следовательно, сингулярные числа равны

$$\mu(A)_{\max} = \sqrt{13}, \quad \mu(A)_{\min} = 0,$$

число обусловленности и определитель равны

$$\text{cond}(A) = \mu(A)_{\max} / \mu(A)_{\min} = \infty, \quad \det(A^*A) = 0,$$

т.е. матрица A^*A является вырожденной и обратная матрица $(A^*A)^{-1}$ не существует; ее норма $\|(A^*A)^{-1}\| = 1/\mu(A)_{\min} = \infty$. Другими словами, СЛАУ (4.48) является неустойчивой (нарушен не только 2-й, но и 3-й пункт корректности по Адамару).

МПОМ применительно к другим уравнениям. Если под запись (4.97) подразумевать не только СЛАУ, но и другие уравнения, а именно, интегральное уравнение Фредгольма I рода (4.67), дифференциальные уравнения (4.54), (4.56), (4.58), (4.60), операторное уравнение (4.46) и др., то изложенный метод псевдообратной матрицы также может быть применен. В этом случае в качестве решения интегрального, дифференциального, операторного и др. уравнения выбирается *нормальное решение* y , удовлетворяющее условию (4.98) и находимое по формуле (4.100), где A^+ – *псевдообратный оператор*, определяемый посредством (4.102) или (4.103). Подробнее – в п. 4.7.

Общий вывод. По материалам данного пункта можно сделать следующий вывод: если в качестве решения брать так называемое *нормальное псевдорешение* (псевдорешение, наиболее гладкое среди возможного множества псевдорешений), то будут выполнены первые два пункта корректности по Адамару (существование и единственность решения).

Однако МНК и МПОМ не решают проблему неустойчивости решения, т.е. 3-й пункт корректности по Адамару для нормального псевдорешения, вообще говоря, не выполняется. Как решить вопрос об устойчивости, изложено в п. 4.7 и 4.8.

4.7. Метод регуляризации Тихонова

В данном пункте изложен один из наиболее эффективных методов решения некорректных задач – метод регуляризации Тихонова [1, 4–8, 10, 16, 19, 29, 34, 44] и продемонстрировано его применение к решению различных уравнений. Метод регуляризации Тихонова является дальнейшим развитием метода наименьших квадратов (МНК) Гаусса, дающего псевдорешение, и метода псевдообратной матрицы (МПОМ) Мура–Пенроуза, дающего нормальное решение.

Существо метода. Сначала рассмотрим метод регуляризации Тихонова применительно к операторному уравнению I рода:

$$Ay = f, \quad y \in Y, \quad f \in F, \quad (4.104)$$

где Y и F – некоторые метрические пространства, $A: Y \rightarrow F$ – заданный линейный оператор, действующий из Y в F , $f \in F$ – заданный элемент (правая часть), $y \in Y$ – искомый элемент (решение). Пусть вместо точных f и A известны их приближения \tilde{f} и \tilde{A} такие, что

$$\begin{aligned} \|\tilde{f} - f\| &\leq \delta, \\ \|\tilde{A} - A\| &\leq \xi, \end{aligned}$$

где $\delta > 0$ и $\xi \geq 0$ – погрешности правой части и оператора (точнее, их верхние оценки). Таким образом, вместо (4.104) решается уравнение

$$\tilde{A}\tilde{y} = \tilde{f}, \quad \tilde{y} \in Y, \quad \tilde{f} \in F. \quad (4.105)$$

Замечание. Если под операторным уравнением (4.104) подразумевать интегральное уравнение, то погрешность δ правой части f обусловлена погрешностями измерений f , а погрешность ξ оператора A – неточностью знания ядра $K(x, s)$ и погрешностью численного метода решения, например, метода квадратур. Если же иметь в виду СЛАУ, то δ – это погрешность вектора правой части, а ξ – это погрешность элементов матрицы A . Если же иметь в виду однородные дифференциальные уравнения (Лапласа, теплопроводности и т.д.), то $\delta = 0$, а ξ – это погрешность краевых условий + погрешность численного метода решения, например, метода сеток.

Для упрощения записи мы будем далее использовать запись (4.104), имея в виду, что в действительности рассматривается уравнение (4.105).

В п. 4.4 приведено классическое (по Адамару) определение корректности. А. Н. Тихонов сформулировал новое определение корректности, которое М. М. Лаврентьев назвал *корректностью по Тихонову*.

Определение. Задача решения уравнения (4.104) называется *условно корректной*, или *корректной по Тихонову*, если [5, с. 18], [6, с. 30], [8, с. 16]:

1) априори известно, что решение y существует и принадлежит некоторому заданному множеству (множеству корректности) $M \subset Y$: $y \in M$,

2) решение единственно в классе функций, принадлежащих M , т.е. оператор A обратим на множестве M ,

3) существует непрерывная зависимость решения y от правой части f , когда бесконечно малым вариациям f , не выводящим решение y за пределы M , соответствуют бесконечно малые вариации решения y , т.е. обратный оператор A^{-1} является непрерывным.

Отличие условной корректности (по Тихонову) от классической (по Адамару) заключается во введении *множества корректности*, существенно сужающего класс возможных решений.

Наиболее характерный пример множества корректности – компакт [10, с. 70, 72] (его примеры: множество функций монотонных ограниченных снизу и сверху, множество ограниченных выпуклых функций и т.д.). Однако в большинстве прикладных задач класс возможных решений Y не является компактом (т.е. нельзя наложить, исходя из физических соображений, жесткие ограничения на решение) и, кроме того, изменения правой части f уравнения (4.104), обусловленные ее погрешностями, могут выводить f за пределы множества $A Y$ – образа множества Y при отображении его с помощью оператора A . Такие задачи называются *существенно некорректными* [8, с. 53].

А. Н. Тихонов разработал [8] принципиально новый подход, дающий устойчивые решения существенно некорректных задач. В его основе лежит понятие *регуляризирующего оператора* (РО) или *регуляризирующего алгоритма* (РА).

В методе регуляризации Тихонова ставятся два условия: условие минимизации невязки типа (4.79), как в МНК Гаусса, и условие минимизации нормы решения типа (4.99), как в МПОМ Мура–Пенроуза. Это – задача условной минимизации и она решается обычно *методом неопределенных множителей Лагранжа*, а именно, вводится условие минимизации *сглаживающего функционала*:

$$\|Ay - f\|^2 + \alpha \|y\|^2 = \min_y, \quad (4.106)$$

где $\alpha > 0$ – параметр регуляризации, играющий роль неопределенного множителя Лагранжа (о способах выбора α см. ниже). Из условия (4.106) вытекает уравнение Эйлера–Тихонова (ср. (4.80)) [8, с. 79]:

$$(\alpha E + A^* A) y_\alpha = A^* f, \quad (4.107)$$

где E – единичный оператор ($E y = y$). Итак, вместо уравнения I рода (4.104) получено уравнение II рода (4.107).

Краткий анализ метода. Проанализируем условие (4.106) и уравнение (4.107).

Если $\alpha = 0$, то метод регуляризации Тихонова переходит в МНК Гаусса с минимальной невязкой $\|Ay - f\|^2$ и крайне неустойчивым решением. С увеличением же α решение y_α становится глаже и устойчивей, т.е. уменьшается норма решения $\|y_\alpha\|^2$, но увеличивается невязка. Истина – посередине, т.е. при некотором умеренном α решение y_α будет иметь и умеренную гладкость, и умеренную невязку.

Если $\delta, \xi \rightarrow 0$, то $\alpha \rightarrow 0$ и

$$y_\alpha = \lim_{\alpha \rightarrow 0} (\alpha E + A^* A)^{-1} A^* f \equiv A^* f \quad (4.108)$$

(см. (4.102)), т.е. решение y_α переходит в нормальное псевдорешение. Таким образом, метод регуляризации Тихонова является обобщением метода наименьших квадратов Гаусса и метода псевдообратного оператора Мура–Пенроуза.

Метод регуляризации Тихонова устойчив, т.е. выполняется 3-й пункт корректности по Адамару, и эту устойчивость можно объяснить математически следующим образом. Оператор $A^* A$ в (4.107) является симметричным положительно определенным, поэтому все его собственные значения вещественны и неотрицательны: $\lambda_i(A^* A) \geq 0$, причем $\lambda(A^* A)_{\min} = 0$. Наличие же слагаемого αE в (4.107) увеличивает все $\lambda_i(A^* A)$ на α , поэтому $\lambda(\alpha E + A^* A)_{\min} = \alpha$. Вследствие этого оператор $\alpha E + A^* A$ становится обратимым, норма обратного оператора $\|(\alpha E + A^* A)^{-1}\| = 1/\alpha \neq \infty$ и задача становится устойчивой.

Решение уравнения (4.107) равно

$$y_\alpha = (\alpha E + A^* A)^{-1} A^* f. \quad (4.109)$$

Важным является вопрос о выборе значения параметра регуляризации α . Разработан ряд способов выбора α в методе регуляризации Тихонова. Приведем два из них.

1-й способ – способ невязки [4, 8, 34]. Согласно нему, α выбирается из условия (при $\xi = 0$):

$$\| Ay_\alpha - f \| = \delta. \quad (4.110)$$

Если $\| f \| \geq \delta$, то решение уравнения (4.110) относительно α существует и является единственным. При $\xi \neq 0$ способ невязки переходит в *обобщенный принцип невязки* [10]. При этом принято делить все способы выбора α , использующие погрешности δ и ξ , на два типа – априорный и апостериорный [44]. Способ невязки и обобщенный принцип невязки относятся к последнему типу.

2-й способ – *способ подбора* [7, 14, 39, 40]. Согласно нему, находятся решения y_α для ряда значений α в «разумном» диапазоне и окончательный выбор α делается на основе не математических, а в основном физиологических критериев восприятия. Способ весьма напоминает выбор контраста телеизображения. Действительно, при уменьшении α повышается неустойчивость решения y_α или контраст изображения, если под изображением подразумевать функцию y_α , и наоборот, при увеличении α повышается гладкость решения y_α или уменьшается контраст изображения. Несмотря на простоту способа, он весьма эффективен при решении таких задач, как реконструкция смазанных и дефокусированных изображений (см. п. 5.1, 5.2), получение рентгеновских томограмм (см. п. 5.3) и др.

Разработаны также следующие способы выбора параметра регуляризации α : способ квазиоптимального (квазинаилучшего), способ отношения, способ независимых реализаций, способ перекрестной значимости, способ моделирования и др. [4, 8, 19].

Ниже будет рассмотрено применение метода регуляризации Тихонова к устойчивому решению различных уравнений (интегральных уравнений, СЛАУ, дифференциальных уравнений и т.д.).

Замечание. Помимо метода регуляризации Тихонова, для решения уравнений I рода (интегральных, алгебраических, дифференциальных) разработаны следующие *устойчивые методы*: методы регуляризации Лаврентьева, Денисова, Бакушинского, итеративной регуляризации Фридмана, Вайникко, Емелина, Красносельского, локальной регуляризации Арсенина, Воскобойникова, квазирешений Иванова, решения на компакте, дескриптивной регуляризации Морозова и др. [4–6, 9, 10, 19, 30, 34]. Это – *методы детерминистской регуляризации*. Разработаны также *методы статистической регуляризации* Турчина, Халфина, Лаврентьева, Жуковского, *оптимальной фильтрации* Калмана–Бьюси, Винера, субоптимальной фильтрации и др. [4, 19, 51]. Однако в данном учебнике мы ограничимся изложением метода регуляризации Тихонова.

Решение интегрального уравнения Фредгольма I рода. Рассмотрим интегральное уравнение Фредгольма I рода:

$$Ay \equiv \int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d. \quad (4.111)$$

Применение метода регуляризации Тихонова к уравнению (4.111) приводит к интегральному уравнению Фредгольма II рода (ср. (4.90), (4.107)):

$$\alpha y_\alpha(t) + \int_a^b R(t, s) y_\alpha(s) ds = F(t), \quad a \leq t \leq b, \quad (4.112)$$

с новым (симметричным положительно определенным) ядром (ср. (4.91))

$$R(t, s) = R(s, t) = \int_c^d K(x, t) K(x, s) dx \quad (4.113)$$

и с новой правой частью (ср. (4.92))

$$F(t) = \int_c^d K(x, t) f(x) dx. \quad (4.114)$$

Численный алгоритм. Рассмотрим вопрос о численном решении интегрального уравнения (4.112). Остановимся на одном из наиболее эффективных алгоритмов его решения – *методе квадратур* [4, с. 249–251], [7, с. 195], [10].

Пусть правая часть $f(x)$ задана таблично на следующей, вообще говоря, неравномерной x -сетке узлов:

$$c = x_1 < x_2 < x_3 < \dots < x_l = d, \quad (4.115)$$

а решение $y_\alpha(s)$ ищется на другой неравномерной s -сетке узлов, совпадающей с t -сеткой:

$$a = s_1 = t_1 < s_2 = t_2 < s_3 = t_3 < \dots < s_n = t_n = b, \quad (4.116)$$

причем $l \leq n$ или $l \geq n$.

Распишем интеграл в (4.112) по некоторой квадратурной формуле, лучше всего по формуле трапеций. Получим:

$$\alpha y_k + \sum_{j=1}^n r_j R_{kj} y_j = F_k, \quad k = 1, \dots, n, \quad (4.117)$$

где $y_k \equiv y_\alpha(t_k)$, $y_j \equiv y_\alpha(s_j)$, $R_{kj} \equiv R(t_k, s_j)$, $F_k \equiv F(t_k)$. Аналогично интегралы в (4.113) и (4.114) аппроксимируем конечными суммами по квадратурной формуле. Получим:

$$R_{kj} = R_{jk} = \sum_{i=1}^l p_i K_{ik} K_{ij}, \quad k, j = 1, \dots, n, \quad (4.118)$$

$$F_k = \sum_{i=1}^l p_i K_{ik} f_i, \quad k = 1, \dots, n, \quad (4.119)$$

где $K_{ik} \equiv K(x_i, t_k)$, $K_{ij} \equiv K(x_i, s_j)$, $f_i \equiv f(x_i)$. В формулах (4.117)–(4.119) r_j и p_i – коэффициенты квадратурной формулы.

Запись (4.117) есть СЛАУ относительно y_j , $j = 1, \dots, n$. Подробности численного алгоритма см. в [4, с. 249–251].

Программы. В [4, с. 371–379] приведены программы ТИKN1, ТИKN2, ТИKN3, ТИKN4, ТИKN5, а в [10, с. 104–107, 157–174] – программы PTIMR, PTIZR на Фортране, реализующие метод регуляризации Тихонова согласно формул типа (4.115)–(4.119) с различными способами выбора параметра регуляризации α .

Численный пример. Приведем результаты решения интегрального уравнения (4.111) методом регуляризации Тихонова. Рассмотрим модельный пример [7, с. 198–199], в котором точное решение задавалось равным сумме пяти гауссиан

$$y(s) = 6.5 e^{-[(s+0.66)/0.085]^2} + 9 e^{-[(s+0.41)/0.075]^2} + \\ + 12 e^{-[(s-0.14)/0.084]^2} + 14 e^{-[(s-0.41)/0.095]^2} + 9 e^{-[(s-0.67)/0.065]^2},$$

$a = -0.85$, $b = 0.85$, $c = -1$, $d = 1$, ядро

$$K(x, s) = \sqrt{q/\pi} e^{-q(x-s)^2/(1+x^2)},$$

где точное $q = 59.924$. Шаги дискретизации $\Delta x = \Delta s = \text{const} = 0.0125$, числа узлов $l = 161$, $n = 137$.

Сначала решалась прямая задача. Значения f_i , $i = 1, \dots, l$, вычислялись численно по квадратурной формуле трапеций

$$f_i \equiv \sum_{j=1}^N p_j K(x_i, s_j) y(s_j)$$

с шагом, значительно меньшим, чем $\Delta s = 0.0125$, т.е. при $N \gg n$. Затем с помощью программы RNDAN (датчика случайных чисел) [7, с. 152] к значениям f_i были добавлены погрешности δf_i , распределенные по нормальному закону с нулевым математическим ожиданием и среднеквадратическим отклонением $\delta = 0.0513$, что соответствует относительной погрешности правой части $\delta_{\text{отн}} \approx 1\%$.

Затем решалась обратная задача. Точное значение q заменялось на приближенное $\tilde{q} = 60$ – это соответствует $\xi_{\text{отн}} \approx 1\%$. При-

мер решался по программам ТІКН2 и ТІКН3. Значение параметра регуляризации α было выбрано способом моделирования (путем решения ряда близких модельных примеров) с помощью программы ТІКН2: $\alpha = 10^{-3.5}$. Затем по программе ТІКН3 найдено решение $y_\alpha(s)$. На рис. 4.3 представлены: точное решение $y(s)$, точная правая часть $f(x)$ и регуляризованное решение $y_\alpha(s)$ (а на рис. 4.2 дано решение этого же примера без регуляризации).

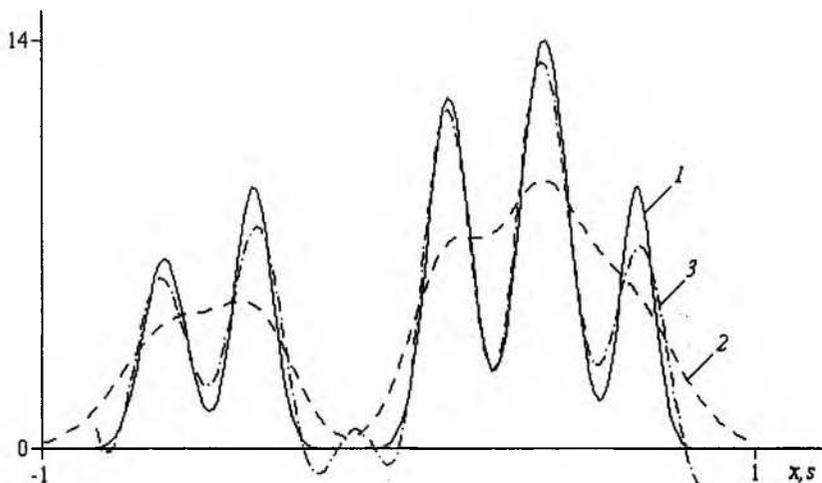


Рис. 4.3. Решение интегрального уравнения Фредгольма I рода методом регуляризации Тихонова
 1 — $y(s)$ 2 — $f(x)$ 3 — $y_\alpha(s)$

Данный пример характерен, например, для задачи гидроакустики (редукции протяженных сигналов [7, с. 114–115]). Функция $y(s)$ – это интенсивность звукового поля в зависимости от направления s , функция $K(x, s)$ – характеристика направленности (по мощности) антенны, измеряющей поле, а $\tilde{f}(x)$ – результат измерений (индикаторный процесс). Задача состоит в том, что по измеренной $\tilde{f}(x)$ и известной (с погрешностью) $\tilde{K}(x, s)$ необходимо математическим путем восстановить входное поле $y(s)$.

При этом искомое поле $y(s)$ имеет значительные флуктуации. Однако измеренная функция $f(x)$, во-первых, практически не имеет флуктуаций за счет конечной ширины характеристики направленности $K(x, s)$, а во-вторых, отягощена погрешностями измерений.

Рис. 4.3 показывает, что метод регуляризации Тихонова позволил вполне удовлетворительно восстановить поле на входе антенны $y(s)$. Это равносильно тому, что измерительное устройство (антенна), сопряженное с вычислительным устройством (компьютером, спецпроцессором и т.п.) с заложенной в него программой, реализующей метод регуляризации Тихонова, трансформируется в новое измерительное устройство с более высокой разрешающей способностью.

Данный пример можно трактовать также как пример из *обратной задачи спектроскопии* [7, с. 79], полагая, что переменная x , а также s – это частота (или длина волны), функция $y(s)$ – истинное распределение интенсивности по спектру, $K(x, s)$ – частотная характеристика (ЧХ) спектрометра, а $f(x)$ – измеренный спектр. В этом случае задача заключается в том, что по измеренному спектру $\tilde{f}(x)$ и известной ЧХ нужно математическим путем восстановить истинный спектр $y(s)$, тем самым повысить разрешающую способность спектрометра.

Данные задачи (гидроакустики и спектроскопии) – это одни из многочисленных задач, иллюстрирующих так называемую *редукционную проблему Рэлея* [7, с. 111].

Рассмотрим важный частный случай уравнения (4.111) – интегральное уравнение Фредгольма I рода типа свертки одномерное (а также двумерное).

Метод регуляризации Тихонова для интегральных уравнений I рода типа свертки. Рассмотрим *одномерное интегральное уравнение Фредгольма I рода типа свертки*:

$$Ay \equiv \int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (4.120)$$

В п. 4.5 рассмотрено его решение методом преобразования Фурье и показано, что такое решение является неустойчивым и эта неустойчивость обусловлена сильной чувствительностью высоких гармоник Фурье к погрешностям правой части $f(x)$.

Устойчивый алгоритм решения уравнения (4.120) дает метод регуляризации Тихонова. Согласно нему, устойчивое решение уравнения находится из условия *минимума сглаживающего функционала* (ср. (4.106))

$$\int_{-\infty}^{\infty} [Ay - f(x)]^2 dx + \alpha \int_{-\infty}^{\infty} M(\omega) |Y(\omega)|^2 d\omega = \min_y, \quad (4.121)$$

где

$$M(\omega) = |\omega|^{2q} \quad (4.122)$$

– регуляризатор q -го порядка, причем $q > 0$ – задаваемый порядок регуляризации, например, $q = 1$.

Из условия (4.121) получается следующее выражение для регуляризованного решения (ср. (4.72)):

$$y_\alpha(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-\omega) F(\omega)}{L(\omega) + \alpha M(\omega)} e^{-i\omega s} d\omega, \quad (4.123)$$

где

$$\lambda(\omega) = \int_{-\infty}^{\infty} K(x) e^{i\omega x} dx, \quad (4.124)$$

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx \quad (4.125)$$

– спектры (или ПФ) ядра $K(x)$ и правой части $f(x)$ и

$$L(\omega) = |\lambda(\omega)|^2 = \lambda(\omega)\lambda(-\omega) = \operatorname{Re}^2 \lambda(\omega) + \operatorname{Im}^2 \lambda(\omega). \quad (4.126)$$

Сравним классическое решение (4.72) и регуляризованное решение (4.123). В (4.123) подынтегральная функция стремится к нулю при $|\omega| \rightarrow \infty$ за счет слагаемого $\alpha M(\omega)$, подавляющего реакцию высоких гармоник Фурье на погрешность правой части $f(x)$, причем подавление тем сильнее, чем больше значения α и q . При этом чем больше q , тем сильнее подавляются высокие гармоники в решении по сравнению с низкими, параметр же α определяет глобальное подавление: с увеличением α сильнее подавляются все гармоники. Поэтому в отношении q следует использовать следующее правило: если искомое решение имеет флуктуации (и значит, спектр решения имеет высокочастотную составляющую), то q следует понизить, например, положить $q = 1$, а если решение гладкое (и его спектр не содержит высоких гармоник), то q следует повысить, например, $q = 2$. Что же касается α , то способы его выбора те же, что и для уравнения (4.112) (способы невязки, подбора и др.).

Численный алгоритм решения $y_\alpha(s)$ получается в результате замены интегралов (4.123)–(4.125) конечными суммами (по формулам прямоугольников, трапеций и т.д.). В результате непрерывные преобразования Фурье (НПФ) заменяются на дискретные преобразования Фурье (ДПФ) или даже на быстрые преобразования Фурье (БПФ) [7, с. 166–170].

Разработаны и опубликованы следующие программы для решения уравнения (4.120) методом регуляризации Тихонова: РТИКР [10,

с. 124–130, 178–179], CONV1, CONV2, CONV3, CONV4, CONV5 [4, с. 379–388] и др.

Численный пример. Приведем результаты решения следующего примера [7, с. 201–202], [51] (близкого к примеру [10, с. 43, 127–129]):

$$\int_0^1 k(t-\tau) y(\tau) d\tau = f(t), \quad 0 \leq t \leq 2, \quad (4.127)$$

где ядро

$$k(t) = e^{-80(t-0.5)^2},$$

точное решение

$$y(\tau) = \left\{ 0.45 \exp \left[- \left(\frac{\tau - 0.29}{0.18} \right)^2 \right] + \exp \left[- \left(\frac{\tau - 0.71}{0.16} \right)^2 \right] \right\} \sqrt{1 - \left(\frac{\tau - 0.5}{0.5} \right)^2},$$

локальные носители (области, вообще говоря, ненулевых значений функций): $\text{supp}k(t) \subseteq [0,1]$, $\text{supp}f(t) \subseteq [0,2]$, $\text{supp}y(\tau) \subseteq [0,1]$. Уравнение (4.127) хотя и имеет конечные пределы изменения τ и t , но является уравнением типа свертки, так как вне пределов изменения τ и t (вне локальных носителей) функции $k(t)$, $f(t)$ и $y(\tau)$ равны нулю, т.е. в действительности пределы изменения τ и t бесконечны.

Согласно численной методике дискретизации задачи, изложенной в [10, с. 39, 123], вместо (4.127) решалось уравнение:

$$\int_{-0.5}^{1.5} k(t-\tau) y(\tau) d\tau = f(t), \quad 0 \leq t \leq 2,$$

с новыми носителями равной длины: $\text{supp}k(t) \subseteq [-0.5,1.5]$, $\text{supp}f(t) \subseteq [0,2]$, $\text{supp}y(\tau) \subseteq [-0.5,1.5]$. Число узлов дискретизации положено равным $N = 64$. Шаг дискретизации равен $h = 2/N$. Узлы дискретизации по k и y равны $t_j = -0.5 + h \cdot (j - 1/2)$, $j = 1, \dots, N$, а по f равны $t_i = h \cdot (i - 1/2)$, $i = 1, \dots, N$.

Вместо точной $f(t)$ использовалась зашумленная $\tilde{f}(t) = f(t) + \delta f(t)$ со случайной погрешностью $\delta f(t)$, распределенной по нормальному закону с нулевым математическим ожиданием и среднеквадратическим отклонением $\delta = 0.0164$, что соответствует относительной погрешности правой части $\delta_{\text{отн}} \approx 10\%$ от $\max_{t \in [0,2]} f(t)$.

На рис. 4.4 приведены точные $k(t)$, $y(\tau)$ и $f(t)$, а также $\tilde{f}(t)$.

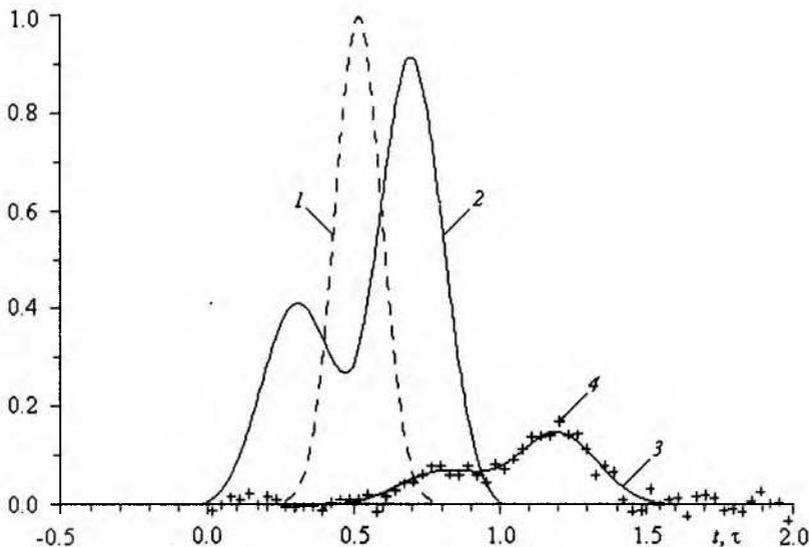


Рис. 4.4. Численный пример: интегральное уравнение Фредгольма 1 рода типа свертки

1 — $k(t)$ 2 — $y(\tau)$ 3 — $f(t)$ 4 — $\tilde{f}(t)$

На рис. 4.5 даны: $y(\tau)$ — точное решение и результаты решения при $\text{supp}k(t) \subseteq [-0.5, 1.5]$, $\text{supp}f(t) \subseteq [0, 2]$, $\text{supp}y(\tau) \subseteq [-0.5, 1.5]$, а именно, $y_{\alpha_1}(\tau)$ — решение методом регуляризации Тихонова со значением α , выбранным по способу невязки ($\alpha = \alpha_1 = 4 \cdot 10^{-4}$) — слишком гладкое (хотя и очень устойчивое) решение и $y_{\alpha_2}(\tau)$ — решение методом регуляризации Тихонова при $\alpha = \alpha_2 \ll \alpha_1$, а именно, $\alpha_2 = 10^{-5}$ — решение с гораздо лучшим разрешением двух максимумов в решении, хотя и более неустойчивое, имеющее на краях ложные знакопеременные флуктуации — проявление так называемого эффекта Гиббса.

Данный пример, как и ряд других примеров ([4, с. 283], [10, с. 34, 47]) показывает, что при больших погрешностях исходных данных ($\sim 10\%$) способ невязки (а также обобщенный принцип невязки) дает завышенное значение α и нужно использовать дополнительные способы, например, способ подбора.

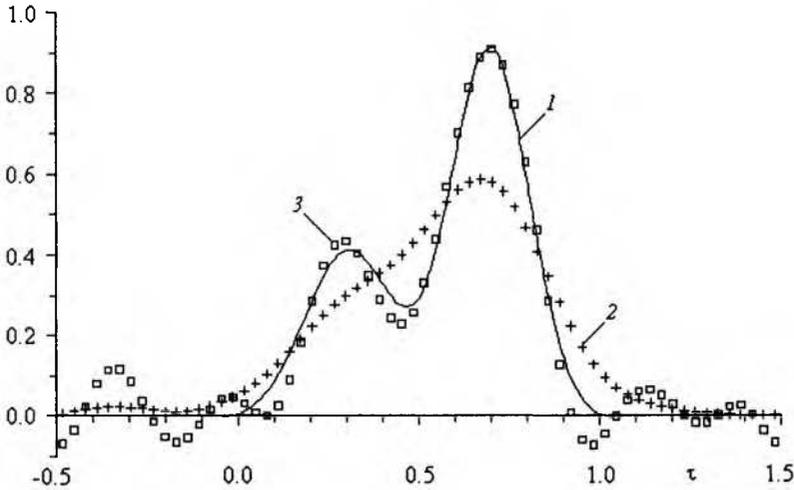


Рис. 4.5. Решение интегрального уравнения Фредгольма I рода типа свертки методом регуляризации Тихонова
 1 — $y(\tau)$ 2 — $y_{\alpha_1}(\tau)$ 3 — $y_{\alpha_2}(\tau)$

Что же касается ложных флуктуаций в решении $y_{\alpha_2}(\tau)$ на краях области $\tau \in [-0.5, 1.5]$, то они могут быть снижены или вообще устранены, например, путем сокращения длин носителей. На рис. 4.6 приведены результаты решения этого же примера при сокращенных (равных) длинах носителей, а именно, при $\text{supp}k(t) \subseteq [0, 1]$, $\text{supp}f(t) \subseteq [0.5, 1.5]$, $\text{supp}y(\tau) \subseteq [0, 1]$ (обозначения те же, что и на рис. 4.5).

Рассмотрим также *двухмерное интегральное уравнение Фредгольма I рода типа свертки*:

$$A y \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x_1 - s_1, x_2 - s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2), \quad -\infty < x_1, x_2 < \infty. \quad (4.128)$$

Применим для его решения метод регуляризации Тихонова. Введем условие *минимума сглаживающего функционала* (ср. (4.121)):

$$\| A y - f \|^2 + \alpha \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(\omega_1, \omega_2) |Y(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2 = \min_y.$$

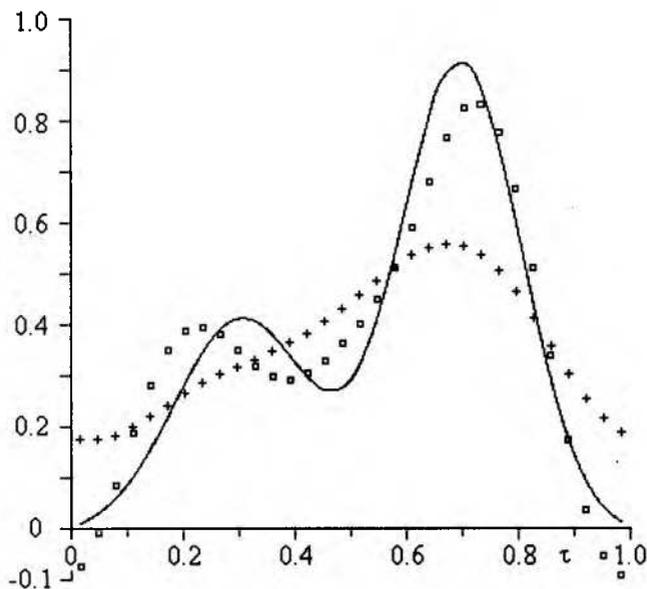


Рис. 4.6

Из этого условия следует *регуляризованное решение* (ср. (4.123))

$$y_{\alpha}(s_1, s_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\lambda(-\omega_1, -\omega_2) F(\omega_1, \omega_2)}{L(\omega_1, \omega_2) + \alpha M(\omega_1, \omega_2)} e^{-i(\omega_1 s_1 + \omega_2 s_2)} d\omega_1 d\omega_2, \quad (4.129)$$

где (ср. (4.124)–(4.126))

$$F(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) e^{i(\omega_1 x_1 + \omega_2 x_2)} dx_1 dx_2, \quad (4.130)$$

$$\lambda(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x_1, x_2) e^{i(\omega_1 x_1 + \omega_2 x_2)} dx_1 dx_2, \quad (4.131)$$

$$\begin{aligned} L(\omega_1, \omega_2) &= |\lambda(\omega_1, \omega_2)|^2 = \lambda(\omega_1, \omega_2) \lambda(-\omega_1, -\omega_2) = \\ &= \operatorname{Re}^2 \lambda(\omega_1, \omega_2) + \operatorname{Im}^2 \lambda(\omega_1, \omega_2). \end{aligned} \quad (4.132)$$

Регуляризатор $M(\omega_1, \omega_2)$ может быть выбран, например, в виде (ср. (4.122)):

$$M(\omega_1, \omega_2) = (\omega_1^2 + \omega_2^2)^2. \quad (4.133)$$

При практической реализации двумерной задачи значения x_1, s_1, x_2, s_2 задаются дискретно и в конечных пределах, двумерные НПФ (4.129)–(4.131) заменяются на двумерные ДПФ или БПФ.

Параметр регуляризации α выбирается теми же способами, что и в случае одномерных уравнений (4.111) и (4.120) (способами невязки, обобщенной невязки, подбора и др.).

В [10, с. 130–136, 185–186] приведена программа RTITR на Фортране для решения двухмерного интегрального уравнения Фредгольма I рода типа свертки (4.128) методом регуляризации Тихонова.

В п. 5.2 изложена задача реконструкции дефокусированных изображений, описываемая двухмерным интегральным уравнением Фредгольма I рода типа свертки. Для его решения использован метод регуляризации Тихонова.

Решение вырожденных и плохо обусловленных СЛАУ. В п. 4.2–4.4 мы уже рассматривали вырожденные и плохо обусловленные СЛАУ.

Запишем систему линейных алгебраических уравнений (СЛАУ) в виде:

$$Ay = f \quad (4.134)$$

или подробнее

$$\sum_{j=1}^n a_{ij} y_j = f_i, \quad i = 1, \dots, m, \quad (4.135)$$

где A – $m \times n$ -матрица, y – n -вектор-столбец, f – m -вектор-столбец.

Эта система может быть неразрешимой, т.е. не имеющей решения (когда она является переопределенной), однозначно разрешимой (когда она является определенной) и вырожденной (имеющей бесконечно много решений, когда она является недоопределенной). Напомним определения, данные в п. 4.2–4.4.

Рангом $r = \text{rang}(A)$ матрицы A называется максимальный порядок отличных от нуля миноров. *Рангом расширенной матрицы* называется величина $\rho = \text{rang}(A|f)$. СЛАУ называется *переопределенной*, если $m > n$, точнее, число линейно независимых строк в (4.135) больше n или $\rho > r$. СЛАУ называется *определенной*, если $m = n = r$ или $\rho = r$. СЛАУ называется *недоопределенной*, если $m < n$ или $r < n$ или $\rho < n$. Переопределенная СЛАУ не имеет решений, определенная СЛАУ имеет единственное решение, а недоопределенная СЛАУ имеет множество решений.

СЛАУ называется *вырожденной*, если ее $m \times n$ -матрица A имеет хотя бы одно нулевое сингулярное число μ . В этом случае число обусловленности $\text{cond}(A) = \mu(A)_{\max} / \mu(A)_{\min} = \infty$. Если же $m = n$, т.е. матрица системы квадратная, то в этом случае СЛАУ называется *вырожденной*, если ее определитель $|A| = 0$. СЛАУ является *плохо обуслов-*

ленной, если число обусловленности $\text{cond}(A)$ ее матрицы A достаточно велико, хотя и конечно.

Псевдорешением СЛАУ (4.134) называют вектор y , минимизирующий невязку $\|Ay - f\|$. Система (4.134) может иметь не одно псевдорешение. Нормальным решением СЛАУ (4.134) называется псевдорешение с минимальной нормой $\|y\|$. Для любой системы вида (4.134) нормальное решение существует и единственно.

Однако нормальное решение может быть неустойчивым и оно тем неустойчивее, чем больше число обусловленности

$$\text{cond}(A) = \mu(A)_{\max} / \mu(A)_{\min}.$$

Метод регуляризации Тихонова решения СЛАУ заключается в получении устойчивого приближения к ее нормальному решению.

Вводится условие минимизации *сглаживающего функционала* (ср. (4.106)):

$$\|Ay - f\|^2 + \alpha \|y\|^2 = \min_y, \quad (4.136)$$

где $\alpha > 0$ – параметр регуляризации. Из условия (4.136) получается новая СЛАУ (ср. (4.107)):

$$(\alpha E + A^* A)y_\alpha = A^* f, \quad (4.137)$$

где E – единичная матрица. Систему (4.137) можно записать иначе (ср. (4.81)):

$$(\alpha E + B)y_\alpha = u, \quad (4.138)$$

где (см. (4.82), (4.83))

$$B = A^* A, \quad (4.139)$$

$$u = A^* f \quad (4.140)$$

– новая матрица B (симметричная и положительно определенная) и новая правая часть u или в случае вещественности A

$$B = A^T A, \quad (4.141)$$

$$u = A^T f. \quad (4.142)$$

Более подробно СЛАУ (4.138) можно записать в виде:

$$\alpha y_i + \sum_{j=1}^n B_{ij} y_j = u_i, \quad i = 1, \dots, n, \quad (4.143)$$

где в случае вещественности A (см. (4.86), (4.87))

$$B_{ij} = \sum_{k=1}^m A_{ki} A_{kj}, \quad i, j = 1, \dots, n, \quad (4.144)$$

$$u_i = \sum_{k=1}^m A_{ki} f_k, \quad i = 1, \dots, n. \quad (4.145)$$

Регуляризованное решение СЛАУ (4.137) или (4.138) имеет вид:

$$y_\alpha = (\alpha E + A^* A)^{-1} A^* f \quad (4.146)$$

или

$$y_\alpha = (\alpha E + B)^{-1} u, \quad (4.147)$$

а практически может быть найдено численными методами путем решения СЛАУ (4.143) с учетом того, что матрица $\alpha E + B$ является симметричной и положительно определенной – методами Краута, Холецкого, Воеводина и др.

При этом параметр регуляризации α может быть выбран одним из ряда способов: способом невязки, обобщенным принципом невязки и др.

Решение системы обыкновенных дифференциальных уравнений с управляющей функцией. Рассмотрим (как и в п. 4.3, 4.4) систему обыкновенных дифференциальных уравнений (ОДУ) с управляющей функцией [8, с. 241]:

$$\frac{dy(t)}{dt} = f(t, y, u), \quad t_0 \leq t \leq T, \quad (4.148)$$

где $y(t) = \{y_1(t), \dots, y_n(t)\}$ – искомая функция, $f(t) = \{f_1(t), \dots, f_n(t)\}$ – правая часть, $u(t) = \{u_1(t), \dots, u_m(t)\}$ – управляющая функция (все вектор-функции). Начальные условия:

$$y(t_0) = y_0,$$

где y_0 – заданный вектор.

Функцию $u(t)$ можно считать заданной – в этом случае управление может быть неоптимальным, а можно считать и искомой, минимизирующей некоторый функционал $F[u]$ (с возможными ограничениями на $u(t)$) – в этом случае управление будет оптимальным.

Рассмотрим конкретный пример задачи Коши для системы ОДУ с оптимальным управлением – *вертикальное движение ракеты* переменной массы в однородной атмосфере с целью достижения максимальной высоты. Это движение описывается следующей системой ОДУ [8, с. 242–245]:

$$\left. \begin{aligned} \frac{dv(t)}{dt} &= \frac{au(t) - cv^2(t)}{m(t)} - g, \\ \frac{dm(t)}{dt} &= -u(t) \end{aligned} \right\} \quad (4.149)$$

с начальными условиями: $m(0) = m_0$, $v(0) = 0$. Здесь $m(t)$ – переменная масса всей ракеты, причем $\mu \leq m(t) \leq m_0$, где μ – масса ракеты без горючего, $v(t)$ – скорость подъема ракеты, $u(t)$ – управляющая функция, равная расходу горючего в функции времени, a , c , g – постоянные (a – скорость истечения газов относительно ракеты, c – коэффициент сопротивления воздуха, g – ускорение силы тяжести).

Максимальная высота подъема ракеты будет равна $H = H[v(u)] = \int_0^T v(t) dt$, причем $v(T) = 0$.

Ставится задача отыскания оптимальной управляющей функции $u_{\text{opt}}(t)$ – такой функции, при которой $H = \max$. Эта задача обычно решается следующим образом. Вводится в рассмотрение так называемый целевой функционал, который, как правило, выбирается в виде:

$$F[u] = H[v(u)] - \tilde{H}, \quad (4.150)$$

где \tilde{H} – оценка искомой максимальной высоты. Задача отыскания $u_{\text{opt}}(t)$ решается из условия минимизации целевого функционала:

$$F[u_{\text{opt}}(t)] = \min_{u(t)}. \quad (4.151)$$

Однако данная задача некорректна: малым вариациям функционала F будут соответствовать сколь угодно большие вариации $u(t)$, т.е. нарушается 3-й пункт определения корректности по Адамару.

Изложим метод регуляризации решения системы ОДУ, делающий задачу устойчивой. Согласно методу регуляризации, нужно минимизировать *сглаживающий функционал*

$$\Phi_\alpha[u(t)] = F[u(t)] + \alpha \Omega[u(t)], \quad (4.152)$$

где *стабилизирующий функционал* $\Omega[u(t)]$ выбирается, например, в виде:

$$\Omega[u(t)] = \|u\|_{L_2}^2 = \int_0^T u^2(t) dt \quad (4.153)$$

или

$$\Omega[u(t)] = \|u\|_{W_2^2}^2 = \int_0^T u'^2(t) dt, \quad (4.154)$$

а $\alpha > 0$ – параметр регуляризации.

При этом минимизацию функционала $\Phi_\alpha[u(t)]$ нужно выполнять численно, например, методом градиентов, Ньютона, хорд и т.д., а α выбирать, например, способом невязки. В [8, с. 244–245] приведен численный пример.

Решение дифференциальных уравнений в частных производных. Сначала рассмотрим приведенное в п. 4.3 и 4.4 двухмерное уравнение Лапласа (уравнение эллиптического типа):

$$\Delta u(x, y) \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (4.155)$$

со следующим типом *граничных условий* (задачей Коши):

$$u(x, 0) = f(x), \quad \left. \frac{\partial u}{\partial y} \right|_{y=0} = \varphi(x), \quad (4.156)$$

где $f(x)$ и $\varphi(x)$ – заданные функции.

Известно [6], что при минимальных требованиях, налагаемых на решение $u(x, y)$, оно существует и является единственным, т.е. выполняются 1-й и 2-й пункты корректности по Адамару. Однако, как показано в п. 4.4, решение является неустойчивым.

Изложим один из возможных алгоритмов построения регулярного (устойчивого) решения задачи Коши для уравнения Лапласа – *алгоритм приведения к интегральному уравнению*.

Для этого уравнение Лапласа (4.155) приведем к интегральному уравнению Фредгольма I рода с ядром Пуассона [6, с. 7], [22]:

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{y^2 + (x - \xi)^2} u(\xi, y) d\xi = u(x, 0), \quad -\infty < x < \infty. \quad (4.157)$$

В этом уравнении y играет роль параметра. Поэтому при каждом фиксированном значении y уравнение является одномерным интегральным уравнением Фредгольма I рода типа свертки. Такое уравнение эффективно решается методом преобразования Фурье с регуляризацией по Тихонову (см. (4.120)–(4.126)).

В [4, с. 30, 146, 152–157, 230] изложен подробно вопрос о приведении различных дифференциальных уравнений (в обыкновенных или в частных производных) или их систем с различными граничными условиями к соответствующим интегральным уравнениям.

Отметим, что уравнение (4.157) широко используется в обратной задаче гравиметрии, а именно, в задаче аналитического продолжения измеренного гравитационного потенциала u с поверхности Земли ($y = 0$) в ее глубь (на глубину y), или в задаче пересчета гравитационного потенциала.

Теперь рассмотрим приведенное в п. 4.3 и 4.4 одномерное уравнение теплопроводности (уравнение параболического типа). Напомним, что рассмотрено четыре варианта постановки задачи его решения.

В 1-м варианте (в прямой задаче, или задаче с прямым ходом вре-

мени) решается уравнение

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}, \quad 0 \leq x \leq l, \quad t > 0, \quad (4.158)$$

с начальным условием: $u(x,0) = \varphi(x)$ и граничными условиями: $u(0,t) = \psi_1(t)$, $u(l,t) = \psi_2(t)$. Прямая (классическая) задача является корректной: решение существует, является единственным и непрерывно зависит от начальных данных $\varphi(x)$.

Во 2-м варианте (в обратной задаче, или задаче с обратным ходом времени) по заданной функции $u(x,t_*)$ в некоторый момент времени $t_* > 0$ необходимо, решая задачу в направлении уменьшения t , определить начальное распределение температуры $u(x,0) = \varphi(x)$. При этом граничные условия $\psi_1(t)$ и $\psi_2(t)$ не заданы.

В 3-м варианте задачи необходимо решить уравнение

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}, \quad -\infty < x < \infty, \quad 0 \leq t \leq T, \quad (4.159)$$

с начальными условиями: $u(x,0) = \varphi(x)$ и $u(x,T) = \chi(x)$, где $\varphi(x)$ и $\chi(x)$ – заданные функции.

В 4-м варианте задачи необходимо решить уравнение (4.158) или (4.159), однако лишь с одним начальным условием: $u(x,0) = \varphi(x)$.

Как отмечено в п. 4.4, варианты 2–4 некорректны. Изложим кратко два устойчивых метода решения вариантов 2–4 задачи решения уравнения теплопроводности.

1-й метод. 2-й и 3-й варианты задачи могут быть решены методом квазиобращения [30]. Согласно нему, вместо уравнения (4.159) решается *регуляризованное уравнение* [5, с. 108], [8, с. 52]:

$$\frac{\partial u_\alpha(x,t)}{\partial t} = \frac{\partial^2 u_\alpha(x,t)}{\partial x^2} + \alpha \frac{\partial^4 u_\alpha(x,t)}{\partial x^4}, \quad -\infty < x < \infty, \quad 0 \leq t \leq T, \quad (4.160)$$

где $\alpha > 0$ – параметр регуляризации. При этом задаем лишь начальное условие $u_\alpha(x,T) = \chi(x)$ и, идя от $t = T$ до $t = 0$ (в сторону уменьшения t), находим устойчивым образом функцию $u_\alpha(x,0) = \varphi_\alpha(x)$.

После этого для проверки, решая корректную задачу Коши для уравнения (4.160) при начальном условии $u_\alpha(x,0) = \varphi_\alpha(x)$ и, идя от $t = 0$ до $t = T$ (в сторону увеличения t), можно найти $u_\alpha(x,t)$ и, в том числе, $u_\alpha(x,T) = \chi_\alpha(x)$. Как показано в [30], имеет место сходимость в среднем: $\chi_\alpha(x) \rightarrow \chi(x)$ при $\alpha \rightarrow 0$.

2-й метод. Это – метод *интегрального уравнения*. Согласно не-

му [6, с. 7], (дифференциальное) уравнение теплопроводности может быть приведено к интегральному уравнению:

$$\frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\xi)^2}{4t}\right] u(\xi, t) d\xi = \varphi(x), \quad -\infty < x < \infty. \quad (4.161)$$

Здесь $\varphi(x) = u(x, 0)$ – начальное условие. Уравнение (4.161) есть интегральное уравнение Фредгольма I рода типа свертки с параметром t . Решая данное уравнение при каждом фиксированном значении t методом одномерного преобразования Фурье и регуляризации Тихонова (см. выше), получим устойчивое решение $u_{\alpha}(\xi, t)$, где $\alpha > 0$ – параметр регуляризации.

4.8. Метод решения на компакте

Существо метода. Возможность построения устойчивых решений некорректных задач основывается на использовании априорной (дополнительной) информации о решении. В п. 4.7 был рассмотрен случай существенно некорректных задач, для решения которых используются методы (метод регуляризации Тихонова и др.), использующие *качественную информацию о решении*, а именно, информацию о его гладкости. Например, в методе регуляризации Тихонова гладкость решения регулируется параметром регуляризации α .

В данном пункте рассмотрим случай, когда имеется *количественная информация о решении*, а именно, информация, позволяющая сузить класс возможных решений до *компактного множества* (компакта^{*)}). Впервые такой подход к решению некорректных задач был сформулирован А. Н. Тихоновым в 1943 г. [41].

Рассмотрим уравнение

$$Ay = f, \quad y \in Y, f \in F, \quad (4.162)$$

где Y и F – некоторые метрические пространства, а A – непрерывный оператор. Пусть задача решения уравнения (4.162) является корректной по Тихонову (см. п. 4.7), для чего в пространстве Y выделено некоторое подпространство $M \subseteq Y$, являющееся множеством корректности (на котором решение существует, является единственным и устойчивым). Полагаем, что M является компактом (примеры компактов см. далее), а f и A заданы точно. В этом случае справедлива следующая теорема [41] (см. также [4, с. 302], [6, с. 31], [8, с. 39–40]).

Теорема 1 (Тихонова). Если отображение $M \rightarrow AM$ множества $M \subseteq Y$ на множество $AM \subseteq F$ (образ множества M) взаимно

^{*} Определение компакта см. в п. 4.1.

однозначно и непрерывно и M является компактом, то обратное отображение $AM \rightarrow M$ также непрерывно, т.е. на множестве $AM \subseteq F$ оператор, обратный к A , непрерывен.

Пусть далее все условия теоремы 1 выполнены и известно, что точное решение уравнения (4.162) $y \in M$, но вместо f известна \tilde{f} такая, что $\rho_F(\tilde{f}, f) \leq \delta$, причем $\tilde{f} \in AM$. В этом случае рассмотрим множество

$$Y_\delta^M = \{ \tilde{y} : \tilde{y} \in M, \rho_F(A\tilde{y}, \tilde{f}) \leq \delta \}.$$

Тогда также будет иметь место непрерывность обратного отображения, что на языке ϵ, δ формулируется в виде следующей теоремы [4, с. 302]:

Теорема 2. Для любого $\epsilon > 0$ существует $\delta_0(\epsilon) > 0$ такое, что $\rho_Y(\tilde{y}, y) < \epsilon$ (где y – точное решение) для всех $\tilde{y} \in Y_\delta^M$, как только $\delta < \delta_0(\epsilon)$.

Из теорем 1 и 2 следует, что во-первых, решение на компакте является устойчивым, а во-вторых, в случае приближенной правой части $\tilde{f} \in AM$ в качестве приближенного решения некорректной задачи следует принять любой элемент $\tilde{y} \in Y_\delta^M$. При этом $\lim_{\delta \rightarrow 0} \rho_Y(\tilde{y}, y) = 0$.

В качестве иллюстрации к изложенному рассмотрим интегральное уравнение Фредгольма I рода:

$$\int_a^b K(x, s) y(s) ds = \tilde{f}(x), \quad c \leq x \leq d, \quad (4.163)$$

где $\tilde{f}(x) \in L_2$. Рассмотрим случай приближенной правой части и точного ядра. Применительно к уравнению (4.163) метод отыскания решения на компакте заключается в минимизации функционала невязки (ср. (4.79))

$$\Phi(y) \equiv \|Ay - \tilde{f}\|_{L_2}^2 = \int_c^d \left[\int_a^b K(x, s) y(s) ds - \tilde{f}(x) \right]^2 dx \quad (4.164)$$

на функциях $\tilde{y} \in Y_\delta^M$, где M – некоторый компакт, задаваемый, исходя из физических предпосылок задачи, в виде ограничений типа неравенств и равенств (см. ниже).

Конкретизируем изложенное примерами компактов [4, с. 303–305], [10, с. 70, 72, 80], причем сделаем это применительно к интегральному уравнению (4.163).

Решение на множестве монотонных ограниченных функций.
 Это – 1-й пример компакта. Пусть априори (из физических соображений) известно, что точное решение \bar{y} некорректной задачи есть *монотонная* (для определенности *невозрастающая*) *функция, ограниченная* снизу и сверху соответственно константами C_1 и C_2 . Итак, полагаем, что искомое точное решение $\bar{y} \in M \in L_2$, где M – множество невозрастающих ограниченных снизу и сверху функций. Обозначим множество невозрастающих ограниченных снизу и сверху функций через $Y \downarrow_{C_1}^{C_2}$ (ср. [10, с. 70]), т.е.

$$Y \downarrow_{C_1}^{C_2} = \{y: y'(s) \leq 0, C_1 \leq y(s) \leq C_2, a \leq s \leq b\}. \quad (4.165)$$

Запись (4.165) означает множество функций, имеющих неположительную первую производную и ограниченных снизу и сверху на интервале $[a, b]$. Другими словами, в 1-м примере компакта наложены ограничения на функцию и ее первую производную.

Практический алгоритм решения задачи в данном случае имеет следующий вид.

Дискретизируем уравнение (4.163), используя, вообще говоря, неравномерные x -сетку узлов:

$$c = x_1 < x_2 < x_3 < \dots < x_l = d \quad (4.166)$$

и s -сетку узлов:

$$a = s_1 < s_2 < s_3 < \dots < s_n = b, \quad (4.167)$$

а также, например, формулу трапеций.

В результате задача условной минимизации функционала (4.164) запишется в виде:

$$\Phi(y) \equiv \sum_{i=1}^l p_i \left[\sum_{j=1}^n r_j K_{ij} y_j - \tilde{f}_i \right]^2 = \min_y, \quad (4.168)$$

с добавлением ограничений в виде неравенств (ср. [4, с. 303], [10, с. 73, 96]):

$$\left. \begin{aligned} y_{j+1} - y_j &\leq 0, & j = 1, 2, \dots, n-1, \\ y_1 - C_2 &\leq 0, & C_1 - y_n \leq 0. \end{aligned} \right\} \quad (4.169)$$

Здесь $K_{ij} \equiv K(x_i, s_j)$, $y_j \equiv y(s_j)$, $\tilde{f}_i \equiv \tilde{f}(x_i)$, а p_i и r_j – коэффициенты квадратурной формулы (см. [4, с. 250]).

При этом, как отмечено в [10, с. 86], нет необходимости находить точный минимум функционала невязки (4.168). Достаточно добиться выполнения условия:

$$\Phi(y) \equiv \sum_{i=1}^l p_i \left[\sum_{j=1}^n r_j K_{ij} y_j - \tilde{f}_i \right]^2 \leq \delta^2, \quad (4.170)$$

где δ – погрешность задания правой части: $\|\tilde{f} - \bar{f}\|_{L_2} \leq \delta$.

В случае дискретизации задачи множество $Y \downarrow_{C_1}^{C_2}$ (4.165) переходит в множество (ср. [10, с. 73])

$$M \downarrow_{C_1}^{C_2} = \left\{ y: y \in R^n, \quad y_{j+1} - y_j \leq 0, \quad j=1, 2, \dots, n-1, \right. \\ \left. y_1 - C_2 \leq 0, \quad C_1 - y_n \leq 0. \right\} \quad (4.171)$$

Итак, задача численного решения в классе монотонно невозрастающих ограниченных снизу и сверху функций заключается в минимизации функционала невязки $\Phi(y)$ (см. (4.168)) или доведении его значения до δ^2 или меньшего значения (см. (4.170)) с учетом ограничений на решение (4.169). Такая задача, называемая нелинейным (а также линейным, квадратичным и т.д.) программированием, может быть решена различными методами условной минимизации: методами проекций сопряженных градиентов, условного градиента и т.д.

Решение на множестве монотонных ограниченных выпуклых функций. Это – 2-й пример компакта. Пусть искомое точное решение \bar{y} принадлежит множеству *монотонно невозрастающих ограниченных* снизу и сверху константами C_1 и C_2 *выпуклых функций*. Обозначим это множество через $\hat{Y} \downarrow_{C_1}^{C_2}$ (ср. [10, с. 72]), т.е.

$$\hat{Y} \downarrow_{C_1}^{C_2} = \{ y: y'(s) \leq 0, \quad y''(s) \leq 0, \quad C_1 \leq y(s) \leq C_2, \quad a \leq s \leq b \}. \quad (4.172)$$

Запись (4.172) означает, что наложены ограничения на функцию, а также ее первую и вторую производные.

В случае дискретизации задачи (см. (4.166)–(4.168), (4.170)) добавляются ограничения в виде неравенств (ср. [10, с. 74]):

$$\left. \begin{aligned} y_{j+1} - y_j &\leq 0, \quad j=1, 2, \dots, n-1, \\ y_{j-1} - 2y_j + y_{j+1} &\leq 0, \quad j=2, 3, \dots, n-1, \\ y_1 - C_2 &\leq 0, \quad C_1 - y_n \leq 0. \end{aligned} \right\} \quad (4.173)$$

В случае дискретизации задачи множество $\hat{Y} \downarrow_{C_1}^{C_2}$ (4.172) переходит в множество $\hat{M} \downarrow_{C_1}^{C_2}$ (ср. [10, с. 73–74]), определяемое ограничениями (4.173).

Решение на множестве ограниченных выпуклых функций. Это – 3-й пример компакта. Пусть искомое точное решение \bar{y} принадлежит множеству *ограниченных* снизу и сверху константами C_1 и

C_2 выпуклых функций. Обозначим это множество через $\widehat{Y}_{C_1}^{C_2}$ (ср. [10, с. 72]), т.е.

$$\widehat{Y}_{C_1}^{C_2} = \{y: y''(s) \leq 0, C_1 \leq y(s) \leq C_2, a \leq s \leq b\}. \quad (4.174)$$

Запись (4.174) означает, что наложены ограничения на функцию, а также ее вторую производную.

В случае дискретизации задачи (см. (4.166)–(4.168), (4.170)) добавляются ограничения в виде неравенств (ср. [10, с. 74]):

$$\left. \begin{aligned} y_{j-1} - 2y_j + y_{j+1} &\leq 0, \quad j = 2, 3, \dots, n-1, \\ y_1 - C_2 &\leq 0, \quad C_1 - y_n \leq 0. \end{aligned} \right\} \quad (4.175)$$

В случае дискретизации задачи множество $\widehat{Y}_{C_1}^{C_2}$ (4.174) переходит в множество $\widehat{M}_{C_1}^{C_2}$ (ср. [10, с. 73–74]), определяемое ограничениями (4.175).

Решение на множестве функций ограниченной вариации. Это – 4-й пример компакта. Пусть искомого точного решения \bar{y} принадлежит множеству функций, имеющих, вообще говоря, флуктуации, но суммарная вариация которых не превосходит заданного значения C . Обозначим это множество через V_C (ср. [10, с. 80]).

В случае дискретизации задачи *вариацией* называется сумма [10, с. 89] $|y_2 - y_1| + |y_3 - y_2| + \dots + |y_n - y_{n-1}|$. Поэтому задача решается с использованием соотношений (4.166)–(4.168), (4.170) и с добавлением ограничения в виде неравенства:

$$|y_2 - y_1| + |y_3 - y_2| + \dots + |y_n - y_{n-1}| \leq C. \quad (4.176)$$

Численные примеры. Для реализации изложенных выше вариантов решения на компакте разработаны программы на Фортране PTIGR, PTIGR1, PTIGR2 и др. [10, с. 136–142, 192–195]. С помощью данных программ был решен ряд примеров.

Решалось интегральное уравнение Фредгольма I рода (ср. [10, с. 87])

$$\int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (4.177)$$

где $a = 0$, $b = 1$, $c = -1$, $d = 2$, ядро

$$K(x, s) = \frac{1}{1 + 100(x-s)^2}. \quad (4.178)$$

Сначала точное решение задавалось в виде (1-й пример)

$$\bar{y}(s) = 1 - s^2, \quad a \leq s \leq b. \quad (4.179)$$

На рис. 4.7 представлены точные $y(s)$, $K(x,0)$ и $f(x)$ 1-го примера.

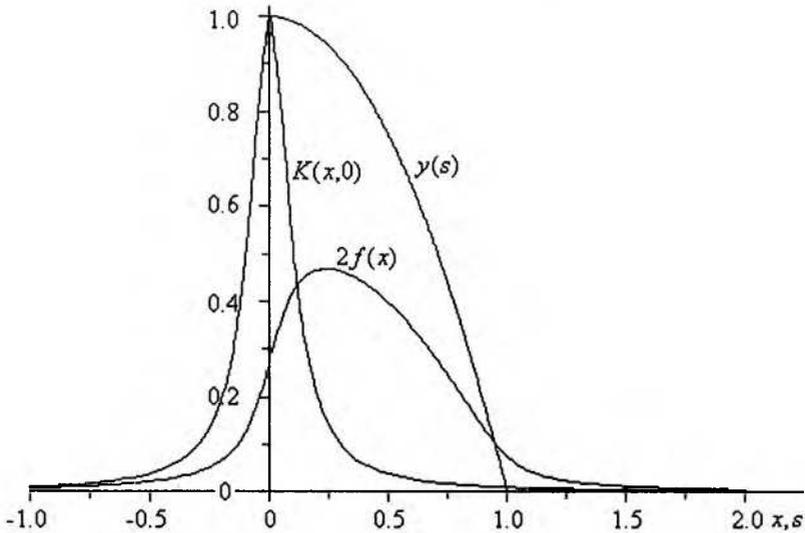


Рис. 4.7. 1-й пример. Точные $y(s)$, $K(x,0)$ и $f(x)$

Характерной особенностью 1-го примера является то, что его решение $y(s)$ является монотонной ограниченной выпуклой функцией при $s \in [a, b]$, т.е. решение в принципе можно искать на множествах $M \downarrow_{C_1}^{C_2}$, $\widehat{M} \downarrow_{C_1}^{C_2}$ или $\widehat{M} \downarrow_{C_1}^{C_2}$ в зависимости от того, какой информацией о решении мы располагаем.

Задача дискретизировалась: число узлов по x положено равным $m = 61$, шаг дискретизации $\Delta x = (d - c)/(m - 1) = 0.05$, число узлов по s взято равным $n = 41$, шаг $\Delta s = (b - c)/(n - 1) = 0.025$.

С помощью программы (датчика случайных чисел) RNDAN [7, с. 152] к правой части $f(x)$ были добавлены погрешности, равномерно распределенные на интервале $[-\delta_0, \delta_0]$, где $\delta_0 = 0.01 \cdot \max_{c \leq x \leq d} f(x) = 0.2325 \cdot 10^{-2}$. Это дает [7, с. 148] $\delta^2 = (d - c) \cdot \delta_0^2 / 3 = 7.21 \cdot 10^{-6}$.

На рис. 4.8 представлены точное решение $\bar{y}(s)$ и приближенное решение $\tilde{y}(s)$, найденное на множестве $M \downarrow_{C_1}^{C_2}$ монотонно невозрастающих ограниченных функций (см. (4.171)) с использованием условия (4.170). При этом $C_1 = 0$, $C_2 = 1$. Начальное приближение положено равным $y_0(s) = 0$. Выполнено 800 итераций, в результате чего

функционал невязки $\Phi(y)$ достиг значения $7.30 \cdot 10^{-6}$, т.е. практически значения δ^2 .

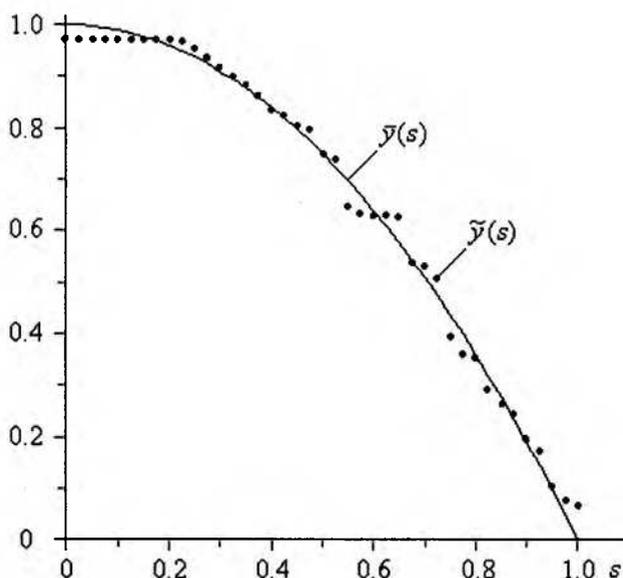


Рис. 4.8. Точное решение $\bar{y}(s)$ и приближенное решение $\tilde{y}(s)$ на множестве $M \downarrow_{C_1}^{C_2}$ монотонно невозрастающих ограниченных функций

На рис. 4.9 и 4.10 представлены результаты поиска решения соответственно на множестве $\tilde{M} \downarrow_{C_1}^{C_2}$ монотонно невозрастающих ограниченных выпуклых функций и на множестве $\tilde{M}_{C_1}^{C_2}$ ограниченных выпуклых функций. При этом решался тот же пример (4.177)–(4.179), однако с большей погрешностью: $\delta_0 = 0.03 \cdot \max_{c \leq x \leq d} f(x) = 0.6975 \cdot 10^{-2}$, $\delta^2 = 6.49 \cdot 10^{-5}$. Выполнено по 800 итераций. Использовано нулевое начальное приближение: $y_0(s) = 0$.

Рисунки 4.8–4.10 демонстрируют, во-первых, что использование дополнительных условий в виде ограничений на решение, помещающих его в компакт (если таковые ограничения следуют из существа задачи), обеспечивают устойчивость решения. Во-вторых, как показывает сравнение рисунков, точность восстановления решения зависит от типа ограничений на решение ((4.169), (4.173), (4.175) и т.д.).

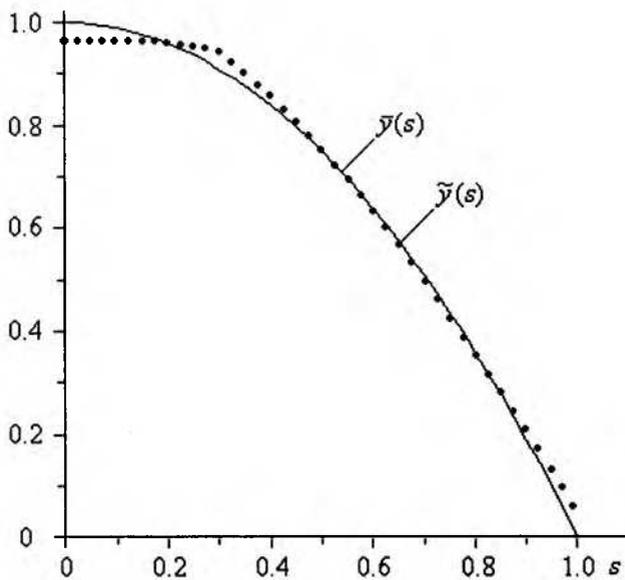


Рис. 4.9. Точное решение $\bar{y}(s)$ и приближенное решение $\tilde{y}(s)$ на множестве $\bar{M}_{C_1}^{C_2}$ монотонно невозрастающих ограниченных выпуклых функций

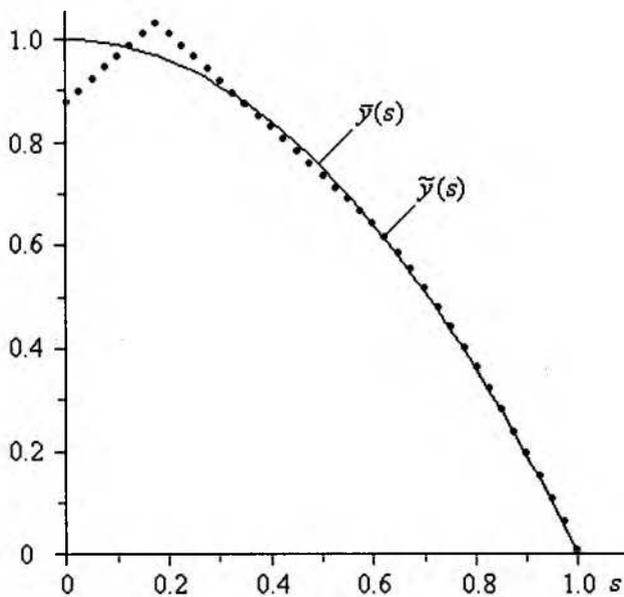


Рис. 4.10. Точное решение $\bar{y}(s)$ и приближенное решение $\tilde{y}(s)$ на множестве $\bar{M}_{C_1}^{C_2}$ ограниченных выпуклых функций

Решался также 2-й пример, в котором точное решение равно $\bar{y}(s) = 4s(1-s)$, $s \in [0,1]$ [10, с. 89]. На рис. 4.11 представлены точные решение, правая часть и ядро 2-го примера.

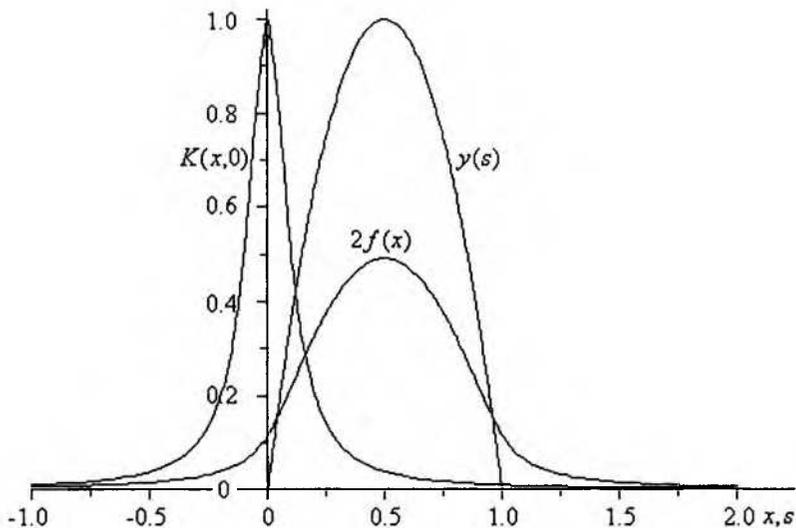


Рис. 4.11. 2-й пример.
Точные решение $y(s)$, правая часть $f(x)$ и ядро $K(x,0)$

В этом примере $\delta_0 = 0.01 \cdot \max f(x) = 0.2457 \cdot 10^{-2}$, $\delta^2 = 8.05 \cdot 10^{-6}$.

Характерной особенностью этого примера является то, что в искомого решение $y(s)$ при $s \in [0,1]$ является ограниченной выпуклой функцией.

На рис. 4.12 представлено точное и приближенное решение 2-го примера, выполненное на множестве $\bar{M}_{C_1}^{C_2}$ ограниченных выпуклых функций. Выполнено 800 итераций, в результате чего функционал невязки $\Phi(y)$ достиг значения $9.99 \cdot 10^{-6}$.

В [10, с. 92, 99] приведены также результаты решения численного примера на множестве функций с ограниченной вариацией V_C .

Данные примеры показывают, что использование дополнительной информации о решении позволяет сделать задачу нахождения решения корректной.

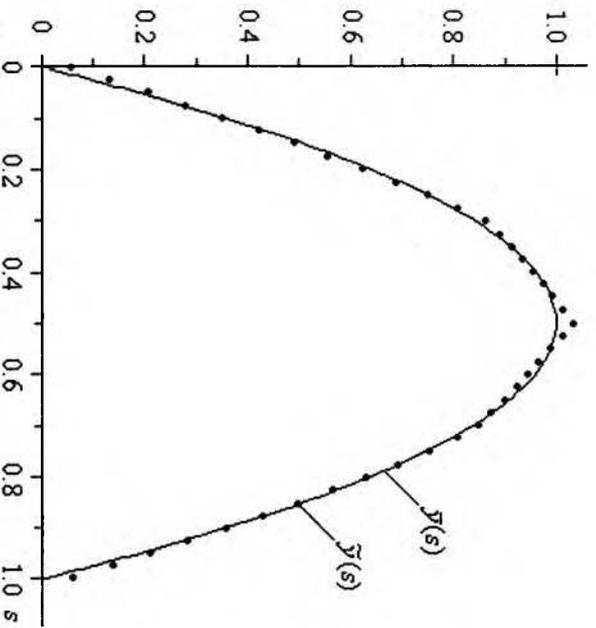


Рис. 4.12. Точное решение $\bar{u}(s)$ и приближенное решение $\bar{y}(s)$ на множестве $\bar{M}_{C_1}^{C_2}$ ограниченных выпуклых функций

Глава 5

ОБРАТНЫЕ ЗАДАЧИ РЕКОНСТРУКЦИИ ИЗОБРАЖЕНИЙ И ТОМОГРАФИИ

В этой главе рассмотрены некоторые некорректные обратные прикладные задачи. Дана их постановка, сделано их математическое описание, применен для их устойчивого решения метод регуляризации Тихонова и приведены результаты компьютерного решения численных примеров.

5.1. Реконструкция смазанных изображений

В п. 5.1 и 5.2 рассмотрим одну из обратных задач оптики – задачу реконструкции (восстановления) искаженных изображений, а именно, смазанных и дефокусированных изображений [1, 2, 7].

Будем под *изображением* подразумевать фотоснимок человека, текста, объекта природы (в том числе, снимок, сделанный из космоса), теле- и киноизображение, телескопический снимок космического объекта и т.д. Однако для определенности под изображением будем далее подразумевать *фотоснимок*.

Полагаем, что выполнена предварительная обработка изображения, а именно, устранены царапины на снимке, подобрана его яркость и контрастность и т.п. – операции, не требующие математической обработки. А мы остановимся на наиболее сложной задаче – на математической реконструкции изображений, искаженных в результате смаза (сдвига, смещения) фотоаппарата или объекта.

Постановка задачи. Рассмотрим данную задачу на примере смазанного (сдвинутого, смещенного) фотоснимка [1, 2, 7, 40, 43]. Пусть фотографируемый объект (полагаемый плоским вследствие его удаленности) и фотопленка фотоаппарата расположены параллельно апертуре тонкой линзы фотоаппарата по разные стороны от линзы на расстояниях соответственно f_1 и f_2 от нее, причем

$$\frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f}, \quad (5.1)$$

где f – фокусное расстояние линзы и $f_1 \geq f$ (см. рис. 5.1). В результате на фотопленке возникнет перевернутое изображение.

Введем в плоскости объекта прямоугольную систему координат $\xi'O'\eta'$, а в плоскости фотопленки $\xi O\eta$. Возьмем на объекте некото-

рую точку $A'(\xi', \eta')$ с интенсивностью излучения $w'(\xi', \eta')$. Лучи, исходящие из нее и прошедшие через линзу, пересекутся в некоторой точке $A(\xi, \eta)$ в плоскости фотопленки. Из подобия прямоугольных треугольников $A'CO'$ и ACO следует:

$$\frac{\overline{O'A'}}{f_1} = \frac{\overline{OA}}{f_2}$$

или в проекциях:

$$\frac{\xi'}{f_1} = -\frac{\xi}{f_2}, \quad \frac{\eta'}{f_1} = -\frac{\eta}{f_2}. \quad (5.2)$$

В результате точка $A(\xi, \eta)$ на фотопленке (соответствующая точке $A'(\xi', \eta')$ на объекте) будет иметь интенсивность w и координаты ξ, η (прямая задача):

$$w(\xi, \eta) = w'(\xi', \eta'), \quad \xi = -\xi'/q, \quad \eta = -\eta'/q, \quad (5.3)$$

где $q = f_1 / f_2$, причем f_2 определяется как

$$f_2 = \left(\frac{1}{f} - \frac{1}{f_1} \right)^{-1}. \quad (5.4)$$

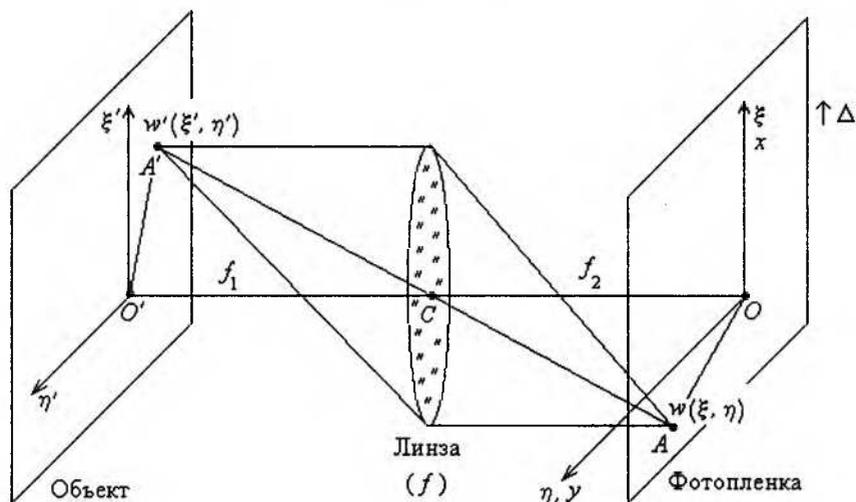


Рис. 5.1. Схема получения смазанного изображения

Таким образом, каждой точке A' на объекте будет соответствовать точка A на фотопленке с той же интенсивностью $w(\xi, \eta) = w'(\xi', \eta')$, но с уменьшенными в $q = f_1 / f_2$ раз (и перевернутыми) координатами

(см. (5.3)).

Пример. Пусть $f_1 = 7$ м, $f = 5$ см, тогда $f_2 = 5.03$ см (согласно (5.4)), а $q = 139.2$, т.е. изображение будет в 139.2 раз меньше объекта (и перевернутым).

По фотоснимку можно восстановить объект (*обратная задача*):

$$w'(\xi', \eta') = w(\xi, \eta), \quad \xi' = -q\xi, \quad \eta' = -q\eta. \quad (5.5)$$

При этом

$$f_1 = \left(\frac{1}{f} - \frac{1}{f_2} \right)^{-1}. \quad (5.6)$$

Далее полагаем, что за время экспозиции τ фотоаппарат совершил прямолинейный и равномерный *сдвиг* (смещение) со скоростью $v = \text{const}$ вдоль оси ξ , т.е. на величину $\Delta = v\tau$ или сдвиг совершил объект (например, быстролетящая цель) на величину $-q\Delta$. В результате изображение на фотопленке будет смазанным (сдвинутым, смещенным) вдоль ξ (см. далее рис. 5.2а).

Вывод интегрального уравнения. Опишем математически задачу смазывания изображения. Введем наряду с неподвижной системой координат $\xi O\eta$ систему координат $x O_y$, связанную с движущейся фотопленкой и совпадающую с $\xi O\eta$ лишь в начальный момент ($\tau = 0$) процесса смазывания (см. рис. 5.1). На некоторую точку (x, y) фотопленки за время τ спроектируется непрерывный набор точек A с абсциссами от $\xi = x$ до $\xi = x + \Delta$ и с различными интенсивностями $w(\xi, y)$, т.е. результирующая интенсивность (обозначим ее через g) в некоторой точке (x, y) фотопленки будет равна сумме (точнее, интегралу) интенсивностей $w(\xi, y)$, $\xi \in [x, x + \Delta]$ (подробности см. в [7, с. 66]):

$$g(x, y) = \frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi. \quad (5.7)$$

Запишем (5.7) иначе:

$$\frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi = g(x, y). \quad (5.8)$$

Соотношение (5.8) является основным в задаче реконструкции смазанных изображений. В нем $g(x, y)$ – распределение интенсивности на фотопленке (на смазанном изображении) в функции прямоугольных координат x, y , причем ось x направлена вдоль сдвига (смаза), Δ – величина смаза, полагаемая известной, а $w(\xi, y)$ – распределение ис-

тинной неискаженной интенсивности на фотопленке (той интенсивности, которая была бы на фотопленке в отсутствие сдвига, т.е. при $\Delta=0$).

Соотношение (5.8) есть одномерное интегральное уравнение Вольтерра I рода относительно $w(\xi, y)$ при каждом фиксированном значении y , играющем роль параметра, другими словами, (5.8) есть совокупность одномерных интегральных уравнений.

В п. 4.4 это уравнение уже кратко рассмотрено (пример 3). По терминологии [12], это уравнение является *неклассическим уравнением Вольтерра I рода* (так как оба предела интегрирования переменны). Задача его решения корректна на паре пространств $(C, C^{(1)})$, т.е. когда $w(\xi, y) \in C$, $g(x, y) \in C^{(1)}$, и некорректна на паре (C, C) , когда $w(\xi, y) \in C$, $g(x, y) \in C$.

Отметим, что в ряде работ ([1, 43] и др.) рассматриваются более сложные постановки данной задачи: неравномерный и/или непрямолинейный сдвиг фотоаппарата или объекта, непараллельность плоскостей объекта и фотопленки и т.д.

Решение интегрального уравнения методом преобразования Фурье и регуляризации Тихонова. Уравнение (5.8) можно записать в виде интегрального уравнения Фредгольма I рода типа свертки [1, 7, 40]:

$$\int_{-\infty}^{\infty} k(x-\xi)w(\xi, y)d\xi = g(x, y), \quad -\infty < x, y < \infty, \quad (5.9)$$

где

$$k(x) = \begin{cases} 1/\Delta & \text{при } x \in [-\Delta, 0], \\ 0 & \text{при } x \notin [-\Delta, 0]. \end{cases} \quad (5.10)$$

Задача решения уравнения (5.9) является некорректной. Для его решения используем методы преобразования Фурье и регуляризации Тихонова [1, 4, 7, 10] (см. п. 4.7). Регуляризованное решение имеет вид (ср. (4.123)):

$$w_{\alpha}(\xi, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_{\alpha}(\omega, y) e^{-i\omega\xi} d\omega, \quad (5.11)$$

где регуляризованный Фурье-спектр решения равен

$$W_{\alpha}(\omega, y) = \frac{K(-\omega) G(\omega, y)}{L(\omega) + \alpha M(\omega)}, \quad (5.12)$$

причем (ср. (4.124), (4.125))

$$K(\omega) = \int_{-\infty}^{\infty} k(x) e^{i\omega x} dx = \frac{\sin(\omega\Delta)}{\omega\Delta} + \frac{\cos(\omega\Delta) - 1}{\omega\Delta} i, \quad (5.13)$$

$$G(\omega, y) = \int_{-\infty}^{\infty} g(x, y) e^{i\omega x} dx \quad (5.14)$$

– Фурье-спектры ядра $k(x)$ и правой части $g(x, y)$ уравнения (5.9), регуляризатор $M(\omega)$ может быть выбран, например, в виде $M(\omega) = \omega^2$, $\alpha > 0$ – параметр регуляризации,

$$L(\omega) = |K(\omega)|^2 = K(\omega)K(-\omega) = \text{Re}^2 K(\omega) + \text{Im}^2 K(\omega).$$

Как мы указывали в п. 4.7, для выбора α существует ряд способов, например, способ невязки и обобщенный принцип невязки [10, 34]. Однако для задачи реконструкции изображений, как показали испытания, наиболее эффективен способ подбора [7, с. 71], [40]. Согласно нему, для ряда значений α вычисляется $w_\alpha(\xi, y)$ по формуле (5.11) (с использованием (5.10), (5.12)–(5.14)), выводится на дисплей решение $w_\alpha(\xi, y)$ в графической форме и выбирается значение α , дающее наилучшее восстановление изображения с точки зрения физиологических (но не математических) критериев восприятия. Этот способ аналогичен способу настройки контраста телеизображения (в этом случае α обратно пропорционален контрасту).

Отметим, что величина смаза Δ априори неизвестна и ее можно определить путем подбора или по величине штрихов на искаженном изображении. Что же касается направления смаза (вдоль которого устанавливается ось x), то его можно определить по направлению штрихов на искаженном изображении (см. ниже рис. 5.2а, 5.4а, 5.5а).

Итак, правильно выбрав направление оси x на искаженном фотоснимке (вдоль смаза) и величину смаза Δ , можно, решив уравнение (5.9) (точнее, совокупность уравнений), используя формулы (5.10)–(5.14) и выбрав α , например, способом подбора, восстановить устойчивым образом неискаженную интенсивность на фотоснимке $w_\alpha(x, y)$ по интенсивности на искаженном снимке $g(x, y)$. При этом на искаженном фотоснимке может быть заложена ценная, но непосредственно нераспознаваемая информация: фотография знатного человека, исторического здания, важного текста, земного объекта, снятого из космоса, быстролетящей цели и т.д. И только математическим путем (с использованием компьютера) можно будет извлечь эту информацию.

Численные результаты. Разработан пакет программ IMAGE [14] (под Windows 98) на языке Visual C++ для решения задачи реконструкции смазанных (а также дефокусированных, см. п. 5.2) изображений методами регуляризации Тихонова и преобразования Фурье согласно формулам (5.9)–(5.14) с подбором α (а также Δ) и выводом

результатов обработки на дисплей. При этом решается как *прямая задача* – моделирование интенсивности на искаженном изображении $g(x, y)$ согласно (5.6), так и *обратная задача* – восстановление неискаженной интенсивности на изображении $w_\alpha(\xi, y)$ согласно (5.11).

Вычисление одномерного обратного и прямого преобразования Фурье (см. (5.11), (5.14)) выполняется в виде дискретного преобразования Фурье (ДПФ). ДПФ реализуется как в виде быстрого преобразования Фурье (БПФ), так и в виде обычного ДПФ. БПФ используется в случае, когда требуется быстрая реконструкция изображения, например, быстро летящей цели, а обычное ДПФ может быть использовано в случае, когда высокая скорость обработки не обязательна, например, при реконструкции текстов, старых фотографий и т.п.

При этом при обработке черно-белых изображений для получения большой градации яркостей использован *серый цвет* (смещение красного, зеленого и синего цветов в одинаковой пропорции), а при обработке цветных изображений использована раздельная обработка в трех цветах и последующее наложение трех изображений. Подробности – в статье [40].

На рис. 5.2а представлено изображение текста в сером цвете, смазанное под некоторым углом по отношению к строкам текста (величина смаза $\Delta = 20$ пикс). Видно, что нетрудно определить направление смаза по штрихам и оценить величину смаза Δ . На рис. 5.2б,в,г представлено регуляризованное решение $w_\alpha(\xi, y)$ соответственно при $\alpha = 0$ (т.е. без регуляризации), при наилучшем подобранном $\alpha = 2.5 \cdot 10^{-3}$ и при завышенном $\alpha = 0.1$. При этом использовано 256 дискретных значений y , а вдоль x при каждом y взято также по 256 дискретных отсчетов g , причем если отсчетов меньше, чем 256, то делалось добавление нулями. Другими словами, использовался алгоритм БПФ, хотя в данном случае более целесообразен алгоритм обычного ДПФ.

На рис. 5.3 приведены результаты реконструкции смазанного изображения неподвижного объекта (здания с решеткой), а на рис. 5.4 и 5.5 – быстро летящей воздушной цели (самолета) и быстро движущейся наземной цели (автомобиля с пассажиром). Все три изображения являются цветными, поэтому их обработка велась раздельно в трех цветах с последующим наложением трех изображений.



а) Смазанное изображение ($\Delta=20$)



б) Восстановленное изображение при $\alpha=0$

finite sum for differ
a general polynomial
expressed as
 $\alpha_1 r^n + \alpha_2 r^{n-1}$
In this context
exponent of the r



в) Восстановленное изображение при $\alpha=2.5 \cdot 10^{-3}$

г) Восстановленное изображение при $\alpha=0.1$

Рис. 5.2. Восстановление смазанного изображения текста в сером цвете

На рисунках 5.3–5.5 даны черно-белые воспроизведения цветных изображений. В данных примерах число дискретных отсчетов вдоль осей x и ω взято равным по 700 пикс. Вдоль же оси y число дискретных отсчетов взято равным 500 пикс. При этом слева и справа от изображений добавлено по 100 нулевых отсчетов для снижения эффекта Гиббса на реконструированных изображениях. В результате общее число отсчетов вдоль x равняется 900 пикс.

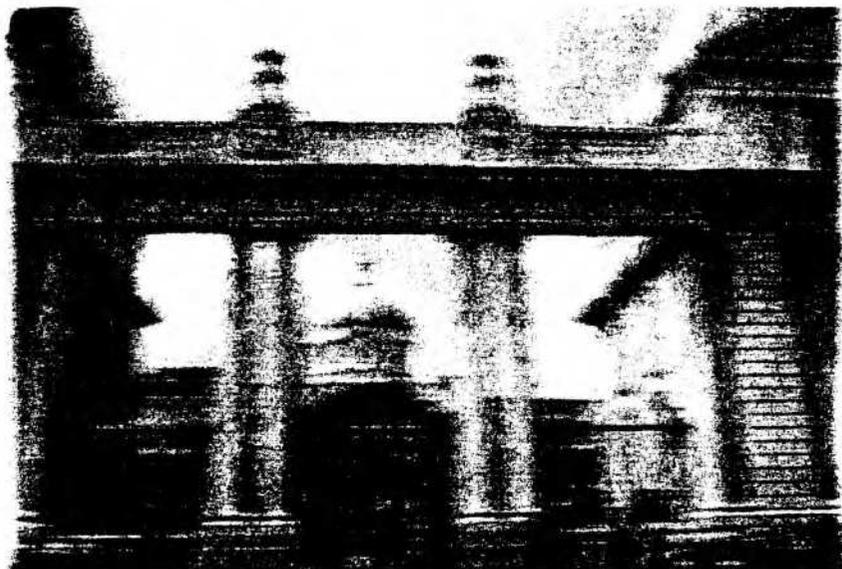


Рис. 5.3а.. Смазанное изображение неподвижной цели

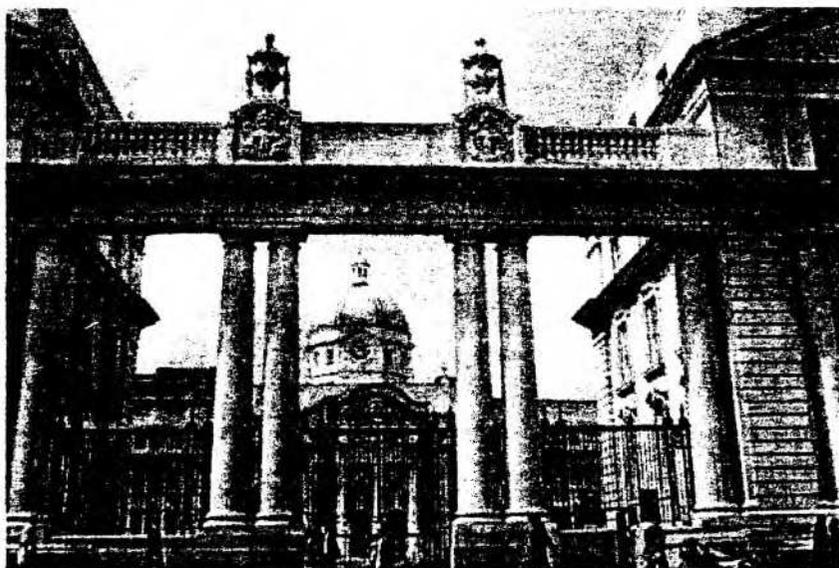


Рис. 5.3б. Реконструированное изображение неподвижной цели



Рис. 5.4а. Смазанное изображение воздушной быстролетающей цели



Рис. 5.4б. Реконструированное изображение воздушной быстролетающей цели

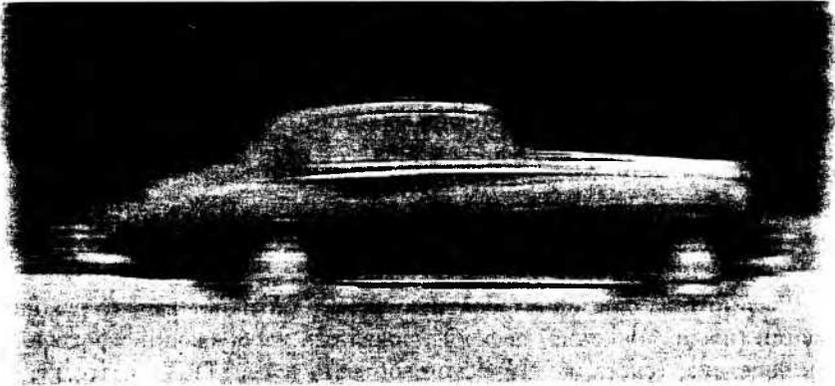


Рис. 5.5а.

Смазанное изображение наземной быстро движущейся цели

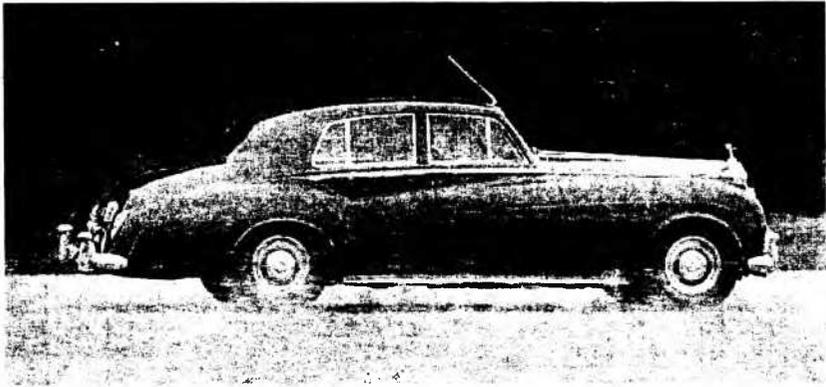


Рис. 5.5б.

Реконструированное изображение наземной быстро движущейся цели

Результаты, приведенные на рис. 5.2–5.5, говорят о следующем. Решение уравнения (5.9) методом преобразования Фурье без регуляризации дает исключительно неустойчивое решение (рис. 5.2б). Применение же метода регуляризации Тихонова при удачно выбранном параметре регуляризации α дает достаточно хороший результат: восстанавливается текст (рис. 5.2в), восстанавливаются изображения неподвижной цели (рис. 5.3б) и подвижных целей (рис. 5.4б и 5.5б) вплоть до мелких деталей (люди, решетка, часы на рис. 5.3б, опознавательные знаки самолета на рис. 5.4б, лицо пассажира в автомобиле на рис. 5.5б), хотя и с погрешностями (шум на рис. 5.2в, эффект Гиббса на фоновой части рисунка 5.4б). При этом, как показал опыт реконструкции

большого числа изображений, параметр регуляризации α довольно надежно и удобно выбирается способом подбора. Что же касается определения направления смаза и его величины Δ , то в одних случаях, как на рис. 5.2а, это можно оценить по штрихам, а в других случаях, как на рис. 5.3а, 5.4а, 5.5а, это можно оценить по размытости левого и правого краев смазанного изображения.

5.2. Реконструкция дефокусированных изображений

Рассмотрим еще одну обратную задачу оптики – задачу реконструкции (восстановления) дефокусированных изображений (фотоснимков человека, текста, космического объекта и т.д.) [1, 2, 7, 16, 17, 40, 43, 49]. Эту задачу рассмотрим на примере дефокусированного фотоснимка.

Постановка задачи. Данная задача имеет много общего с предыдущей задачей (реконструкция смазанных изображений), но имеет и существенные отличия.

Пусть снимаемый объект (полагаемый плоским) и фотопленка расположены параллельно тонкой линзе (по разные стороны от нее) на расстояниях от линзы f_1 и $f_2 + \delta$ соответственно, где δ – погрешность фокусировки изображения (см. рис. 5.6).

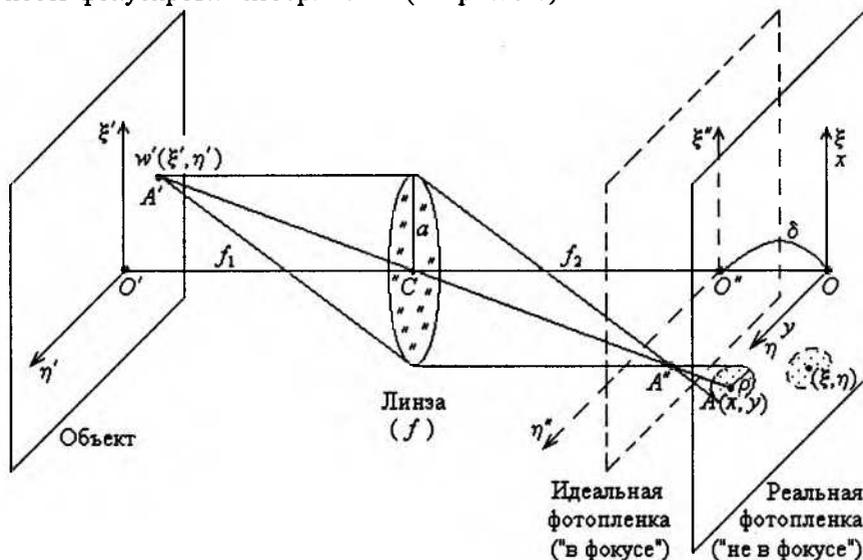


Рис. 5.6. Схема получения дефокусированного изображения

При этом, как и в предыдущей задаче, имеет место соотношение (5.1), где f – фокусное расстояние линзы.

Введем в плоскости объекта прямоугольную систему координат $\xi'O'\eta'$, на «идеальной» фотопленке, расположенной «в фокусе» ($\delta = 0$), – систему координат $\xi''O''\eta''$, а на реальной фотопленке, расположенной «не в фокусе» ($\delta \neq 0$), – систему координат $\xi O\eta$, а также совпадающую с ней xOy . Обозначим через $w'(\xi', \eta')$ интенсивность, исходящую из некоторой точки $A'(\xi', \eta')$ объекта. Точка A' отобразится на «идеальной» фотопленке также в точку A'' с интенсивностью $w''(\xi'', \eta'') = w'(\xi', \eta')$ и с координатами $\xi'' = -\xi'/q$, $\eta'' = -\eta'/q$, где $q = f_1 / f_2$ (ср. (5.2)), причем f_2 определяется согласно (5.3).

На реальной же фотопленке точка A' отобразится не в точку, а в дифракционный круг радиуса

$$\rho = \frac{a\delta}{f_2} \quad (5.15)$$

с центром в точке $A(x, y)$, где a – радиус апертуры линзы, а

$$x = -\frac{f_2 + \delta}{f_1} \xi', \quad y = -\frac{f_2 + \delta}{f_1} \eta' \quad (5.16)$$

(ср. (5.2)).

Вывод основного уравнения. Опишем математически задачу дефокусировки изображения. Рассмотрим, помимо дифракционного круга с центром в точке $A(x, y)$, также некоторый другой круг с центром в точке (ξ, η) (см. рис. 5.6). Радиусы этих (а также других) кругов одинаковы и равны ρ (см. (5.15)), а площади кругов равны $S = \pi\rho^2$. В результате интенсивность $w(\xi, \eta)$, соответствующая точке (ξ, η) , будет «размазана» по кругу радиуса ρ и площади $S = \pi\rho^2$ с плотностью интенсивности $w(\xi, \eta) / \pi\rho^2$ (в первом приближении постоянной в пределах дифракционного круга).

Интенсивность в точке $A(x, y)$ будет результатом суммирования (интегрирования) по все тем кругам, которые накрывают точку $A(x, y)$. Условие накрытия точки $A(x, y)$ кругом с центром в точке (ξ, η) и радиусом ρ есть

$$\sqrt{(x - \xi)^2 + (y - \eta)^2} \leq \rho. \quad (5.17)$$

В результате интенсивность в точке $A(x, y)$ будет равна

$$g(x, y) = \iint_{\sqrt{(x-\xi)^2 + (y-\eta)^2} \leq \rho} \frac{w(\xi, \eta)}{\pi \rho^2} d\xi d\eta. \quad (5.18)$$

Соотношение (5.18) является основным в задаче реконструкции дефокусированных изображений. Запишем его в виде уравнения:

$$\iint_{\sqrt{(x-\xi)^2 + (y-\eta)^2} \leq \rho} \frac{w(\xi, \eta)}{\pi \rho^2} d\xi d\eta = g(x, y). \quad (5.19)$$

Уравнение (5.19) является *основным уравнением* в задаче реконструкции дефокусированных изображений. В нем $g(x, y)$ – измеренная функция, а $w(\xi, \eta)$ – искомая функция.

Приведение основного уравнения к стандартной форме. Уравнение (5.19) есть двумерное интегральное уравнение I рода относительно $w(\xi, \eta)$. Однако оно имеет нестандартную форму, что создает сложности с его решением. Преобразуем его к стандартной форме. Запишем (5.19) в виде [7, с. 75]:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(x-\xi, y-\eta) w(\xi, \eta) d\xi d\eta = g(x, y), \quad -\infty < x, y < \infty, \quad (5.20)$$

где

$$k(x-\xi, y-\eta) = \begin{cases} \frac{1}{\pi \rho^2} & \text{при } \sqrt{(x-\xi)^2 + (y-\eta)^2} \leq \rho, \\ 0 & \text{при } \sqrt{(x-\xi)^2 + (y-\eta)^2} > \rho, \end{cases} \quad (5.21)$$

или

$$k(x, y) = \begin{cases} \frac{1}{\pi \rho^2} & \text{при } \sqrt{x^2 + y^2} \leq \rho, \\ 0 & \text{при } \sqrt{x^2 + y^2} > \rho. \end{cases} \quad (5.22)$$

Соотношение (5.20) есть двумерное интегральное уравнение Фредгольма I рода типа свертки. В нем $g(x, y)$ – интенсивность на дефокусированном фотоснимке, $w(\xi, \eta)$ – искомая интенсивность на неискаженном снимке (на снимке, который был бы получен при $\delta = 0$), $k(x, y)$ есть ядро интегрального уравнения, называемое *функцией рассеяния точки*.

При этом ρ можно определить согласно (5.15), где a и f_2 известны, а δ может быть подобрано. Однако целесообразнее подбирать непосредственно ρ без использования формулы (5.15). На рис. 5.7а приведен пример дефокусированного изображения [7, с. 75], [40], из которого видно, что ρ можно оценить как радиус дифракционных кру-

гов на дефокусированном изображении.

Отметим, что в работе [43] рассмотрена также задача реконструкции дефокусированного изображения для случая непараллельности плоскости объекта и плоскости фотопленки.

После решения уравнения (5.20) можно восстановить исходное изображение в плоскости объекта (*обратная задача*, ср. (5.4)):

$$w'(\xi', \eta') = w(\xi, \eta), \quad \xi' = -\frac{f_1}{f_2 + \delta} \xi, \quad \eta' = -\frac{f_1}{f_2 + \delta} \eta.$$

Решение уравнения методом двумерного ПФ и регуляризации Тихонова. Уравнение (5.20) как двумерное интегральное уравнение Фредгольма I рода типа свертки может быть решено методом двумерного преобразования Фурье (ПФ), или методом инверсной фильтрации, аналогично одномерному уравнению (4.71) или (5.9). Однако задача решения уравнения (5.20), как и уравнения (5.9), является некорректной [1, 4, 16, 19]. Это связано с тем, что функция $g(x, y)$ измеряется с погрешностями и это ведет к сколь угодно большим погрешностям решения $w(\xi, y)$.

Для устойчивого решения задачи используем помимо метода двумерного ПФ еще метод регуляризации Тихонова. Тогда решение уравнения (5.20) методами двумерного ПФ и регуляризации Тихонова запишется в виде (см. (4.129)–(4.133)):

$$w_\alpha(\xi, \eta) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_\alpha(\omega_1, \omega_2) e^{-i(\omega_1 \xi + \omega_2 \eta)} d\omega_1 d\omega_2, \quad (5.23)$$

где

$$W_\alpha(\omega_1, \omega_2) = \frac{K(-\omega_1, -\omega_2) G(\omega_1, \omega_2)}{L(\omega_1, \omega_2) + \alpha M(\omega_1, \omega_2)}, \quad (5.24)$$

причем

$$K(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy, \quad (5.25)$$

$$G(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy, \quad (5.26)$$

$$L(\omega_1, \omega_2) = |K(\omega_1, \omega_2)|^2 = K(\omega_1, \omega_2) K(-\omega_1, -\omega_2), \quad (5.27)$$

$$M(\omega_1, \omega_2) = (\omega_1^2 + \omega_2^2)^{2p}, \quad (5.28)$$

$\alpha > 0$ – параметр регуляризации, $p = 1, 2, 3, \dots$ – порядок регуляризации.

Обычно для выбора параметра регуляризации α используется способ невязки или же обобщенный принцип невязки [10, с. 13]. Однако

для задачи реконструкции дефокусированных изображений (как и для задачи реконструкции смазанных изображений, см. п. 5.1) более эффективен способ подбора [7, с. 77].

Численные результаты. Разработан пакет программ IMAGE на языке Visual C++ для реконструкции смазанных (см. п. 5.1) и дефокусированных изображений. В задаче реконструкции дефокусированных изображений решаются: прямая задача – моделирование дефокусированного изображения и обратная задача – реконструкция исходного изображения по дефокусированному.

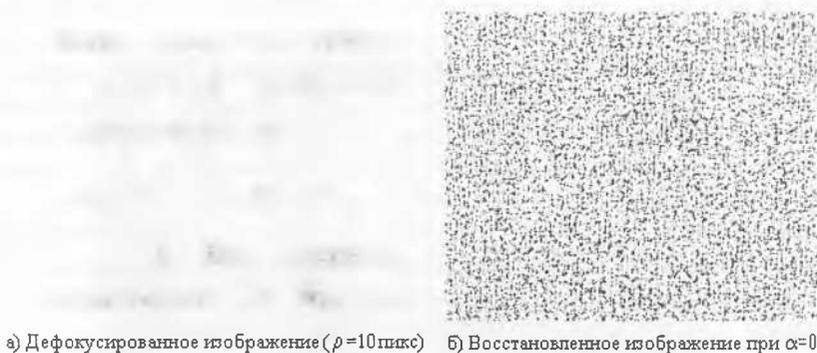
В обратной задаче решается уравнение (5.20) методами двумерного ПФ и регуляризации Тихонова согласно (5.23)–(5.28). При этом двумерные непрерывные ПФ (5.23) и (5.26) вычисляются как двумерные дискретные ПФ, которые в свою очередь реализуются в виде быстрых ПФ по программе FTFTC [10, с. 190]. Величина ρ (радиус дифракционных кругов) полагается оцененной по дефокусированному изображению и, если необходимо, затем уточняется подбором. Параметр регуляризации α выбирается способом подбора с выводом на дисплей решений $w_\alpha(\xi, \eta)$.

При этом, как и в задаче реконструкции смазанных изображений, для получения большой градации яркостей при обработке черно-белых изображений использован серый цвет (смесь красного, зеленого и синего цветов в одинаковой пропорции), а при обработке цветных изображений использована раздельная обработка в трех цветах и последующее наложение трех изображений.

На рис. 5.7 приведены результаты решения модельного примера. На рис. 5.7а – дефокусированное изображение текста в сером цвете (прямая задача). Взято 256×256 дискретных отсчетов. На рис. 5.7б – восстановленное изображение (обратная задача) при $\alpha = 0$, т.е. без регуляризации – неустойчивое решение. На рис. 5.7в – восстановленное изображение при $\alpha = 5 \cdot 10^{-4}$ – наилучшем α , подобранном с помощью чисто визуального анализа дисплейного изображения. А на рис. 5.7г – восстановленное изображение при $\alpha = 5 \cdot 10^{-2}$ – завышенном α . Везде $p = 1$. Данный способ выбора α аналогичен настройке (выбору) контраста телеизображения, когда используются не столько математические, сколько физиологические критерии.

На рис. 5.8 представлен результат восстановления дефокусированного цветного изображения.

Итак, с помощью устойчивых соотношений (5.23)–(5.28), подобрав ρ и α , можно восстановить неискаженный снимок с интенсивностью $w_\alpha(\xi, \eta)$, по дефокусированному снимку с интенсивностью $g(x, y)$.



а) Дефокусированное изображение ($\rho=10$ пикс) б) Восстановленное изображение при $\alpha=0$

finite sum for diff.,
a general polynoi
expressed as

$$a_1 X^n + a_2 X^m$$

In this contex
exponent of the X

в) Восстановленное изображение при $\alpha=5 \cdot 10^{-4}$ г) Восстановленное изображение при $\alpha=5 \cdot 10^{-2}$

Рис. 5.7. Восстановление дефокусированного изображения текста в сером цвете



а) Дефокусированный снимок ($\rho=10$ пикс) б) Восстановленное изображение ($\alpha=5 \cdot 10^{-3}$, $\rho=2$)

Рис. 5.8. Восстановление дефокусированного цветного изображения

В [49] представлены примеры восстановления изображений космических объектов (планет, туманностей, галактик и т.д.) по их дефокусированным изображениям, полученным с помощью космического телескопа «Хаббл» (КТХ). Дефокусирование было обусловлено ошибкой кривизны зеркала КТХ при его изготовлении. Применение же математической (и компьютерной) обработки космических фотоснимков позволило восстановить неискаженные изображения объектов, т.е. компенсировать техническую ошибку измерительного устройства. Правда, при этом был использован, помимо метода двумерного ПФ, не метод типа метода регуляризации Тихонова, а метод типа усечения спектра частот Фурье и сглаживающих окон [4, с. 259–260].

5.3. Задача рентгеновской томографии

В этом пункте рассмотрим некоторые задачи рентгеновской томографии (РТ) [7, с. 17–33], [9], [11, т. 1], [35, 42, 46], а именно, ее математическую часть, связанную с некорректным интегральным уравнением. Заметим, что существуют также ядерно-магнитно-резонансная (ЯМР-) томография [7, с. 33–62], [11, т. 2], [13, 20, 33, 51], ультразвуковая томография [11, т. 2], позитрон-эмиссионная томография и т.д.

Слово «томография» происходит от греческих слов $\tau\omicron\mu\eta$ – сечение и $\gamma\rho\alpha\phi\omega$ – пишу, т.е. «пишу по сечениям». *Суть* всех типов томографии одина: по суммарной информации, полученной от некоторого сечения вещества, нужно определить локальную информацию, а именно, распределение плотности вещества по сечению $c(x, y)$, где x, y – координаты в сечении, а затем по плотностям $c_z(x, y)$ в ряде сечений, где z – координата, перпендикулярная сечению, получить (сконструировать) объемную плотность $c(x, y, z)$. В разных типах томографии суммарная информация качественно различна: в РТ это интенсивность на детекторах (приемниках), в ЯМР-томографии это эхо-сигналы и т.д. Различны также математические описания: в РТ задача описывается интегральным уравнением Радона или Фредгольма, в ЯМР-томографии – двумерным преобразованием Фурье и т.д. Однако *конечная цель* одина – получение распределения плотности $c(x, y)$ по сечениям. Поэтому, например, рентгеновские и ЯМР-томограммы (отображение на прозрачных пленках плотности $c(x, y)$) выглядят практически одинаково, хотя физика, техника и математика, требуемая для их получения, различны. На рис. 5.9 приведен пример ЯМР-томограмм.

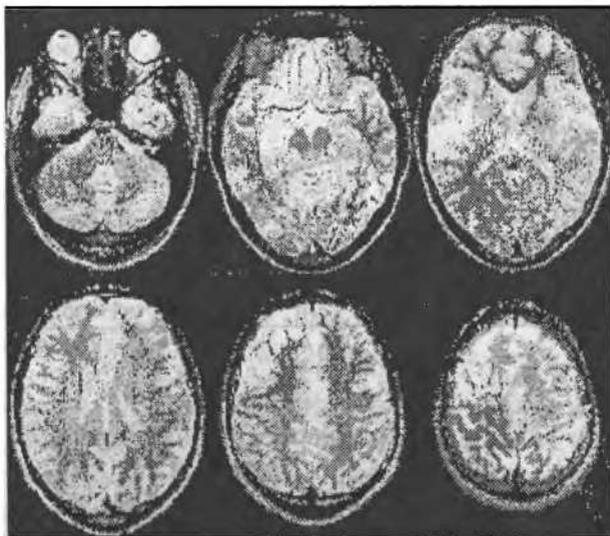


Рис. 5.9. Томограммы шести горизонтальных сечений головного мозга

Постановка эксперимента. На рис. 5.10 представлена схема рентгеновского томографа 2-го поколения (существуют рентгеновские томографы 1–5 поколений [7, с. 24–26], [11, т. 1, с. 142–146]).

На раме помещен ряд рентгеновских трубок (источников), излучающих остронаправленные параллельные рентгеновские лучи. Лучи проходят через сечение объекта (например, мозга) и их ослабленные интенсивности фиксируются соответствующими детекторами (приемниками). Затем вся рама с источниками и приемниками поворачивается на новый угол θ и эксперимент (излучение и прием) повторяется.

Закон Бера. Согласно закону Бера [7, с. 18], [11, т. 1, с. 138], интенсивность рентгеновского луча, принятого детектором, равна

$$I(l, \theta) = I_0 \cdot \exp \left[- \int_{L(l, \theta)} c(x, y) ds \right], \quad (5.29)$$

где l – координата детектора, θ – угол поворота рамы, I_0 – интенсивность излучающей трубки, $c(x, y)$ – плотность вещества на луче $L(l, \theta)$, являющемся прямой, уравнение которой:

$$x \cos \theta + y \sin \theta = l. \quad (5.30)$$

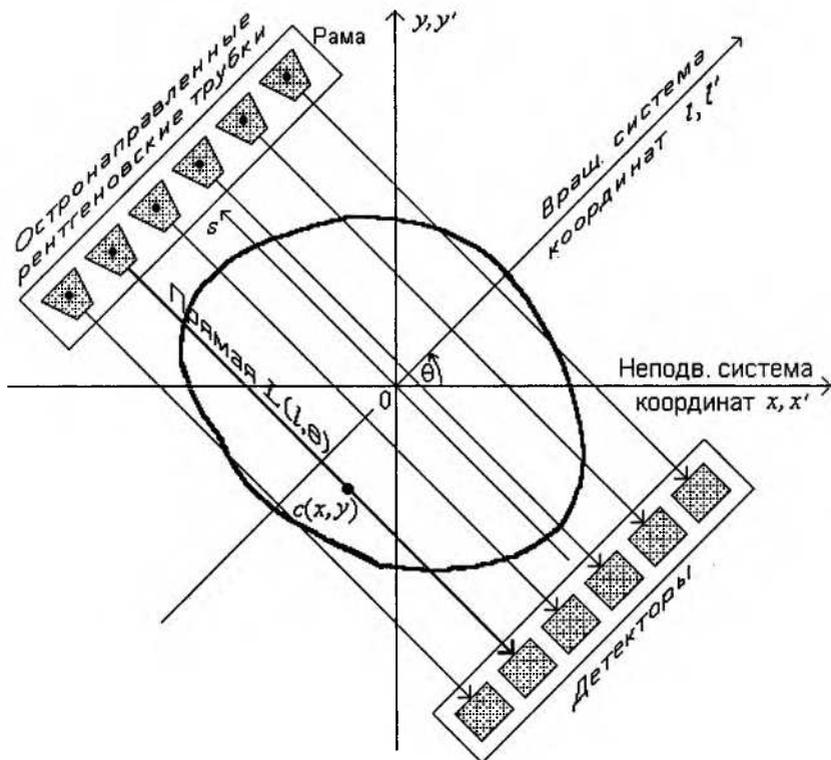


Рис. 5.10. Постановка эксперимента в РТ
(на примере рентгеновского томографа 2-го поколения)

Интегрирование в (5.29) ведется по лучу $L(l, \theta)$. Интеграл $\int_{L(l, \theta)} c(x, y) ds$ есть масса вещества на луче. Чем больше значение этого интеграла, тем меньше принятая интенсивность $I(l, \theta)$.

Преобразование и уравнение Радона. Запишем (5.29) иначе:

$$I(l, \theta) / I_0 = \exp[-q(l, \theta)], \quad (5.31)$$

где

$$q(l, \theta) = \int_{L(l, \theta)} c(x, y) ds. \quad (5.32)$$

Определение. Выражение (5.32), где $L(l, \theta)$ есть некоторый луч, проходящий через вещество, $c(x, y)$ – плотность вещества на этом луче, а координата s направлена вдоль этого луча, называется *преобразованием Радона* (1917 г.).

Логарифмируя (5.31), получим:

$$q(l, \theta) = -\ln[I(l, \theta) / I_0]. \quad (5.33)$$

Определение. Функция $q(l, \theta)$ называется *поглощением*.

Поглощение $q(l, \theta)$ может принимать значения от 0, когда $I(l, \theta) = I_0$, т.е. среда прозрачна, до ∞ , когда $I(l, \theta) = 0$, т.е. среда абсолютно непрозрачна.

Определение. Отношение $I(l, \theta) / I_0$ называется *прозрачностью*.

Прозрачность $I(l, \theta) / I_0$ может принимать значения от 0 (среда абсолютно непрозрачна) до 1 (среда прозрачна).

Запишем (5.32) иначе:

$$\int_{L(l, \theta)} c(x, y) ds = q(l, \theta). \quad (5.34)$$

Соотношение (5.34) называется *уравнением Радона*. Оно является основным уравнением в РТ. Здесь двухмерная функция $q(l, \theta)$ определяется согласно (5.33), где $I(l, \theta)$ получается в результате измерений. А двухмерная функция $c(x, y)$ является искомой. Поэтому (5.34) можно рассматривать как *интегральное уравнение* относительно $c(x, y)$ по измеренной правой части $q(l, \theta)$. Решение уравнения (5.34) позволяет в принципе найти распределение плотности $c(x, y)$ в некотором сечении объекта (например, мозга) по измеренной $I(l, \theta)$, а значит, и $q(l, \theta)$. Такая задача называется *реконструкцией рентгеновского изображения*.

Впервые задачу реконструкции рентгеновского изображения рассмотрел Радон в 1917 г.

Одно из решений уравнения (5.34) (с помощью преобразования Фурье) имеет вид [46, с. 33]:

$$c(x, y) = \frac{1}{\pi} \int_0^\pi d\theta \int_{-\infty}^{\infty} q(l, \theta) p(x \cos \theta + y \sin \theta - l) dl, \quad (5.35)$$

где

$$p(t) = \pi \int_{-\infty}^{\infty} |\omega| e^{i\omega t} d\omega \quad (5.36)$$

– так называемая *импульсная реакция фильтра* с частотной характеристикой $\pi |\omega|$. Исследуем кратко решение (5.35)–(5.36). Воспользуемся формулой Эйлера $e^{i\omega t} = \cos \omega t + i \sin \omega t$ и тем, что $\int_{-\infty}^{\infty} |\omega| \cos \omega t d\omega = 2 \int_0^{\infty} |\omega| \cos \omega t d\omega$ как интеграл от четной функции и

$\int_{-\infty}^{\infty} |\omega| \sin \omega t d\omega = 0$ как интеграл от нечетной функции. Получим:

$$p(t) = \pi \int_{-\infty}^{\infty} |\omega| \cos \omega t d\omega + i \pi \int_{-\infty}^{\infty} |\omega| \sin \omega t d\omega = 2\pi \int_0^{\infty} \omega \cos \omega t d\omega.$$

Последний интеграл не имеет сходимости при любом $t \in (-\infty, \infty)$ и как следствие, не имеет также сходимости и выражение (5.35) при любых $x, y \in (-\infty, \infty)$. Другими словами, решение (5.35)–(5.36) есть расходящееся решение даже при точной $q(l, \theta)$. Ниже приведен сходящийся и устойчивый вариант этого решения.

Другой вариант решения уравнения (5.34) [9, с. 40]:

$$c(x, y) = -\frac{1}{2\pi^2} \int_0^{\pi} d\theta \int_{-\infty}^{\infty} \frac{\partial q(l, \theta)}{\partial l} \frac{dl}{l - (x \cos \theta + y \sin \theta)}. \quad (5.37)$$

Решение (5.37) является неустойчивым – это связано с необходимостью численного вычисления производной $\partial q(l, \theta) / \partial l$ от обычно зашумленной функции $q(l, \theta)$. Кроме того, интеграл в (5.37) является сингулярным, так как знаменатель $l - (x \cos \theta + y \sin \theta)$ может обращаться в нуль.

Таким образом, задача решения уравнения (5.34) является некорректной.

Перечень устойчивых методов решения уравнения Радона.

Одна из первых устойчивых методик решения уравнения Радона была предложена английской фирмой ЕМІ и реализована Г. Хаунсфилдом в первой коммерческой системе (сканере головного мозга – томографе СТ-1010). Был разработан также метод Шеппа–Логана [35]. Эти и подобные методы (использующие для устойчивости прием усечения спектра частот Фурье, сглаживающие окна и т.п.) принято называть методами «интуитивной регуляризации». Эти методы хотя и обладают некоторой устойчивостью, но все же недостаточно точны (см. дальше рис. 5.11, 5.12). Более эффективный результат получается при применении метода регуляризации Тихонова и еще более точного метода – метода локальной регуляризации Арсенина, использованного в первом отечественном рентгеновском компьютерном томографе СРТ-1000 [42]. Остановимся на использовании метода регуляризации Тихонова для решения задачи реконструкции рентгеновских изображений.

Устойчивое решение задачи методом регуляризации Тихонова.

Сначала остановимся на решении (5.35). Его регуляризованный вариант имеет вид [46, с. 33]:

$$c_{\alpha}(x, y) = \frac{1}{\pi} \int_0^{\pi} d\theta \int_{-\infty}^{\infty} \tilde{q}(l, \theta) p_{\alpha}(x \cos \theta + y \sin \theta - l) dl, \quad (5.38)$$

где $\tilde{q}(l, \theta)$ – функция с погрешностями, а

$$p_{\alpha}(l) = \pi \int_{-\infty}^{\infty} |\omega| W_{\alpha}(|\omega|) e^{i\omega l} d\omega, \quad (5.39)$$

причем (ср. (4.123), (4.122))

$$W_{\alpha}(|\omega|) = \frac{H^2(|\omega|)}{H^2(|\omega|) + \alpha \omega^{2p}}, \quad H(|\omega|) = \frac{1}{\pi |\omega|}, \quad (5.40)$$

$\alpha > 0$ – параметр регуляризации, $p \geq 0$ – порядок регуляризации. Еще проще выражение для $W_{\alpha}(|\omega|)$ может быть записано в виде:

$$W_{\alpha}(|\omega|) = \frac{1}{1 + \alpha \omega^{2p}}.$$

Теперь рассмотрим другой вариант тихоновской регуляризации. Уравнение Радона (5.34) может быть приведено к более привычному типу уравнения, а именно, к *двухмерному интегральному уравнению Фредгольма I рода типа свертки* [9]:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{c(x', y') dx' dy'}{\sqrt{(x-x')^2 + (y-y')^2}} = S(x, y), \quad -\infty < x, y < \infty, \quad (5.41)$$

где

$$S(x, y) = \frac{1}{\pi} \int_0^{\pi} q(x \cos \theta + y \sin \theta, \theta) d\theta. \quad (5.42)$$

Уравнение (5.41) имеет стандартную форму уравнения, методика решения которого хорошо отработана [4, 8, 10]. В нем ядро равно

$$K(x, y) = \frac{1}{\sqrt{x^2 + y^2}}, \quad (5.43)$$

искомая функция есть $c(x, y)$, а правая часть $S(x, y)$ может быть вычислена численно по известной $q(l, \theta)$ согласно (5.42).

В п. 4.7 изложен метод регуляризации Тихонова решения двухмерного интегрального уравнения Фредгольма I рода типа свертки. Согласно нему, регуляризованное решение уравнения (5.41) имеет вид (см. (4.129)–(4.133)):

$$c_{\alpha}(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{c}_{\alpha}(\omega_1, \omega_2) e^{-i(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2, \quad (5.44)$$

где регуляризованный спектр решения равен

$$\hat{c}_\alpha(\omega_1, \omega_2) = \frac{\hat{K}(-\omega_1, -\omega_2) \hat{S}(\omega_1, \omega_2)}{L(\omega_1, \omega_2) + \alpha M(\omega_1, \omega_2)}, \quad (5.45)$$

при этом спектр правой части

$$\hat{S}(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy \quad (5.46)$$

вычисляется численно, а спектр ядра может быть найден аналитически:

$$\hat{K}(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy = \frac{2\pi}{\sqrt{\omega_1^2 + \omega_2^2}} = \frac{2\pi}{\omega}. \quad (5.47)$$

Здесь

$$L(\omega_1, \omega_2) = |\hat{K}(\omega_1, \omega_2)|^2 = \hat{K}(\omega_1, \omega_2) \hat{K}(-\omega_1, -\omega_2) = \frac{4\pi^2}{\omega^2}, \quad (5.48)$$

где $\omega = \sqrt{\omega_1^2 + \omega_2^2}$, $\alpha > 0$ – параметр регуляризации. Регуляризатор $M(\omega_1, \omega_2)$ в работе [9, с. 68] выбран в виде:

$$M(\omega_1, \omega_2) = 4\pi^2 \left[(\omega_1^2 + \omega_2^2)^2 + 1 \right] = 4\pi^2 (\omega^4 + 1). \quad (5.49)$$

В результате регуляризованный спектр решения получается равным [9, с. 68]

$$\hat{c}_\alpha(\omega_1, \omega_2) = \frac{1}{2\pi} \frac{\omega}{1 + \alpha \omega^2 (\omega^4 + 1)} \hat{S}(\omega_1, \omega_2). \quad (5.50)$$

Для выражения (5.50) характерно следующее. Во-первых, $\hat{c}_\alpha(\omega_1, \omega_2) \rightarrow 0$ при $\omega \rightarrow \infty$ и интеграл (5.44) сходится. Во-вторых, подавление высоких частот ω в (5.50) происходит более аккуратно, чем в методах, использующих усечение верхних частот или даже сглаживающие окна [4, с. 259–260]. Дело в том, что с одной стороны, высокие частоты наиболее сильно реагируют на погрешности и поэтому их нужно подавлять, а с другой стороны, высокие частоты нужны для высокого разрешения томограмм (разрешения близких деталей на них), поэтому это подавление должно быть умеренным. Обоим этим критериям (в пропорции, регулируемой параметром α) метод регуляризации Тихонова удовлетворяет.

Численные иллюстрации. Приведем примеры реконструкции рентгеновских изображений, выполненные согласно различных методик. На рис. 5.11 приведены результаты решения модельного примера [42].

На рис. 5.11а представлено решение согласно метода Шеппа–Логана (относимого к методам «интуитивной регуляризации»), а на рис.

5.11б – согласно метода локальной регуляризации Арсенина [9, 42] (метода, еще более точного, чем метод регуляризации Тихонова). Видим, каким неустойчивым, зашумленным является решение методом Шеппа–Логана и каким высокоточным – решение методом локальной регуляризации.

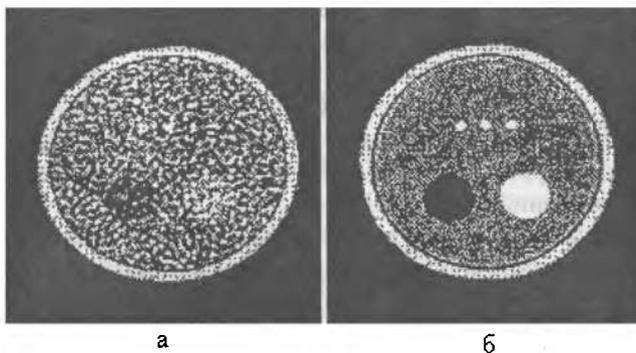


Рис. 5.11. Реконструкция рентгеновского изображения (модельный пример) методом Шеппа–Логана (а) и методом локальной регуляризации Арсенина (б)

На рис. 5.12 приведены томограммы [42] одного и того же сечения головного мозга человека, полученные по методике фирмы ЕМІ (томограф СТ-1010) – см. рис. 5.12а и по методу локальной регуляризации (томограф СРТ-1000) – см. рис. 5.12б. Для этого сечения характерно то, что плотности вещества в соседних участках отличаются лишь на $\approx 0.5\%$ и методика ЕМІ (рис. 5.12а) их не разрешает, в то время как методом локальной регуляризации (рис. 5.12б) они разрешаются.

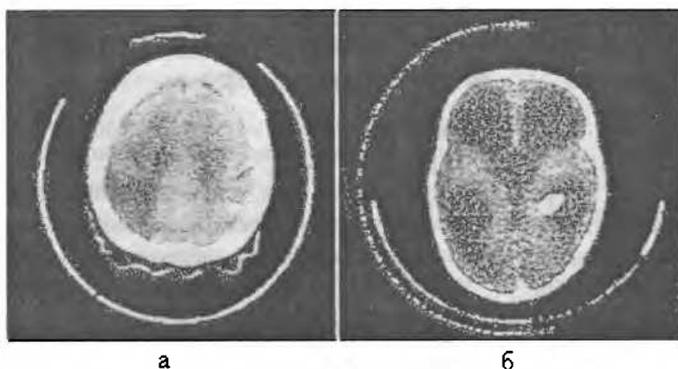


Рис. 5.12. Томограммы сечения мозга по методике фирмы ЕМІ (а) и по методу локальной регуляризации (б)

Приведены примеры также в [9], [11, т. 1, с. 141], [35] и др.

Об алгоритмах и программах. На практике изложенные методы реконструкции рентгеновских изображений реализуются в виде численных алгоритмов и компьютерных программ. Рассмотренные выше уравнения являются интегральными уравнениями типа свертки и их решения выражаются через преобразования Фурье (с регуляризацией). Поэтому значительная часть вычислений сводится к вычислению непрерывных преобразований Фурье (НПФ) прямых и обратных, одно- и двухмерных (5.36), (5.39), (5.44), (5.46), (5.47) и др. В свою очередь, НПФ при практических вычислениях заменяются дискретными преобразованиями Фурье (ДПФ), а ДПФ обычно реализуются в виде быстрых преобразований Фурье (БПФ).

Для компьютерной реализации БПФ разработано много стандартных программ на различных универсальных языках программирования (Fortran, C, Matlab, MathCAD и др.), например, программы FFT [38] и FTF1C [10, с. 183, 190] – программы вычисления одномерного БПФ и FTF2C [10, с. 190] – программа вычисления двухмерного БПФ на языке Fortran.

Разработаны также программы для решения одно- и двухмерных интегральных уравнений Фредгольма I рода типа свертки методом преобразования Фурье и регуляризации. Это – программы RTIKR [10, с. 178], CONV1 и др. [4, с. 379–388] – для решения одномерных уравнений; RTITR [10, с. 185] – для решения двухмерных уравнений и т.д.

Однако в программное обеспечение реальных томографов обычно «зашиваются» программы не на универсальных языках программирования, а например, на Ассемблере, а также используются спецпроцесоры БПФ.

5.4. Задача синтеза магнитного поля в ЯМР-томографе

Среди многочисленных задач ядерно-магнитно-резонансной томографии (ЯМР-томографии) [7, с. 33–62], [9], [11, т. 2, с. 105–231], [13, 20, 31, 32, 33, 52] рассмотрим одну из них, являющуюся некорректной и требующую применения некоторого регулярного метода для ее устойчивого решения. Это – задача синтеза магнитного поля в ЯМР-томографе [7, 13, 20, 31, 32, 39, 50].

Формулировка задачи. В ЯМР-томографии исключительно важной является задача создания высокооднородного магнитного поля внутри ЯМР-томографа. Относительная неоднородность напряженности H основного (статического) поля в пределах рабочего объема томографа должна быть $\Delta H/H \sim 10^{-5} - 10^{-6}$ [11, т. 2], [20, 32, 52]. Это

позволит с высоким разрешением решать основную задачу ЯМР-томографии – задачу реконструкции ЯМР-изображений [7, с. 45–51], [11, т. 2], [52, р. 270–274].

Как правило, задача создания высокооднородного магнитного поля решается путем введения соленоидальных корректирующих катушек Гельмгольца различного порядка [20]. Однако это довольно сложная техническая задача. Рассмотрим другой подход к решению задачи создания высокооднородного поля в ЯМР-томографе. Этот подход основан на расчете такого закона распределения тока вдоль обмотки катушки томографа, что он обеспечит высокую однородность магнитного поля внутри катушки. Данная задача может быть *прямая*, или *задача анализа* (расчет магнитного поля по распределению тока) и *обратная*, или *задача синтеза* (расчет распределения тока по заданному магнитному полю). Обратная задача является значительно более сложной, чем прямая [13, с. 17].

Рассмотрим *задачу синтеза магнитного поля* внутри катушки ЯМР-томографа [7, с. 55–61], [13, с. 17], [32, 39, 50], причем в качестве примера рассмотрим ее простейший вариант – синтез на оси цилиндрической катушки с бесконечно тонкой обмоткой.

Одним из первых задачу синтеза магнитного поля сформулировал с учетом ее некорректности и решил с применением метода регуляризации Тихонова К. Адамиак [50]. Эту задачу рассмотрел также Л. Луганский [31] и др. Однако в работе [50] был допущен ряд неточностей, в частности, при выборе параметра регуляризации α в методе регуляризации Тихонова был использован *способ минимальной невязки*, что привело к очень малому значению $\alpha \sim 10^{-13}$, т.е. по-существу к отказу от регуляризации и, как следствие, к неустойчивому распределению тока. В работе [39] ряд положений работы [50] сформулирован заново, а также выполнено дальнейшее развитие идеи К. Адамиака.

Итак, рассмотрим следующую обратную задачу ЯМР-томографии: определить распределение плотности тока $J(a)$ вдоль бесконечно тонкой обмотки цилиндрической катушки по заданному магнитному полю (напряженности) $H(z)$ на ее оси. Это – *задача интегрального синтеза магнитного поля* на оси катушки ЯМР-томографа [32]. Здесь a – координата вдоль обмотки катушки, z – координата вдоль ее оси. Рассмотрим случай $H(z) = \text{const}$.

На рис. 5.13 представлена цилиндрическая катушка с бесконечно тонкой обмоткой, где $a \in [-l, l]$ – расстояние вдоль обмотки катушки, $z \in [-l, l]$ – расстояние от центра катушки вдоль ее оси, l – полудлина катушки, $H(z) = H = \text{const}$, $z \in [-l, l]$ – заданная напряженность маг-

нитного поля на оси катушки, $J(a)$, $a \in [-l, l]$ – искомое распределение тока вдоль обмотки катушки, R – радиус катушки.

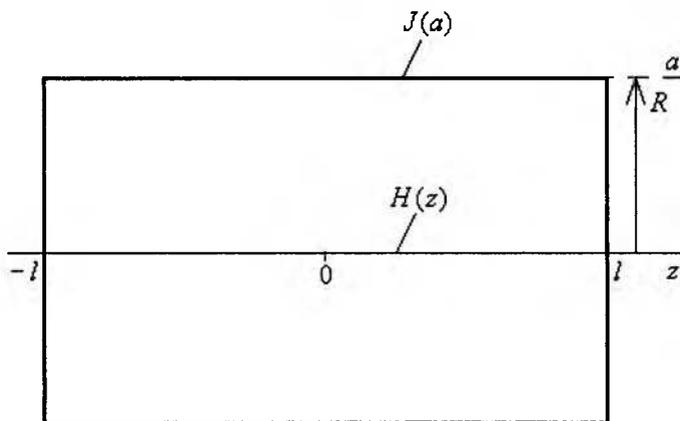


Рис. 5.13. Цилиндрическая катушка с бесконечно тонкой обмоткой

Согласно [33, с. 37], обмотка в случае $J(a) \neq \text{const}$ называется *соленоидом с переменной плотностью тока*, или с *неоднородным распределением тока*. Реализовать неоднородное распределение тока можно, например, путем разбиения соленоида на несколько отдельных секций. Другие способы (создание изолированных витков с индивидуальной подводкой тока через единый источник или создание соленоида с единым, намотанным на цилиндр проводом, но имеющим на каждом витке свое сопротивление) предложены в работе [39].

Вывод интегрального уравнения задачи синтеза. Как известно [13, с. 20], [20], [24, с. 327], [33, с. 23], [48, с. 330–333], [50], напряженность магнитного поля на оси тонкого витка (кругового тока) равна (см. рис. 5.14)

$$H(z) = k \frac{J R^2}{\sqrt{[R^2 + (z - a)^2]^3}}, \quad -\infty < z < \infty,$$

где k – коэффициент пропорциональности, J – ток в витке, R – радиус витка, z – координата точки, в которой вычисляется H , a – z -координата центра витка (ось z направлена перпендикулярно плоскости витка). Для упрощения дальнейших записей будем полагать: $k = 1$.

Тогда

$$H(z) = \frac{J R^2}{\sqrt{[R^2 + (z - a)^2]^3}}, -\infty < z < \infty. \quad (5.51)$$

Далее полагаем, что имеется бесконечное множество витков, намотанных на цилиндр радиуса R и полудлины l и имеющих плотность тока $J = J(a)$ (см. рис. 5.13). В этом случае, интегрируя по виткам с напряженностью (5.51), получим следующее выражение для суммарной напряженности магнитного поля в точке с координатой z на оси цилиндра с бесконечно тонкой обмоткой (см. рис. 5.13):

$$H(z) = \int_{-l}^l \frac{J(a) R^2 da}{\sqrt{[R^2 + (z - a)^2]^3}}, -\infty < z < \infty. \quad (5.52)$$

Вычисление поля $H(z)$ по заданному току $J(a)$ согласно (5.52) есть *прямая задача*. Сделаем важное замечание, что формула (5.52) справедлива как для точек внутри цилиндра ($|z| \leq l$), так и вне его ($|z| \geq l$). Выполним краткий анализ прямой задачи.

Если $J(a) = J = \text{const}$, то интегрирование (5.52) дает:

$$H(z) = \left[\frac{z+l}{\sqrt{R^2 + (z+l)^2}} - \frac{z-l}{\sqrt{R^2 + (z-l)^2}} \right] J, -\infty < z < \infty. \quad (5.53)$$

Частные случаи (5.53): поле в центре катушки $H(0) = 2J/\sqrt{1 + (R/l)^2}$, поле на краю катушки $H(l) = J/\sqrt{1 + (R/2l)^2}$. Если $R=l$, то $H(0) = 2J/\sqrt{2}$, $H(l) = 2J/\sqrt{5}$, т.е. напряженность в центре катушки в $\sqrt{5/2} \approx 1.58$ раз больше, чем на ее краях. Если $z \rightarrow \infty$, то [48, с. 333] $H(z) \rightarrow 2J R^2 l / z^3$, т.е. вне катушки поле $H(z)$ падает как $\sim z^{-3}$. Такая асимптотика справедлива и для тонкого витка (см. (5.51)), а также будет иметь место и для цилиндрической катушки при $J(a) \neq \text{const}$.

Если $l \gg R$ (бесконечно длинная катушка), то $H(0) = 2J$, $H(l) = J$. Видим что в этом случае напряженность в центре катушки вдвое больше, чем на ее краях.

Если, например, $l = 10R$, т.е. длина катушки в 20 раз больше ее радиуса, то $H(0) = 1.9901 J$, $H(R) = 1.9898 J$ и в результате

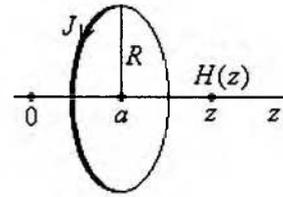


Рис. 5.14. Напряженность на оси тонкого витка

$|H(R) - H(0)|/H(0) \approx 1.5 \cdot 10^{-4}$. Видим, что в случае длинной катушки в принципе можно получить неоднородность поля порядка 10^{-4} в пределах $|z| < R$, однако такую катушку технически трудно изготовить.

Вышеприведенный краткий анализ прямой задачи показывает, что при $J(a) = \text{const}$ поле $H(z)$ падает от центра катушки к ее краям, а вне катушки $H(z) \sim |z|^{-3}$ при $|z| \rightarrow \infty$.

Перейдем к рассмотрению *обратной задачи*. Запишем (5.52) в виде:

$$\int_{-l}^l \frac{R^2}{\sqrt{[R^2 + (z-a)^2]^3}} J(a) da = H(z), \quad -l \leq z \leq l. \quad (5.54)$$

Соотношение (5.54) есть *интегральное уравнение Фредгольма I рода*, где $H(z)$ – задаваемая правая часть (напряженность магнитного поля на оси катушки), например, $H(z) = H = \text{const}$, а $J(a)$ – искомая функция (распределение тока вдоль бесконечно тонкой обмотки цилиндрической катушки ЯМР-томографа). Задача решения уравнения (5.54) является некорректной (см. п. 4.4).

Используем безразмерные переменные: $s = a/R$, $x = z/R$, $s_0 = l/R$. Тогда (5.54) можно записать в виде

$$\int_{-s_0}^{s_0} K(x, s) J(s) ds = H(x), \quad -s_0 \leq x \leq s_0, \quad (5.55)$$

где

$$K(x, s) = \frac{1}{\sqrt{[1 + (x-s)^2]^3}} \quad (5.56)$$

– ядро интегрального уравнения.

Решение уравнения (5.55) и техническая реализация решения позволят в принципе создать на оси катушки ЯМР-томографа поляризуемое поле с заданным законом изменения его напряженности, в частности, $H(x) = H = \text{const}$.

Решение уравнения методом регуляризации с ограничениями. Рассмотрим вопрос о решении уравнения (5.55).

Вначале выполним качественный анализ. Заметим, что запись (5.55)–(5.56) справедлива не только для $x \in [-s_0, s_0]$, но и для $x \notin [-s_0, s_0]$.

Пусть $H(x) = \text{const}$ при $x \in [-s_0, s_0]$. Тогда при $x \notin [-s_0, s_0]$ поле $H(x) \neq \text{const}$ и будет монотонно падать с ростом $|x|$, переходя при

$|x| \rightarrow \infty$ в асимптотику $H(x) \sim |x|^{-3}$ (см. рис. 5.15).

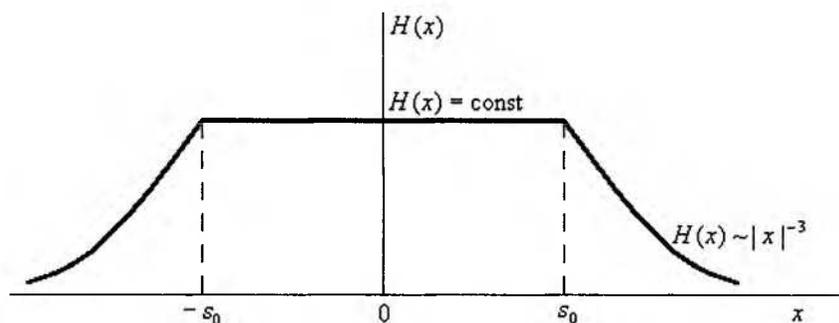


Рис. 5.15. Поведение напряженности поля $H(x)$ внутри и вне катушки

Видим, что функция $H(x)$ имеет разрывные производные $H'(x)$, $H''(x)$, ... при $x = \pm s_0$. Отсюда следует, что искомая функция $J(s)$ должна быть сингулярной при $s = -s_0 + 0$ и $s = s_0 - 0$, а при $|s| > s_0$ ток $J(s) = 0$, как это следует из физической постановки задачи. Кроме того, $J(-s) = J(s)$, если $H(-x) = H(x)$.

Таким образом, функция $J(s)$ должна быть сингулярной и симметричной, т.е. иметь вид, изображенный на рис. 5.16.

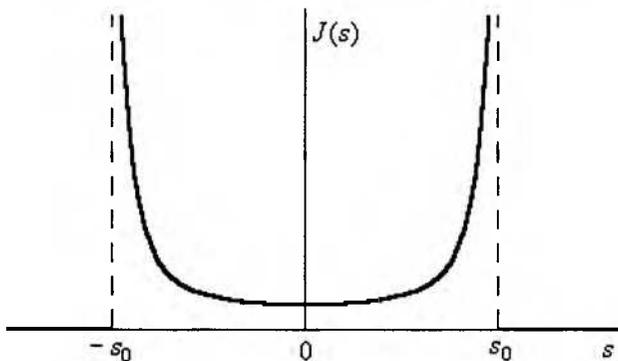


Рис. 5.16. Принципиальный вид функции $J(s)$

Теперь перейдем к количественному алгоритму решения.

Уравнение (5.55) в принципе может быть решено аналитически.

Однако аналитического решения получить не удастся. Поэтому рассмотрим вопрос о численном решении уравнения (5.55). Учитывая некорректность уравнения (5.55), воспользуемся методом регуляризации Тихонова (см. п. 4.7).

В работе [50] также использовался метод регуляризации Тихонова. При этом, как показало численное решение, при малых значениях параметра регуляризации α регуляризованное решение $J_\alpha(s)$ имеет большие знакопеременные флуктуации. В то же время, параметр регуляризации α должен быть действительно очень мал (порядка $10^{-5} - 10^{-10}$). Это обусловлено тем, что погрешность задания $H(x)$ отсутствует (есть только погрешность численного алгоритма). Кроме того, регуляризованное решение $J_\alpha(s)$ должно иметь очень большой диапазон значений (из-за сингулярности точного решения). При этом решение $J_\alpha(s)$ должно быть монотонно возрастающим от центра катушки ($s = 0$) к ее краям ($s = \pm s_0$).

Чтобы регуляризованное решение $J_\alpha(s)$ было таковым, а именно, неотрицательным и монотонно возрастающим от центра катушки к ее краям, нужно воспользоваться *методом регуляризации Тихонова с ограничениями на решение* [10, с. 118].

Ограничения состоят в том, что решение $J_\alpha(s)$ ищется на множестве неотрицательных монотонно невозрастающих функций [10, с. 118]. Однако такому условию удовлетворяет лишь левая половина решения $J_\alpha(s)$, т.е. при $s \in (-s_0, 0]$. Поэтому модифицируем уравнение (5.55), учитывая симметрию решения: $J(-s) = J(s)$.

Модификация уравнения (5.55). Запишем (5.55) в виде:

$$\begin{aligned} H(x) &= \int_{-s_0}^0 \frac{J(s) ds}{\sqrt{[1 + (x-s)^2]^3}} + \int_0^{s_0} \frac{J(s) ds}{\sqrt{[1 + (x-s)^2]^3}} = \\ &= \int_{-s_0}^0 \frac{J(s) ds}{\sqrt{[1 + (x-s)^2]^3}} + \int_{-s_0}^0 \frac{J(-s) ds}{\sqrt{[1 + (x+s)^2]^3}}. \end{aligned}$$

Поскольку $J(s) = J(-s)$, т.е. распределение тока симметрично, то $H(x) = H(-x)$, т.е. магнитное поле $H(x)$ также симметрично. Справедливо и обратное: если $H(x) = H(-x)$, то $J(s) = J(-s)$.

В результате уравнение (5.55) может быть записано в виде:

$$\int_{-s_0}^0 R(x, s) J(s) ds = H(x), \quad -s_0 \leq x \leq 0, \quad (5.57)$$

где новое ядро равно

$$R(x,s) = \frac{1}{\sqrt{[1+(x+s)^2]^3}} + \frac{1}{\sqrt{[1+(x-s)^2]^3}}. \quad (5.58)$$

Итак, $J(s)$ будем искать в левом полупространстве путем решения уравнения (5.57), т.е. при $s \in (-s_0, 0]$, после чего в правом полупространстве $J(s) = J(-s)$, $s \in [0, s_0]$.

При этом дополнительно будем учитывать, что функция $J(s)$, $s \in [-s_0, 0]$, является неотрицательной монотонно невозрастающей типа изображенной на рис. 5.17.

Для получения численного решения $J_\alpha(s)$ методом регуляризации Тихонова на множестве неотрицательных монотонно невозрастающих функций можно воспользоваться программой РТИРР на Фортране [10, с. 118, 174]. В этой программе выполняется минимизация сглаживающего функционала методом проекции сопряженных градиентов на множество неотрицательных монотонно невозрастающих функций при каждом значении параметра регуляризации $\alpha > 0$.

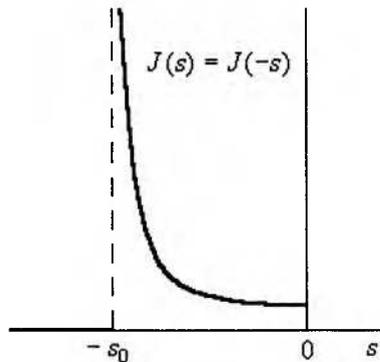


Рис. 5.17

Пример [39]. С помощью программы РТИРР был решен следующий пример (подобный примерам, рассмотренным в работах [7, с. 57], [50]): $s_0 = 1$, шаг дискретизации $h = \Delta s = \Delta x = 0.05$, число шагов дискретизации в левом полупространстве $n = s_0/h = 20$, общее число витков катушки $N = 2n + 1 = 41$, напряженность поля $H(x) = H = \text{const} = 1$, $-s_0 \leq x \leq s_0$.

Задача решалась для ряда значений параметра регуляризации α . На рис. 5.18 отображены полученные решения $J_\alpha(s)$ при $\alpha = 10^{-4}$ (кривая 1), $\alpha = 10^{-5}$ (кривая 2) и $\alpha = 10^{-6}$ (кривая 3).

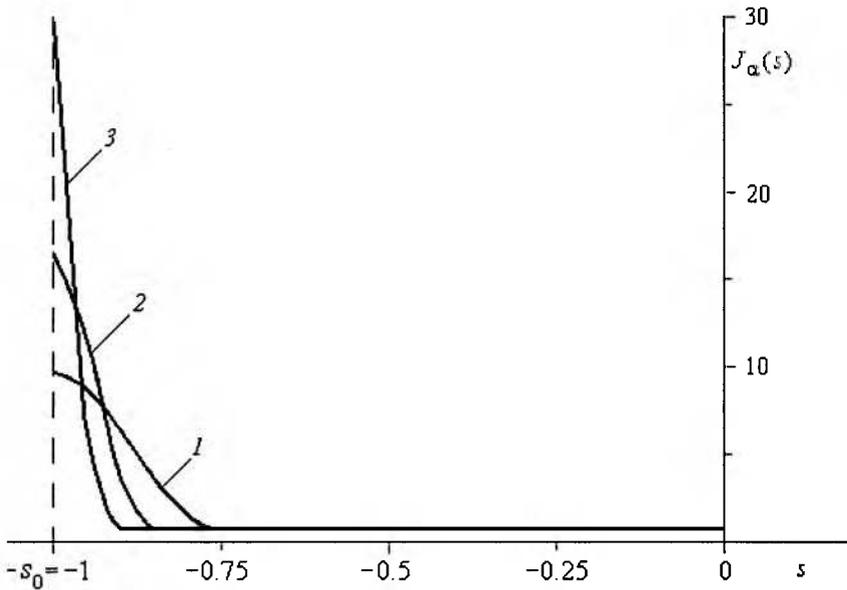


Рис. 5.18. Регуляризованные решения $J_\alpha(s)$ с ограничениями:

$$1 — \alpha = 10^{-4}; \quad 2 — \alpha = 10^{-5}; \quad 3 — \alpha = 10^{-6}$$

На рис. 5.19 отображены соответствующие найденным решениям $J_\alpha(s)$ распределения напряженности магнитного поля (ср. (5.57))

$$H_\alpha(x) = \int_{-s_0}^0 R(x,s) J_\alpha(s) ds, \quad -s_0 \leq x \leq 0, \quad (5.59)$$

при $\alpha = 10^{-4}$ (кривая 1) и $\alpha = 10^{-6}$ (кривая 3).

При этом интеграл в (5.57) расписывался, а в (5.59) вычислялся по формуле трапеций с постоянным шагом h .

Рис. 5.18 показывает, что с уменьшением α отношение

$J_\alpha(-s_0)/J_\alpha(0)$ возрастает и решение $J_\alpha(s)$ в пределе при $\alpha \rightarrow 0$ переходит в сингулярную функцию. А рис. 5.19 показывает, что с уменьшением α относительная неоднородность поля

$$\Delta H_{\alpha \text{ отн}} = \frac{|H_\alpha - H|}{H}$$

убывает.

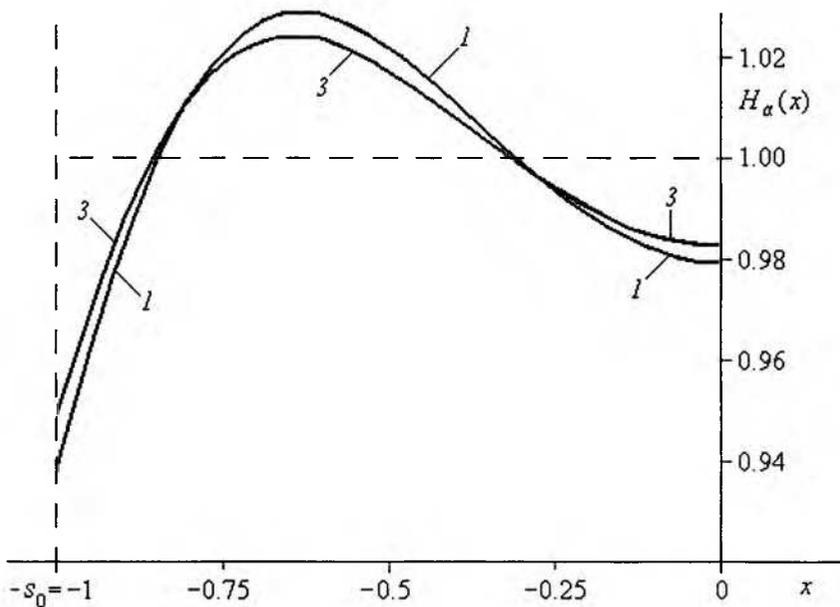


Рис. 5.19. Регуляризованная напряженность $H_\alpha(x)$ магнитного поля:

$$1 - \alpha = 10^{-4}; \quad 2 - \alpha = 10^{-6}$$

Можно сделать следующие выводы: 1) применение метода регуляризации Тихонова без ограничений на решение хотя и стабилизирует решение, но при малых значениях параметра регуляризации α решение $J_\alpha(s)$ имеет флуктуации и тем больше, чем меньше α , как видно из работы [50] (см. также [7, с. 58]); 2) при использовании метода регуляризации Тихонова с ограничениями на решение флуктуации в решении $J_\alpha(s)$ исчезают даже при очень малых значениях α (см. рис. 5.18); 3) как видно из рис. 5.18, при $\alpha \rightarrow 0$ решение $J_\alpha(s)$ переходит в сингулярную функцию типа δ -функции, что практически означает, что катушка должна состоять из соленоида с однородным распределением тока за исключением нескольких крайних витков, в которых ток идет по нарастающей к краю катушки.

Решение путем приближения к δ -функции. Данный пример был решен также следующим способом [39]. Искомая функция $J(s)$ полагалась равной (приближение к δ -функции)

$$J(s) = \begin{cases} J_0, & s = -s_0, \\ J, & s \neq -s_0, \end{cases} \quad (5.60)$$

где $J = \text{const} = 1$, а J_0 варьировалось. Технически это равносильно катушке в виде соленоида с током J и паре тонких витков (колец Гельмгольца) с током J_0 на ее краях (соленоид и витки имеют одинаковый радиус).

При этом интеграл в (5.57) расписывался по формуле трапеций и напряженность $H(x)$ рассчитывалась как

$$H(x) = h \sum_{j=0}^n p_j R(x, s_j) J(s_j), \quad (5.61)$$

где $s_j = -s_0 + h \cdot j$, $p_j = \begin{cases} 0.5, & j = 0 \text{ или } j = n, \\ 1, & 0 < j < n. \end{cases}$

На рис. 5.20 приведены рассчитанные напряженности $H(x)$ при $J_0 = 25$ (кривая 1), $J_0 = 30$ (кривая 2) и $J_0 = 35$ (кривая 3).

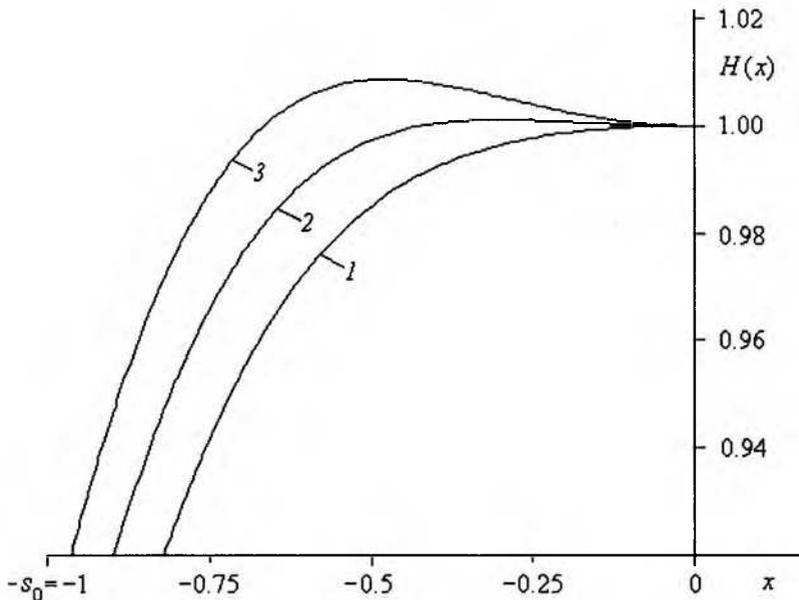


Рис. 5.20. Напряженности $H(x)$, рассчитанные согласно (5.61):
1 — $J_0 = 25$; 2 — $J_0 = 30$; 3 — $J_0 = 35$

Видим, что при некотором J_0 однородность поля (кривая 2 на рис. 5.20 при $|x| \leq 0.4 s_0$ – в рабочем объеме) в среднем порядка 10^{-4} . При этом с увеличением числа витков N неоднородность поля уменьшается.

Можно сделать следующий вывод. Если распределение тока $J(s)$ вдоль обмотки катушки ЯМР-томографа сделать однородным за исключением крайних витков (т.е. типа (5.60)), то это позволит повысить однородность магнитного поля $H(x)$ на оси катушки. Данный способ (приближение к δ -функции) хотя и является более простым, чем предыдущий (метод регуляризации Тихонова с ограничениями на решение), но приводит к более точному результату (к более однородному полю). Это показывает, как важно использовать дополнительную информацию при решении некорректных задач и на ее основе строить адекватные алгоритмы. Хотя оба способа используют информацию о решении, но второй способ дает более адекватное математическое описание задачи и, как следствие, более точное решение.

Второй способ не является новым. Его, по-существу, использовали Л. Луганский [31] и др. Однако цель нашего изложения не состоит в получении наиболее точных практических результатов, а в демонстрации различных вариантов устойчивых методик решения некорректных задач.

ЛИТЕРАТУРА к Части II

Основная

1. *Бакушинский А.Б., Гончарский А.В.* Некорректные задачи. Численные методы и приложения. – М.: Изд-во МГУ, 1989. 199 с.
2. *Бейтс Р., Мак-Доннелл М.* Восстановление и реконструкция изображений. – М.: Мир, 1989. 336 с.
3. *Бронштейн И.Н., Семендяев К.А.* Справочник по математике для инженеров и учащихся втузов. 13-е изд. – М.: Наука, 1986. 544 с.
4. *Верлань А.Ф., Сизиков В.С.* Интегральные уравнения: методы, алгоритмы, программы. – Киев: Наук. думка, 1986. 544 с.
5. *Иванов В.К., Васин В.В., Танана В.П.* Теория линейных некорректных задач и ее приложения. – М.: Наука, 1978. 206 с.
6. *Лаврентьев М.М., Романов В.Г., Шишатский С.П.* Некорректные задачи математической физики и анализа. – М.: Наука, 1980. 287 с.
7. *Сизиков В.С.* Математические методы обработки результатов измерений. – СПб: Политехника, 2001. 240 с. (Электронный вариант: Сизиков В.С. Устойчивые методы обработки результатов измерений. Электронный учебник: <http://de.ifmo.ru/library/sizikov.pdf>).
8. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. 3-е изд. – М.: Наука, 1986. 288 с.
9. *Тихонов А.Н., Арсенин В.Я., Тимонов А.А.* Математические задачи компьютерной томографии. – М.: Наука, 1987. 160 с.
10. *Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г.* Численные методы решения некорректных задач. – М.: Наука, 1990. 232 с.
11. *Уэбб С. (ред).* Физика визуализации изображений в медицине. В 2-х т. – М.: Мир, 1991. Т. 1, 408 с. Т. 2, 408 с.

Дополнительная

12. *Апарцин А.С.* Неклассические уравнения Вольтерра I рода: теория и численные методы. – Новосибирск: Наука, 1999. 193 с.
13. *Афанасьев Ю.В., Студенцов Н.В., Хорев В.Н.* и др. Средства измерений параметров магнитного поля. – Л.: Энергия, 1979. 320 с.
14. *Белов И.А., Сизиков В.С.* Программа реконструкции смазанных и дефокусированных электронных изображений. – Свидетельство об официальной регистрации программы для ЭВМ № 2001610743. 18.06.2001.
15. *Брейсуэлл Р.* Преобразование Хартли. – М.: Мир, 1990. 175 с.
16. *Василенко Г.И.* Теория восстановления сигналов. – М.: Сов. радио, 1979. 272 с.

17. *Васильев В.Н., Гуров И.Н.* Компьютерная обработка сигналов в приложении к интерферометрическим системам. – СПб: БХВ–Санкт-Петербург, 1998. 240 с.
18. *Воеводин В.В.* Линейная алгебра. – М.: Наука, 1980. 400 с.
19. *Воскобойников Ю.Е., Преображенский Н.Г., Седельников А.И.* Математическая обработка эксперимента в молекулярной газодинамике. – Новосибирск: Наука, 1984. 240 с.
20. *Галайдин П.А., Замятин А.И., Иванов В.А.* Расчет и проектирование электромагнитных систем магниторезонансных томографов. Уч. пособие. – СПб: Изд-во ИТМО, 1998. 29 с.
21. *Гантмахер Ф.Р.* Теория матриц. – М.: Наука, 1966. 576 с.
22. *Гласко В.Б., Мудрецова Е.А., Страхов В.Н.* Обратные задачи гравиметрии и магнитометрии // Некорректные задачи естествознания / Под ред. А.Н. Тихонова, А.В. Гончарского. – М.: Изд-во МГУ, 1987, с. 89–102.
23. *Данилович Я.Б., Петров Ю.П.* О необходимости расширения понятия эквивалентности математических моделей // Докл. АН, 2000, т. 371, № 4, с. 473–475.
24. *Дружкин Л.А.* Задачи теории поля. – М.: Изд-во МИРГЭ, 1964. 462 с.
25. *Канторович Л.В., Акилов Г.П.* Функциональный анализ в нормированных пространствах. – М.: Физматгиз, 1959. 684 с.
26. *Колмогоров А.Н., Фомин С.В.* Элементы теории функций и функционального анализа. 5-е изд. – М.: Наука, 1981. 544 с.
27. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. – М.: Наука, 1968. 720 с.
28. *Курош А.Г.* Курс высшей алгебры. 11-е изд. – М.: Наука, 1975. 431 с.
29. *Лаврентьев М.М.* Некорректные задачи для дифференциальных уравнений. – Новосибирск: Изд-во НГУ, 1981. 74 с.
30. *Латтес Р., Лионс Ж.-Л.* Метод квазиобращения и его приложения. – М.: Мир, 1970. 334 с.
31. *Луганский Л.Б.* Расчет соленоида с заданной формой магнитного поля // Журн. техн. физики, 1985, т. 55, вып. 7, с. 1263–1271.
32. *Лухвич А.А., Чурило В.Р.* Источники поляризующего магнитного поля и его градиентов для я.м.р.-томографии (обзор) // ПТЭ, 1987, № 5, с. 172–173.
33. *Монтгомери Д.Б.* Получение сильных магнитных полей с помощью соленоидов. – М.: Мир, 1971. 359 с.
34. *Морозов В.А.* Регулярные методы решения некорректно поставленных задач. – М.: Наука, 1987. 240 с.
35. *Наттерер Ф.* Математические аспекты компьютерной томографии.

- М.: Мир, 1990. 288 с.
36. *Петров Ю.П.* Третий класс задач физики и техники – промежуточных между корректными и некорректными. – СПб.: Изд-во СПбГУ, 1998. 30 с.
 37. *Понтрягин Л.С.* Обыкновенные дифференциальные уравнения. 5-е изд. – М.: Наука, 1982. 331 с.
 38. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов. – М.: Мир, 1978. 848 с.
 39. *Сизиков В.С., Ахмадулин Р.И., Николаев Д.Б.* Синтез магнитного поля вдоль оси катушки ЯМР-томографа // Изв. вузов. Приборостроение, 2002, т. 45, № 1, с. 52–57.
 40. *Сизиков В.С., Белов И.А.* Реконструкция смазанных и дефокусированных изображений методом регуляризации // Оптич. ж., 2000, т. 67, № 4, с. 60–63.
 41. *Тихонов А.Н.* Об устойчивости обратных задач // Докл. АН СССР, 1943, т. 39, № 5, с. 195–198.
 42. *Тихонов А.Н., Арсенин В.Я., Рубашов И.Б., Тимонов А.А.* Первый советский компьютерный томограф // Природа, 1984, № 4, с. 11–21.
 43. *Тихонов А.Н., Гончарский А.В., Степанов В.В.* Обратные задачи обработки фотоизображений // Некорректные задачи естествознания / Под ред. А.Н. Тихонова, А.В. Гончарского. – М.: Изд-во МГУ, 1987, с. 185–195.
 44. *Тихонов А.Н., Леонов А.С., Ягола А.Г.* Нелинейные некорректные задачи. – М.: Наука, 1995. 550 с.
 45. *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. – М.: Наука, 1972. 596 с.
 46. *Троицкий И.Н.* Статистическая теория томографии. – М.: Радио и связь, 1989. 240 с.
 47. *Уилкинсон Дж. Х.* Алгебраическая проблема собственных значений. – М.: Наука, 1970. 564 с.
 48. *Фриш С.Э., Тиморева А.В.* Курс общей физики. Т. 2. – МЛ: ГИТТЛ, 1952. 616 с.
 49. *Чейссон Э. Дж.* Первые результаты с космического телескопа «Хаббл» // В мире науки, 1992, № 8, с. 6–14.
 50. *Adamiak K.* Method of the magnetic field synthesis on the axis of cylinder solenoid // Appl. Phys., 1978, v. 16, p. 417–423.
 51. *Brunner H., Sizikov V.* On a suboptimal filtration method for solving convolution-type integral equations of the first kind // J. Math. Anal. and Appl., 1998, v. 226, № 2, p. 292–308.
 52. *Cho Z.H., Jones J.P., Singh M.* Foundations of medical imaging. – NY: Wiley, 1993.

Приложение О «ГРАММАТИКЕ» НАУКИ

Цикл лекций для студентов, аспирантов и научных работников

Лекция первая

1. Введение

Опыт работы с аспирантами и молодыми научными сотрудниками показал, что основные проблемы в их подготовке не связаны обычно со слабым знанием основных положений и фактов – со знанием того, что можно условно назвать «словарем» науки. С этим чаще всего все в порядке.

Основные пробелы в их подготовке связаны, чаще всего, со слабым знанием того, что можно (разумеется, условно) назвать «грамматикой» науки, знанием того, как связаны между собой факты и законы, как на самом деле, в реальной научной жизни, устанавливаются новые законы и теоремы и как они опровергаются, как должен относиться аспирант и молодой научный сотрудник к бурному и стремительному потоку публикаций, потоку текущей научной информации, в котором так легко захлебнуться. А ведь надо не только плыть (не захлебываясь) в этом потоке, но и самому понемногу приумножать его, приумножать, строго соблюдая этику научной публикации и научной дискуссии, этику, выработанную опытом веков, опытом наших великих предшественников.

В освоении «грамматики» науки (не менее важной, чем ее «словарь») могли бы очень помочь учебные курсы по истории науки, истории различных научных дисциплин. К сожалению, эти курсы читаются далеко не везде, что и приводит к серьезным пробелам в подготовке молодых ученых (положение, возможно, изменится в 2003 году, когда будут введены в кандидатский минимум экзамены по истории науки (по научным дисциплинам); решение об этом принято в 2002 году).

Настоящие «лекции» имеют целью частично восполнить эти пробелы. «Лекции» построены на материале научных исследований, выполненных автором и его сотрудниками в 1990–2002 годах на одной из кафедр факультета Прикладной Математики – Процессов Управления (факультет ПМ-МУ) Санкт-Петербургского Государственного Университета (СПбГУ) (частично изложенных в первой и второй главах) и на опыте последующего обсуждения результатов этих исследований.

2. Теоремы и контрпримеры

Все началось с того, что мною в 1990 году была обнаружена неполнота весьма важного и широко применяемого во всех проектно-конструкторских организациях метода проверки параметрической устойчивости. Хорошо известно, что для надежной работы любых технических систем и устройств (а также объектов экономики, биологии, медицины) необходимо, чтобы они были не только устойчивы, но и сохраняли устойчивость при неизбежных на практике малых отклонениях (вариациях) параметров системы. Свойство сохранения устойчивости при вариациях параметров называют, в последнее время, коротко: «параметрической устойчивостью».

Еще в первой половине 20 века была постепенно сформулирована и сразу начала применяться на практике важная теорема: «если характеристический полином линейной системы имеет все корни в левой полуплоскости комплексной переменной и далеко от мнимой оси, то система параметрически устойчива» (дается осовремененная формулировка). Доказательство было основано на другой хорошо известной теореме о непрерывной зависимости корней полинома от его коэффициентов.

Но в 1990 году неожиданно было обнаружено, что известная еще с первой половины 20 века теорема о параметрической устойчивости не верна. Были продемонстрированы примеры систем, корни характеристических полиномов которых лежали в левой полуплоскости далеко от мнимой оси, а системы, тем не менее, теряли устойчивость при сколь угодно малых, неизбежных на практике, вариациях параметров. Вот простой пример:

система (где $D = \frac{d}{dt}$)

$$\begin{aligned} (D^3 + 4D^2 + 5D + 2)x_1 - (D^2 + 2D + 1)x_2 &= 0, \\ (D^2 + 4D + 5)x_1 - (kD + 1)x_2 &= 0 \end{aligned} \quad (1)$$

при $k=1$ имеет характеристический полином $\Delta = \lambda^3 + 5\lambda + 7\lambda + 3$ с корнями $\lambda_1 = -3$; $\lambda_2 = \lambda_3 = -1$; однако, если коэффициент (параметр) k отличается от $k=1$ на сколь угодно малую положительную величину, то устойчивость теряется.

Полученный результат, опубликованный в [1], был очень важен. Он раскрывал причину таинственных до этого аварий, происходящих от неполноты традиционных методов расчета, позволял усовершенствовать, дополнить традиционных методы, уменьшить вероятность аварий самолетов, атомных реакторов и других технических объектов.

Тем не менее, этот важный результат был встречен с большим недоверием. Мне говорили: «Юрий Петрович, у Вас все уж слишком просто. Привели примеры – и все. Нет, Вы покажите: где, в каком мес-

те, содержится ошибка в доказательстве теоремы о параметрической устойчивости».

Это возражение сразу показывает, насколько слабо ориентированы молодые научные сотрудники в соотношениях между теоремами, их доказательствами и контрпримерами, т.е. примерами, доказывающими неверность ранее сформулированных теорем. А ведь еще в 30-х годах 20 века выдающийся венгерский математик Д. Пойа (в старых переводах – Полия) справедливо подчеркивал: «Математика состоит из двух вещей – теорем и контрпримеров» (т.е. контрпримеры не менее важны, чем теоремы и их доказательства).

Действительно, в реальной работе математика теорема возникает чаще всего как обобщение подтверждающих утверждение теоремы примеров. Так, легко проверить, что существует много примеров систем с корнями, лежащими в левой полуплоскости комплексного переменного и обладающими параметрической устойчивостью. Отсюда возникает догадка (будущая теорема): *все* такие системы обладают параметрической устойчивостью. Но еще до того, как начать эту догадку доказывать, превращать в теорему, исследователь проверяет догадку контрпримером – проверяет, а не существует ли систем, с корнями, лежащими в левой полуплоскости, но параметрически неустойчивых. Если хотя бы одна такая система обнаружилась – догадке конец, пытаться доказывать ее как теорему – пустая трата времени (а ведь сколько времени тратят впустую те научные работники, к сожалению, многие, которые недооценивают роль контрпримера!). Только если ни одного контрпримера не обнаружено, исследователь начинает доказывать теорему.

А когда теорема (и ее доказательство) опубликованы, то коллеги автора теоремы начинают, прежде всего, с поиска опровергающих ее контрпримеров. Начинают потому, что найти ошибку в чужом доказательстве трудно, очень трудно. Если сомневаешься в утверждении теоремы, то самое простое и эффективное – это искать контрпример. Так в реальной научной жизни чаще всего и поступают; именно так, через контрпример, чаще всего и рушатся опубликованные, но неверные теоремы (а доказательства анализируются редко).

Исторический пример: в 1821 году признанный лидер математики того времени О. Коши опубликовал (с доказательством, разумеется) теорему: «Сумма любого сходящегося ряда непрерывных функций непрерывна». А через пять лет Н. Абель нашел (и опубликовал) всего один сходящийся ряд непрерывных функций, не имеющий непрерывной суммы. Несмотря на то, что Абелю было тогда всего 24 года и он был почти никому не известен в математическом мире, после публикации его контрпримера теорема великого Коши была немедленно при-

знана ошибочной. И только через 22 года Зейделем и Стоксом была вскрыта причина ошибки в доказательстве Коши. Как и всякая ошибка гения, она оказалась поучительной и глубокой – Коши не учел (да и не мог учесть) неизвестного в 1821 году требования *равномерной сходимости ряда*.

Любопытна, разумеется, и непростая причина ошибки в доказательстве теоремы о параметрической устойчивости. Я расскажу о ней в последней лекции, но особой необходимости знакомства с этой непростой причиной нет. Если продемонстрирован контрпример, то теорема все равно неверна, и в причину можно не вникать. Необходимо лишь, разумеется, заменить опровергнутую теорему другой, более узкой, но верной. Я предложил использовать в расчетах устойчивости следующую теорему (доказательство приведено в [2]): параметрической устойчивостью обладают линейные системы, у которых корни характеристического полинома лежат в левой полуплоскости далеко от мнимой оси, *кроме* систем особых, т.е. систем, у которых определитель «матрицы степеней» обращается в нуль. Необходимо, конечно, рассказать, что такое «матрица степеней», но о ней подробно рассказано в известной публикации [2] – во всех трех изданиях книги [2], о «матрице степеней» рассказано на стр. 74–82.

Из этой новой теоремы следует простое практическое правило надежного, достоверного расчета: надо сначала проверить, не является ли система особой (а это очень легко сделать), после чего уже можно применять традиционные методы.

3. О правилах научной дискуссии

В феврале 2002 года новая теорема о параметрической устойчивости в кругу других интересных вопросов, о которых расскажу далее, обсуждалась на расширенном научном семинаре, на факультете ПМ–ПУ, СПбГУ, и один из наиболее авторитетных присутствующих профессоров выдвинул следующее возражение. По его мнению, «опубликованные в [1, 2] контрпримеры не опровергают старой теоремы, поскольку все они относятся к вырожденным системам». Напомню, что вырожденными называют системы дифференциальных уравнений, порядок которых ниже, чем у систем того же вида, но с другими коэффициентами. Вырожденные системы часто встречаются в приложениях.

Данное возражение произвело впечатление на присутствующих на семинаре молодых (да и не только молодых) научных работников. Между тем, при самом простом анализе оно показывает свою несостоятельность. Действительно, старая теорема утверждала: параметрической устойчивостью обладают *все* системы с «хорошими» корнями. С учетом опровергающих ее контрпримеров, я предложил уточненную

формулировку – параметрической устойчивостью обладают системы с «хорошими» корнями, *кроме* систем особых. Возражение уважаемого профессора равносильно другой формулировке: параметрической устойчивостью обладают системы с «хорошими» корнями, *кроме* систем вырожденных. Остается лишь доказать, что класс вырожденных систем совпадает с классом систем особых (возражающим это не было доказано) – и мы приходим к формулировке, уже опубликованной в [2].

Таким образом, *все* присутствующие на семинаре 11.02.2002 г. и молодые, и не молодые научные сотрудники не заметили, что в качестве возражения выдвигается не доказанное утверждение, а если оно будет доказано, то оно не только не опровергнет, а лишний раз докажет оспариваемое возражающим положение, ранее опубликованное в [2].

Данное обстоятельство лишний раз говорит о том, насколько оказалось утраченной культура научной дискуссии, без которой серьезная наука невозможна. А причина простая: во второй половине XX века научные журналы редко публиковали дискуссии, и с каждым десятилетием все реже и реже. Серьезная письменная и печатная дискуссия подменялись ее суррогатом – устным обсуждением на конференциях и семинарах. Но в устной речи, если вопрос сложный, трудно уследить за логикой собеседника, да и высказанные доводы быстро забываются (как сформулировал В. Маканин, «когда я просто говорю – я за свои слова не отвечаю. Я отвечаю только за то, что написано моим пером»). Необходимо возродить публикацию дискуссий в научных журналах и шире рассказывать о великих научных дискуссиях прошлого в курсах истории науки. Только тогда мы вырастим настоящих научных работников-исследователей, которых не повергнут в смущение фразы о «вырожденных системах».

Приведем дополнительные примеры.

Формируя правила научной дискуссии, полезно вспомнить очень старый известный ответ Лежандра (1752–1833) Гауссу (1777–1855) во время их дискуссии о приоритете: «Не существует открытия, писал Лежандр, которое нельзя было бы приписать себе, сказав, что те же вещи были найдены мной на несколько лет раньше; но если не дано этим словам доказательства, состоящего в указании места, где они были опубликованы, то это утверждение становится беспредметным и представляет собой только обиду для истинного автора открытия».

«В математике случается очень часто, продолжает Лежандр, что находят заново те самые вещи, которые уже были открыты другими. Подобное случалось со мной много раз, но я никогда не упоминал об этом и не называл «нашим принципом» принцип, который другой опубликовал раньше меня».

А вот другой пример, уже из XX века: в 1960 году член-

корреспондент Академии наук СССР Александр Михайлович Летов предложил метод «аналитического конструирования», т.е., собственно, метод расчета регуляторов вида

$$u = \sum_i k_i x_i, \quad (2)$$

где u – управляющее воздействие, k_i – постоянные коэффициенты, x_i – все переменные, от первой, x_1 , до последней, x_n , присутствующие в математической модели объекта управления (опубликовано в [3]). А.М. Летов показал, что можно так рассчитать величины коэффициентов k_i , что будет обеспечена устойчивость системы управления, ее параметрическая устойчивость, а также хорошие переходные процессы. При этом все коэффициенты k_i в регуляторе (2) рассчитывались аналитически (а не подбором, как ранее), почему А.М. Летов и назвал свой метод «аналитическим конструированием».

Поскольку регуляторы вида (2) легко реализовывались и обеспечивали хорошее качество регулирования, они быстро стали широко использоваться на практике. Сообщения об использовании их то на одном, то на другом объекте управления десятками появлялись в 60-х гг. на страницах технических журналов того времени.

А потом произошел большой конфуз: «аналитически сконструированные» регуляторы стали причиной нескольких серьезных аварий из-за потери устойчивости, хотя по расчету казалось, что устойчивость обеспечена. Причина заключалась в том, что у многих объектов управления не все переменные x_i доступны для использования в регуляторе и их заменяли измеряемыми, пользуясь, разумеется, только эквивалентными преобразованиями, не изменяющими решений системы и ее переходных процессов. В те годы еще не знали, что эквивалентные (в классическом смысле) преобразования могут изменять параметрическую устойчивость, что и было истинной причиной аварий. Но этого в те годы не знали, и причины аварий списывали на «недостатки» метода аналитического конструирования. В результате «аналитическое конструирование» просто перестали применять, заменив его более сложными методами оптимального управления, не порождающими столь частые аварии.

Таким образом, на практике уже приходилось иметь дело с авариями, причиной которых были ошибки расчета, неразличение преобразований, эквивалентных в классическом смысле и в расширенном (о необходимости такого различения я расскажу в следующем разделе). Но тогда, в 60-70-е годы XX века эта причина еще не была выяснена. В результате убрали только наиболее явный источник аварий – исполь-

зование метода «аналитического конструирования» с исключением части переменных, но аварии, порождаемые не различием преобразований, эквивалентных в различных смыслах, хотя и более редко, но все равно происходили (см. публикацию [2]).

Истинная причина аварий не была раскрыта в те годы потому, что не было тогда проведено на страницах научных журналов полноценной широкой дискуссии по поводу неожиданной и парадоксальной потери параметрической устойчивости. Научное руководство в те годы совсем не поощряло научные дискуссии в печати, на страницах журналов и книг. Я хорошо помню, как мы, специалисты, в своем кругу устно, очень горячо обсуждали интереснейшее и таинственное явление потери устойчивости при «аналитическом конструировании», но до истинной причины добраться тогда не удалось (впервые все прояснилось лишь много лет спустя, в 1987 году, в [4]).

Устное обсуждение, даже самое горячее и искреннее, не могло заменить тогда полноценной дискуссии в печати. Если бы такая дискуссия прошла тогда, то различие между преобразованиями, эквивалентными в классическом смысле и в расширенном, было бы, наверное, обнаружено много раньше и многих аварий удалось бы избежать.

Эти примеры, взятые как из давней, так и из совсем недавней истории науки, лишний раз свидетельствуют о важности письменного (а еще лучше – печатного) оформления научной дискуссии, когда аргументы оппонентов четко фиксируются на бумаге и могут быть в дальнейшем беспристрастно оценены сторонними лицами. Именно они – сторонние лица – выносят окончательный приговор – кто был прав и кто не прав в научной дискуссии, но сделать это они могут, опираясь только на печатный (в крайнем случае – рукописный) материал.

Лекция вторая

4. Знание, которое необходимо всем

Если материал первой лекции представлял интерес для аспирантов и исследователей, встретившихся с проблемой устойчивости, то во второй лекции я расскажу о недавно обнаруженных новых важных явлениях, касающихся практически всех аспирантов и исследователей в области естественных наук. Речь пойдет об эквивалентных (равносильных) преобразованиях математических моделей исследуемых систем – о преобразованиях, без которых не обходится почти ни один исследователь.

Напомню, что согласно общепринятому определению *эквивалентными* (или, что то же самое, *равносильными*) называют *преобразования*, не изменяющие решений. Примеры: перенос членов из левой час-

ти в правую с изменением знака, умножение всех членов на число не равное нулю, замена любого члена на равный ему и т.п.

Правила эквивалентных преобразований известны с XVII века. Они изучаются еще в средней школе и самым широчайшим образом применяются во всех исследованиях и расчетах. За века их применения незаметно сложилось твердое убеждение – раз эквивалентные преобразования не меняют решений преобразуемой системы, то они вообще ничего не меняют: «если что-то изменилось, значит, использованное Вами преобразование было, на самом деле, не эквивалентным; ищите ошибку». Вот ход мысли сегодняшних аспирантов (да и не только аспирантов), с которым я неоднократно встречался. А между тем наши великие предшественники, великие математики прошлых веков, выбиравшие названия для математических операций, различали преобразования *тождественные* и преобразования *эквивалентные*. Они догадывались, наверное, что эквивалентные преобразования могут что-то менять. Но что именно? Вот в этом-то и вопрос, который, как ни странно, веками не поднимался и поэтому до последнего времени не получал ответа.

Не всегда обращают внимание на то, что чаще всего нам нужно не просто решение той или иной задачи, а надежное, корректное решение. Дело в том, что почти все коэффициенты и параметры математической модели определяются из опыта и измерения с неизбежными, хотя и малыми, погрешностями. Если эти неизбежные малые погрешности ведут к большим погрешностям в решении, то решение никакого практического смысла не имеет. Подобные задачи называют некорректными и стараются их избежать; они требуют совершенно особого подхода. Точные математические формулировки корректности желающих может найти в [5, 2].

Пример: для системы, не обладающей параметрической устойчивостью, задача проверки – устойчива система или нет – некорректна. Неизбежные малые погрешности в знании коэффициентов и параметров сделают результат расчета совершенно ненадежным: устойчивую систему мы можем принять за неустойчивую и наоборот.

Как раз при анализе параметрической устойчивости и был получен ответ на вопрос, поставленный великими математиками прошлых веков: *эквивалентные преобразования (в отличие от тождественных) могут изменять корректность решаемой задачи.*

Действительно, рассмотрим систему (1). При $k=1$ она, как мы уже показали, на первой лекции, не обладает параметрической устойчивостью. Задача проверки – устойчиво ли решение $x_1(t)$ системы (1) при $k=1$ или нет – некорректна. Теперь исключим переменную $x_2(t)$ из системы (1) для $k=1$ путем эквивалентных преобразований. Получим

уравнение:

$$(D^3+5D^2+7D+3)x_1=0. \quad (3)$$

Нетрудно проверить, что уравнение (3) имеет то же самое решение:

$$x_1(t) = C_1 e^{-3t} + (C_2 t + C_3) e^{-t}, \quad (4)$$

что и система (1). Поэтому преобразование системы (1) в уравнение (3) для задачи нахождения решения $x_1(t)$ является эквивалентным преобразованием. Однако решение (4) уравнения (3) параметрически устойчиво по отношению ко всем коэффициентам и параметрам уравнения (3), а то же самое решение $x_1(t)$ системы (1) при $k=1$ параметрически неустойчиво; задача проверки устойчивости системы (1) для $k=1$ некорректна.

Таким образом, мы убедились, что существуют эквивалентные преобразования, способные изменить корректность решаемой задачи. Многочисленные более сложные примеры подобных преобразований содержатся в монографии [2].

С учетом существования подобных преобразований, в работе [2] было предложено уточнить представление об эквивалентных преобразованиях и разделить их на:

1) *преобразования, эквивалентные в классическом смысле*, т.е. те, которые (согласно классическому определению) не изменяют решений преобразуемой системы;

2) *преобразования, эквивалентные в расширенном смысле*, т.е. те, которые, во-первых, эквивалентны в классическом смысле, во-вторых, не изменяют корректности решаемой задачи.

Из приведенного примера (и других примеров, опубликованных в [2]) следует, что существуют преобразования, эквивалентные в классическом смысле, но не в расширенном.

Теперь мы видим, что материал, изложенный в настоящей лекции (а более подробно – в публикациях [1, 2]), действительно важен и необходим практически для всех аспирантов и научных работников (точнее – для всех тех, кто использует эквивалентные преобразования; очень трудно найти работу по естественным наукам, в которой они не использовались бы). А необходим этот материал потому, что только использование преобразований, эквивалентных в расширенном смысле, обеспечивает достоверность и страхует от ошибок. Более подробно, с многочисленными примерами, это показано в [2].

5. Наиболее важное следствие

Как только были открыты преобразования, эквивалентные в классическом смысле, но не в расширенном, сразу обнаружилось, что многие традиционные методы расчета, использующие эти преобразования,

недостовверны. Они могут приводить к грубым ошибкам и для восстановления достоверности нужно проводить дополнительные расчеты и проверки, описанные в [2].

Наиболее важными из методов расчета, неполнота которых была обнаружена, являются методы численного решения обыкновенных дифференциальных уравнений, очень часто используемых на практике. Хорошо известно, что первичной, непосредственно вытекающий из законов физики формой записи дифференциальных уравнений, описывающих изучаемый объект, является система, состоящая из уравнений различных порядков. Так, например, чаще всего математическую модель получают из известных уравнений Лагранжа второго рода, а это приводит к системе из уравнений второго порядка. В дальнейшем почти всегда любую систему, состоящую из уравнений различных порядков, приводят к нормальной форме Коши, т.е. к системе уравнений первого порядка; приводят хотя бы потому, что все стандартные программы численного решения уравнений составлены для нормальной формы. Составлять большое количество программ для всего обширного многообразия разнообразных систем было бы слишком громоздко. Гораздо проще привести исходную систему к нормальной форме, разумеется, с помощью эквивалентных преобразований. Это гарантирует, что решения преобразованной системы совпадут с решениями системы исходной.

Поскольку почти все коэффициенты и параметры математической модели, как уже говорилось, получаются из опыта и имеют ограниченную точность, то необходимым (хотя и недостаточным) условием достоверности расчета является непрерывная зависимость решений от параметров (в частном случае – от коэффициентов). Поэтому основной теоремой теории дифференциальных уравнений, лежащей в основе всех практических расчетов, является теорема о непрерывной зависимости решений от параметров.

Однако, обратите внимание, в учебниках по дифференциальным уравнениям эта теорема доказана лишь для систем в нормальной форме (в некоторых учебниках приводится еще и доказательство для одного уравнения n -го порядка). А другие системы, системы не в нормальной форме? У них у всех решения зависят от параметров непрерывно, или же нет? Я задавал этот вопрос многим уважаемым специалистам по дифференциальным уравнениям и всегда получал уверенный ответ: «Да, конечно. Во всяком случае, для систем, которые можно эквивалентными преобразованиями привести к нормальной форме, непрерывная зависимость решений от параметров существует, никаких сомнений нет».

Эта всеобщая уверенность (в свою очередь, опирающаяся на уве-

ренность в том, что эквивалентные преобразования «ничего не меняют») привела к тому, что никто не искал особые системы, не обладающие свойством непрерывной зависимости решений от параметров. А такие системы существуют. Рассмотрим еще раз систему (1). Нетрудно проверить (подробный разбор приведен в публикации [6] и в третьем издании книги [2]), что при $k=1$ зависимость решений $x_1(t=t_0)$ для любого $t_0>0$ от коэффициента k терпит разрыв и поэтому вблизи $k=1$ любые решения совершенно не достоверны. Посчитайте и убедитесь сами, что при переходе, например, от $k=1.0001$ к $k=1.00011$ решение $x_1(10)$, т.е. значение $x_1(t)$ при $t=10$ изменится во много раз.

Для обеспечения достоверности результатов любого расчета, использующего дифференциальные уравнения, необходимо, прежде всего, проверить, не является ли исследуемая система особой, не имеющей непрерывной зависимости решений от коэффициентов и параметров. Методы такой проверки не сложны. Они описаны подробно в публикациях [6] и [2] (третье издание).

Эти результаты важны для очень широкого круга лиц – для всех аспирантов, инженеров, исследователей, использующих в своей работе дифференциальные уравнения.

6. Справедлива ли для аспирантов «презумпция невиновности»?

Когда я рассказывал о результатах предыдущего раздела аспирантам СПбГУ, использующим в своих диссертациях дифференциальные уравнения, и советовал проверить – для обеспечения достоверности – не являются ли исследуемые нами системы особыми, то мне пришлось столкнуться со стойким непониманием: « Юрий Петрович, у нас все правильно, мы считали добросовестно, покажите нам, где у нас ошибка. Если она есть, то тогда (и только тогда) мы заинтересуемся Вашими результатами и посмотрим их». Эта позиция аспирантов была обусловлена позицией их научных руководителей, утверждавших что «аспирант, как подсудимый в суде, должен пользоваться «презумпцией невиновности» – пока не доказано, что его результат неверен, следует считать, что в диссертации все верно и дополнительных проверок можно не проводить».

Мы убеждаемся теперь, что с «грамматикой науки» плохо знакомы не только аспиранты, но и часть научных руководителей. «Презумпция невиновности» у аспирантов, как и у любого научного работника, нет и быть не может. Скорее уж можно говорить о «презумпции виновности»: любой инженер-практик, желающий использовать новый результат, полученный исследователем, а также и любой рецензент диссертации имеют полное право проявить законное недоверие, потребовать, чтобы им доказали как верность, так и достоверность полученного ис-

следователем (или аспирантом) результата его работы. Если у практика (или рецензента диссертации) возникли сомнения – они в любом случае будут истолкованы не в пользу аспиранта или исследователя. На это не нужно обижаться, а нужно просто стараться как можно более доходчиво и ясно доказать достоверность своих результатов.

Если рецензент показал, что ошибка *возможна*, то долг аспиранта *доказать*, что эта возможность в его диссертации *не* реализуется, что на самом деле все верно и возможной ошибки, по причине, которую должен четко обосновать диссертант, на самом деле в его диссертации все же нет.

7. Необходимость согласия в определениях

Одно из основных положений «грамматики науки», без которого невозможно ни понимание чужих научных работ, ни плодотворная дискуссия – это необходимость согласия с автором (или оппонентом в дискуссии) в определениях. Если такого согласия нет – все дальнейшее бесполезно. Можно спорить о фактах, о методах и т.д. – нельзя спорить об определениях. Перед любой дискуссией, любым обсуждением, надо проверить, одинаково ли понимают обсуждающие определения тех терминов, которые будут использованы в дискуссии. Иначе обсуждение будет бесплодным. Несмотря на элементарность этого правила «грамматики науки», оно все же часто нарушается.

Пример из моей собственной практики: я рассказываю об открытии интереснейшего нового (третьего) класса задач математики, физики и техники, о задачах, которые нельзя отнести ни к корректным, ни к некорректным, поскольку эти задачи меняют корректность при эквивалентных, в классическом смысле, преобразованиях, используемых при решении. Тут же, на доске, я демонстрирую примеры таких задач. Один из слушающих возражает: «Юрий Петрович, у Вас все неверно, приведенные Вами преобразования не эквивалентны, поскольку не являются взаимно однозначными». Я удивленно переспрашиваю: «А почему они должны быть взаимно однозначными? Ведь в определении, которое я привел в начале доклада, было ясно сказано, что единственным требованием, выполнение которого делает преобразование эквивалентным (равносильным) является неизменность решений преобразуемой системы. Никаких других требований нет. Кроме того, используемое мною определение общепринято – загляните в «Математическую энциклопедию», изд-е 1984 г., том 4, стр. 800. Вы найдете такое же определение без всяких требований взаимной однозначности». Возражающий упорствует: «А для меня «Математическая энциклопедия» не авторитет». Что же, в этом случае мне остается только посочувствовать и пожалеть – кто не признает определений «Математической эн-

циклопедии» и других авторитетных изданий, тот автоматически ставит себя вне большой сегодняшней науки. Согласие в определениях – первое и неперемное условие успеха своей научной работы и понимания научных работ других.

Дополнительные немалые трудности вносит неоднозначность в определениях некоторых терминов, когда разные исследователи один и тот же термин определяют по-разному. Так, термин «аналитическое конструирование» первоначально определялся как метод аналитического определения коэффициентов k_i регуляторов вида (2) для обеспечения хороших переходных процессов в линейных системах. Впоследствии стало иногда употребляться второе определение «аналитического конструирования» – как метода, позволяющего аналитически определить структуру и коэффициенты регуляторов, не обязательно вида (2), и не обязательно для линейных систем, а это уже совершенно меняет смысл термина.

Наличие нескольких определений часто приводит к недоразумениям. Так, в одном из моих докладов я рассказывал о неоднократных авариях, происходивших из-за того, что при расчетах систем управления на основе «аналитического конструирования» в 60-70-х гг. XX века еще не знали о необходимости различать преобразования, эквивалентные в классическом и в расширенном смысле, и что эти аварии подорвали репутацию методов «аналитического конструирования». Один из весьма квалифицированных слушателей моего доклада возразил: «Не могу согласиться с докладчиком, поскольку «аналитическое конструирование» регуляторов успешно используется в наше время». Недоразумение возникло потому, что я пользовался первым определением понятия «аналитическое конструирование», а возражающий мне – вторым.

Для преодоления недоразумений необходимо, чтобы участники обсуждения сверили определения используемых ими понятий и понимали их одинаково. Иначе любая дискуссия утонет во взаимных недоразумениях.

8. Что обязан и что не обязан читать и знать аспирант

Еще 10 лет назад в ответе на вопрос, «что обязан регулярно читать и что обязан знать аспирант», не было разногласий. И научные руководители, и члены ученых советов, и сами аспиранты – все были согласны с тем, что аспирант обязан просмотреть ведущие научные журналы и книги по своей специальности хотя бы за последние 5–10 лет. Он обязан уже не просмотреть, а внимательно прочесть те немногие из этих статей и книг, которые имеют непосредственное отношение к теме его диссертации. Если во время защиты обнаруживалось, что дис-

сертант не знаком с положениями ранее опубликованной статьи или книги, теми положениями, которые могли повлиять на выводы диссертации, то диссертация Ученым Советом безоговорочно отклонялась.

Такое требование обязательного знания *всей* литературы по теме диссертации отнюдь не является обременительным или чрезмерно жестким. Список ведущих журналов, относящихся к теме диссертации, очень невелик (обычно, это всего 5–6 названий), а просматривая их, аспирант сразу убеждается, что примерно 99% статей в журналах посвящены частным вопросам, не имеющим отношения к его теме. Такие статьи можно лишь бегло просматривать и откладывать в сторону. Важно лишь не пропустить и внимательно изучить те очень немногие статьи, которые либо непосредственно относятся к твоей теме, либо рассказывают о научных открытиях общего характера, не учитывать которые нельзя. То же самое относится и к книгам.

За последние 10 лет положение аспирантов осложнилось: аспирантская стипендия совершенно недостаточна для самого скромного существования. Подавляющее большинство аспирантов работает и защитить диссертацию в срок, совмещая учебу и работу, стало очень трудно.

В этих условиях некоторые научные руководители и ученые советы ослабили требования к аспирантам, исключив такое важное требование, как знакомство с научной литературой по теме диссертации. Да, я безоговорочно согласен, что положение сегодняшних аспирантов очень трудное, согласен с тем, что им надо помогать, но ведь не такой же ужасной ценой!

Пример: я спрашиваю одного из аспирантов: «Вот Вы используете математическую модель исследуемого Вами явления в виде системы дифференциальных уравнений, приводите ее перед решением к нормальной форме эквивалентными преобразованиями. Почему Вы не проверили, не является ли Ваша система особой, для которой такое преобразование недопустимо, поскольку может привести к ошибкам? Ведь такая проверка совсем не сложна, а о необходимости ее рассказано, например, в публикации [7], вышедшей 2 года назад и Вам доступной» (а также в публикациях [2], [6] и др.). Аспирант отвечает: «Научный руководитель сказал мне, что не обязательно читать все публикации, относящиеся к теме диссертации. Можно читать только те, которые нравятся». Проверяю у его научного руководителя и убеждаюсь, что такое «разрешение не читать» действительно было дано.

Я работаю в науке 50 лет (считая с момента написания первой научной работы), много чего слышал, но это меня просто испугало. Вот простое сравнение: вы приходите к врачу, нужен анализ крови и врач собирает колоть вас неодноразовым шприцем. Вы робко спрашиваете

те: «А вирус СПИДа? Я не заражусь?» Врач же вам спокойно отвечает: «Да, я слышал что-то про вирус, знаю, что о нем писали. Но разве можно всю литературу перечитать? А про СПИД мне читать не нравится, я и не читал. Так что, давайте руку, больной, сделаю Вам укол».

Как будете реагировать на такого врача? Ясно, что убежите поскорей, а затем, скорее всего, привлечете такого врача к суду, чтобы он не губил других людей. Правильно сделаете. Но почему же вы думаете, что менее опасным будет пренебрежение аспиранта к новым научным результатам, нежелание (причем, одобренное его научным руководителем!) знакомиться с ними. Ведь завтра этот аспирант, став кандидатом наук, будет участвовать в проектировании и расчете, ну, например, самолета. Привыкнув игнорировать «не нравящиеся» ему научные результаты, он будет делать то же самое и при проектировании. Самолет окажется дефектным, потерпит аварию, загубит экипаж и пассажиров. Кто будет больше виноват? Научный руководитель? Или все-таки сам аспирант? Ведь никто не мешал ему знакомиться с «грамматикой науки», а точнее – с ее неперменными требованиями о том, что должен знать аспирант, требованиями, выработанными на опыте поколений. Литературы по истории науки достаточно, читать ее никто не запрещает.

Могут, конечно, возразить: ошибка участника проектирования самолета совсем не обязательно приводит к аварии и гибели людей, она лишь увеличивает (может быть даже немного) вероятность аварии. Да, это так. Но ведь и укол многократным шприцем совсем не обязательно (и даже редко) ведет к смерти пациента. Однако врача за такой укол судят, и правильно делают. А вот аспиранта (и его научного руководителя) мы не осуждаем. И это напрасно.

Вот печальный недавний пример: в мае 2002 г. в Ученом Совете факультета прикладной математики – процессов управления СПбГУ С.А. Сумачевым защищалась диссертация по исследованию газотурбинного двигателя для самолетов. Двигатель содержал несколько контуров управления, и диссертант рассчитывал их устойчивость традиционными, устаревшими методами, давно доказавшими свою неполноту. Диссертант явно не знал, что неполнота этих методов в 1994 году была доказана мной [1], в 1997 году – профессорами А.Р. Гайдуком [8] и В.А. Подчукаевым [9], в 2000 году – академиком РАН Я.Б. Данилевичем [7]. При этом работы этих авторов были опубликованы в известных авторитетных научных журналах, безусловно, доступных диссертанту. В этих публикациях приводились методы дополнительных проверок, обеспечивающих достоверность расчета. Диссертант мог воспользоваться любым из этих методов – либо моим, либо методом профессора А.Р. Гайдука, либо иным. Он не воспользовался ни одним. В

результате выводы диссертации оказались недостоверными, необходимое требование Высшей Аттестационной Комиссии (ВАК) о достоверности результатов оказалось не выполненным, о чем говорилось и на защите. В данном случае Ученый Совет все же присудил диссертанту степень, хотя и рисковал, что решение будет отменено во время контрольной проверки диссертации в ВАК. Рассчитывали на то, что в последние годы ВАК отправляет на контрольную проверку лишь небольшую долю кандидатских диссертаций, так что риск был не очень велик.

Непонятно другое – зачем надо было Ученому Совету идти на неприятный риск? Не проще ли было потребовать с аспиранта (лучше всего на стадии предварительного обсуждения диссертации) проведения дополнительной проверки по одной из опубликованных [2], [6], [8], [9] методик? Тогда и риска не было бы.

Мы убеждаемся, что игнорирование многих общепринятых ранее правил науки (составляющих в совокупности ее «грамматику») стало в последние годы довольно частым явлением и это не может не огорчать.

9. Как получить достойную зарплату за свой труд?

Сегодняшняя зарплата научных работников России очень низка, недопустимо, нищенски низка. То, что ее необходимо крупно (в разы) повышать, нет никаких сомнений. Это бесспорно! Но, уважаемые мои коллеги, научные работники, давайте будем совершенно откровенны: первосортную зарплату естественно требовать за науку первого сорта. А если мы ради облегчения себе жизни потихоньку игнорируем сегодня одно правило «грамматики науки», завтра – другое правило, то во что же быстро превращается наша наука? Она превращается в «науку второго сорта», за которую и зарплата будет соответствующая – второсортная.

Если хотим получать хорошую зарплату, достойное материальное обеспечение, то лучший путь к этому – заниматься наукой «первого сорта», не поддаваться на соблазны облегчить себе жизнь, проигнорировать пусть обременительные, но необходимые правила «грамматики науки». Ведь эти правила появлялись не случайно, они выросли из опыта столетий развития научных исследований – правила обязательного знания научной литературы предшествующих лет, обязательного отказа от доказательства, опровергнутого контрпримером, согласия в определениях, печатного (в крайнем случае, письменного) оформления научной дискуссии.

Отказ от хотя бы одного из этих правил ведет к потере научной репутации. Если один научный сотрудник позволяет себе отказаться от

общепризнанных правил науки, то погибнет репутация этого конкретного человека и не более того, хотя для него самого это трагедия. Потерять репутацию легко – восстановить очень трудно.

Гораздо хуже, если в стенах какого-нибудь научного учреждения игнорирование правил науки не встречает отпора, становится привычным, часто встречающимся. Тогда может сильно пострадать репутация научного учреждения в целом.

Я, разумеется, рассказал далеко не обо всех правилах «грамматики» науки, остановив внимание на тех, которые нарушались чаще всего на моих глазах.

Не мешает отметить еще и такое важное правило, как возможность подтверждения или опровержения. Высказанное кем-либо утверждение остается гипотезой до тех пор, пока оно не будет:

- 1) доказано (вариант – подтверждено ясными экспериментами);
- 2) опровергнуто.

Любопытно, что и это простое правило иной раз бывает нарушено. Пример: астрологические прогнозы. Вы часто можете слышать прогнозы типа: «Родившихся под знаком Овна на этой неделе ожидает успех в делах», «У родившихся под знаком Рака на этой неделе возможны неприятности, старайтесь воздерживаться от принятия решений». Но тщетно вы будете искать подтверждения прогнозов типа: «На прошлой неделе мы проследили за 100 предпринимателями, родившимися под знаком Овна. У 80 из них наблюдались успехи в делах. А вот из проверенных 100 человек, родившихся под знаком Рака, у 75 на прошедшей неделе были неприятности».

Вы нигде не найдете подобных проверок астрологами своих «прогнозов». Причина проста: когда подобные проверки проводили независимые исследователи, результаты неизменно заканчивались конфузом: прогнозы не сбывались. Тогда астрологи отказались от проверок и от самого принципа проверяемости. Свои прогнозы астрологи постоянно публикуют, а сбываются «прогнозы» или нет – это их не интересует.

Но тогда понятно, что отказавшись от такого важнейшего правила «грамматики науки» как принцип проверяемости, астрология перешла уже даже не на положение «науки второго сорта», а в «науку вне сортов», или, коротко, на положение лженауки. Таковой она остается и до настоящего времени.

Пренебрежение теми правилами «грамматики науки», о которых я рассказывал в предыдущих разделах, не имеет таких катастрофических последствий, но все же тот, кто пренебрегает ими, превращается в научного работника «второго сорта» с соответствующим к нему отношением, а в дальнейшем и с соответствующей зарплатой.

Лекция третья

10. Объяснение парадоксов

В этой лекции я постараюсь дать наиболее простое и наглядное объяснение недавно открытому явлению, о котором говорилось в предыдущих лекциях – изменению некоторых важных характеристик математической модели при эквивалентных, в классическом смысле, преобразованиях ее. При преобразованиях, не изменяющих ее решений, но меняющих некоторые их свойства. Парадоксальность этого явления заключается в том, что эквивалентными (их называют также равносильными) преобразованиями пользовались сотни лет тысячи математиков и никто не замечал и не верил, что эти, привычные еще со средней школы преобразования могут что-то менять.

Внимательное исследование показало, что эквивалентные преобразования могут изменять параметрическую устойчивость, непрерывную зависимость решений от коэффициентов и параметров. То есть, обобщая, они могут изменять корректность решаемой задачи. Парадокса здесь нет, объяснение недавно открытого явления совсем не сложно.

Действительно, рассмотрим внимательнее, что означает, например, безусловно верное утверждение:

«Решение дифференциального уравнения

$$a_1 \dot{x} + a_0 x = 0 \quad (5)$$

непрерывно зависит от коэффициентов a_1 и a_0 » ?

Ведь фактически это совсем не является утверждением о самом уравнении (5). По существу, это является утверждением о свойствах решений *семейства уравнений*, семейства вида

$$a_1(1 + \varepsilon_1)\dot{x} + a_0(1 + \varepsilon_0)x = 0, \quad (6)$$

где ε_1 и ε_0 – любые числа, малые по сравнению с единицей. Это семейство является окрестностью уравнения (5), окрестностью в пространстве коэффициентов.

Утверждение о непрерывной зависимости решений уравнения (5) от коэффициентов a_1 и a_0 справедливо тогда и только тогда, когда все решения семейства (6) при любых, сколь угодно малых ε_1 и ε_0 будут сколь угодно мало отличаться от решения исходного уравнения (5).

Теперь сразу делается ясным, что эквивалентные преобразования исходного уравнения (5), которые, по определению, не меняют его (и только его!) решений, совсем не обязаны не изменять решений всего большого семейства уравнений (6). Эквивалентное преобразование может, конечно, не изменить эти решения, но может и изменить. Парадокса (или противоречия) в этом изменении нет. (Впервые это рассуждение, обосновавшее возможность существования преобразований, эквивалентных в классическом смысле, но не в расширенном, было

опубликовано в [4]. Там же приведены и первые примеры подобных преобразований.) А обнаружены эти преобразования были так поздно просто потому, что они встречаются редко. Большинство преобразований, эквивалентных в классическом смысле, эквивалентны и в расширенном. Особые преобразования, эквивалентные в классическом смысле, но не в расширенном, преобразования, изменяющие корректность задачи, а в частных случаях изменяющие параметрическую устойчивость, непрерывную зависимость решений от параметров и т.п. (все это – частные случаи корректности), встречаются редко. Однако это совсем не означает, что им не нужно уделять самого серьезного и пристального внимания. Ведь каждая неожиданная, непредвиденная встреча с таким преобразованием может стать причиной ошибки в расчете, а значит – причиной аварии. Об этом подробно рассказано в публикациях [2, 6], где приведены примеры аварий, произошедших по этой причине. Поэтому необходимо знать о существовании преобразований, эквивалентных в классическом смысле, но не в расширенном. А еще лучше не просто знать, а подробно изучить их свойства, опираясь на известные публикации [1, 2, 6, 7]. Это уберезет от множества ошибок.

Теперь рассмотрим еще раз причину, почему было трудно открыть существование преобразований, эквивалентных в классическом смысле, но не в расширенном, и почему они были открыты так поздно.

Они были открыты в 1987–1994 гг. в ходе исследования проблемы оптимизации управления для различных технических объектов.

Рассмотрим пример – систему, состоящую из объекта управления (электропривода постоянного или переменного тока) и регулятора, поддерживающего постоянство частоты вращения электропривода. Математическая модель электропривода, работающего на механизм, колебания момента сопротивления которого являются случайным стационарным процессом с дробно-рациональным спектром второго порядка, описываются, как известно, системой трех дифференциальных уравнений с четырьмя переменными:

$$\left. \begin{aligned} m\dot{x}_1 &= a_1x_1 + a_2x_2 + a_3x_3, \\ \dot{x}_3 &= a_4x_4, \\ \dot{x}_4 &= a_5x_3 + a_6x_4. \end{aligned} \right\} \quad (7)$$

В этих уравнениях x_1 – это частота вращения (точнее, отклонение частоты вращения от ее номинального значения), x_2 – вращающий момент электропривода, играющий роль управления, x_3 – отклонение момента сопротивления исполнительного механизма электропривода от номинального значения, x_4 – производная переменной x_3 , m – важнейший параметр электропривода, его механическая постоянная времени,

численно равная времени разгона электропривода от нуля до номинальной частоты вращения при номинальном вращающем моменте и равном нулю моменте сопротивления. Коэффициенты a_1, a_2, \dots, a_6 принимают различные числовые значения у разных электроприводов. Мы в дальнейшем рассмотрим электропривод со значениями коэффициентов $a_1 = -2, a_2 = 1, a_3 = 1, a_4 = 1, a_5 = -1, a_6 = -2$, т.е. рассмотрим электропривод, математическая модель которого имеет вид:

$$\left. \begin{aligned} m\dot{x}_1 &= -2x_1 + x_2 + x_3, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= -x_3 - 2x_4. \end{aligned} \right\} \quad (8)$$

Коэффициенты математической модели (8) близки к коэффициентам одного из реальных электроприводов, но для удобства дальнейших расчетов они округлены до целых чисел. Время t мы будем в дальнейшем выражать в долях от механической постоянной времени номинального режима. Это означает, что в номинальном режиме будет $m=1$, но мы будем учитывать также и возможные вариации параметра m . Так, при изменениях температуры изменяются линейные размеры всей деталей электропривода и при этом неизбежны малые изменения (вариации) параметра m .

В три уравнения (8) входят четыре переменные, поэтому процессы в электроприводе получают определенность при учете регулятора – устройства, которое вырабатывает управляющее воздействие x_2 на основе измерения остальных переменных. Выберем регулятор, математической моделью которого является уравнение:

$$x_2 = -x_1 - 2x_3 - x_4. \quad (9)$$

Его можно рассматривать как вырожденное дифференциальное уравнение, превратившееся в конечное соотношение между переменными.

Система уравнений (8)–(9) описывает процессы, протекающие в системе: «объект управления – регулятор».

Характеристический полином системы уравнений (8)–(9) равен определителю

$$\begin{vmatrix} m\lambda + 2 & 1 & 1 & 0 \\ 0 & 0 & \lambda & -1 \\ 0 & 0 & 1 & \lambda + 2 \\ 1 & 1 & 2 & 1 \end{vmatrix} = (m\lambda + 3)(\lambda + 1)^2, \quad (10)$$

имеющему корни: $\lambda_1 = -3/m, \lambda_2 = \lambda_3 = -1$.

При $m=1$ (т.е. в номинальном режиме) будет $\lambda_1 = -3, \lambda_2 = \lambda_3 = -1$; общее решение системы (8)–(9) имеет вид:

$$x_1 = c_1 e^{-3t} + (c_2 t + c_3) e^{-t}. \quad (11)$$

При малых отклонениях параметра m от номинального значения $m=1$ решения изменятся мало. Таким образом, мы убеждаемся, что решения системы (8)–(9) параметрически устойчивы (т.е. сохраняют устойчивость при вариациях параметра m) и непрерывно зависят от m .

Измерение и использование в регуляторе переменной x_3 и особенно переменной x_4 затруднительно и поэтому желательно исключить переменные x_3 и x_4 из системы (8)–(9) с помощью эквивалентных преобразований.

Выполнив исключение переменных x_3 и x_4 по общим правилам, приведем систему (8)–(9) к виду:

$$[mD^3 + (3 + 2m)D^2 + (6 + m)D + 3] x_1 - (D^2 + 2D + 1) x_2 = 0, \quad (12)$$

$$[mD^2 + (2 + 2m)D + 5] x_1 - (D + 1) x_2 = 0. \quad (13)$$

Уравнение (12) – это уравнение объекта управления, уравнение (13) – это уравнение регулятора. Характеристический полином системы уравнений (12)–(13) равен определителю

$$\begin{vmatrix} m\lambda^3 + (3 + 2m)\lambda^2 + (6 + m)\lambda + 3 & -(\lambda^2 + 2\lambda + 1) \\ m\lambda^2 + (2 + 2m)\lambda + 5 & -(\lambda + 1) \end{vmatrix} = (m\lambda + 3)(\lambda + 1)^2, \quad (14)$$

т.е. будет тем же самым, что и у системы (8)–(9). Поэтому мы должны, казалось бы, сделать вывод о том, что и после исключения переменных x_3 и x_4 рассматриваемая нами система уравнения будет параметрически устойчивой, и ее решения будут непрерывно зависеть от параметра m .

Но этот вывод будет ошибочным! Делая его, мы не учли, что параметры регулятора вполне могут не зависеть от вариаций механической постоянной времени электропривода. Действительно, регулятор – это отдельное устройство. Оно стоит рядом с электроприводом и вариация механической постоянной времени на его коэффициенты вполне может совсем не влиять. Поэтому более правильным и более соответствующим физическому смыслу будет для получения уравнения регулятора исключать переменные x_3 и x_4 из уравнений (8) и (9), в которых предварительно принято, что $m=1$. При этом мы получаем уравнение регулятора в виде:

$$(D^2 + 4D + 5) x_1 - (D + 1) x_2 = 0. \quad (15)$$

Оно совпадает с уравнением (13) при $m=1$.

Теперь его коэффициенты, что соответствует физическому смыслу, не зависят от вариаций параметра m в электроприводе, и поэтому уравнение (15) более правильно, чем уравнение (13), отражает истинные процессы, происходящие в системе «электропривод – регулятор» при вариациях параметра m .

В целом, система уравнений (12)–(15) более правильно описывает реальные процессы в системе управления, по сравнению с уравнениями (12)–(13). Теперь рассмотрим, в чем же заключаются различия. Характеристический полином системы (12)–(15) равен определителю

$$\begin{vmatrix} m\lambda^3 + (3 + 2m)\lambda^2 + (6 + m)\lambda + 3 & -(\lambda^2 + 2\lambda + 1) \\ \lambda^2 + 4\lambda + 5 & -(\lambda + 1) \end{vmatrix} = \quad (16) \\ = [(1 - m)\lambda^2 + (2 - m)\lambda + 3](\lambda + 1)^2.$$

При $m=1$ определители (14) и (16) совпадают. Это означает, что при $m=1$ решения системы (12)–(15) тоже имеют вид (11), и поэтому исключение переменных x_3 и x_4 из уравнений (8)–(9), с учетом реальной физической картины отсутствия связи между вариациями параметров электропривода и регулятора, является преобразованием, эквивалентным в классическом смысле. В то же время, эквивалентности в расширенном смысле нет, поскольку при $m=1+\varepsilon$, где $\varepsilon>0$, характеристический полином (16) получает четвертый большой положительный корень. Это означает, что в решениях системы (12)–(15) при $\varepsilon>0$ появляется быстрорастущий член вида $c_4 e^{t/\varepsilon}$ – растущий тем быстрее, чем меньше ε . Это говорит о том, что система (12)–(15) теряет устойчивость при сколь угодно малом отклонении параметра m от расчетного номинального значения $m=1$. Ни параметрической устойчивостью, ни непрерывной зависимостью решений от параметра m система уравнений (12)–(15) не обладает, в отличие от уравнений (8)–(9).

Реальное поведение системы управления, в которой регулятор формируется на основе измерений и преобразований переменной x_1 с исключенными переменными x_3 и x_4 , соответствует уравнениям (12)–(15), а не (12)–(13). Реальная система, как легко проверить на опыте, параметрической устойчивостью и непрерывной зависимостью решений от параметров не обладает.

Если же мы приведем систему (12)–(15) к нормальной форме Коши, то мы придем к уравнениям (8)–(9). Исследуя их, мы придем к ошибочному выводу о том, что исследуемая система управления якобы обладает параметрической устойчивостью и непрерывной зависимостью решений от параметра m , хотя сами решения, как таковые, при номинальном значении $m=1$ после перехода к нормальной форме Коши будут вычислены правильно.

Теперь мы получаем окончательный вывод: преобразование системы дифференциальных уравнений к нормальной форме Коши является все же «несколько грубоватым» преобразованием. Сами решения системы оно, безусловно, сохраняет, но некоторые тонкие (но важные!) свойства решений, такие, как параметрическая устойчивость и

непрерывная зависимость параметров, это «немного грубоватое» преобразование к нормальной форме может не сохранять. И, как показывает рассмотренный пример, иногда (хотя и не всегда!) действительно не сохраняет.

Для правильного установления многих важных свойств решений необходимо пользоваться преобразованиями, эквивалентными не только в классическом, но и в расширенном смысле. Для этого необходимо знать свойства этих преобразований, которые, к сожалению, много сложнее, чем простые свойства преобразований, эквивалентных в классическом смысле. Некоторые интересные свойства преобразований, эквивалентных в расширенном смысле, описаны в [2, 6], но исследование их еще далеко не закончено.

Становится понятным, почему преобразования, эквивалентные в классическом смысле, но не в расширенном, сохраняющие решения, но изменяющие ряд тонких свойств решений были открыты так поздно. Хотя свойства преобразований – это чисто математический вопрос, но правильное решение его требует, как мы убедились, четкого понимания характера взаимодействия различных элементов и узлов технических устройств, математические модели которых мы исследуем.

Становится так же понятным, почему часть наиболее консервативно настроенных математиков до последних лет (а некоторые и до сих пор!) отвергала существование преобразований, эквивалентных в классическом смысле, но не в расширенном. Им казалось неправомерным, при обсуждении математического вопроса, применение технических соображений (независимость вариаций параметров регулятора от параметров электродвигателя), которые использовались в [2] при выводе соотношений (12)–(15). Однако позже реальность преобразований, не эквивалентных в расширенном смысле, была показана уже и на чисто математических примерах (об этом будет рассказано далее).

11. Важные следствия

Обнаружение преобразований, эквивалентных в классическом смысле, но не в расширенном, повлекло за собой важные следствия (частично мы уже говорили о них в предыдущих лекциях).

1. Обнаружилась неполнота традиционных и широко применяемых методов расчета устойчивости. Оказалось, что они требуют совершенствования и уточнения, а иначе могут сделаться причиной опасных аварий. Об этом мы уже говорили.

2. Выявились, что для ряда систем дифференциальных уравнений традиционные методы вычисления решений через стандартные программы с использованием преобразования системы к нормальной форме Коши ведут к ошибкам. Для предотвращения ошибок нужно перед

расчетом проверять – не является ли рассматриваемая система особой. Об этом мы уже тоже говорили.

3. Третье следствие относится к тем многочисленным методам вычислений, которые используют цепочки эквивалентных преобразований. Так, наиболее распространенный метод вычисления определителей высокого порядка заключается в последовательном понижении порядка. Определитель, например, седьмого порядка преобразуют в определитель шестого порядка, потом – пятого и т.д. эквивалентными преобразованиями (домножениями и сложениями).

Теперь посмотрим, что будет, если среди этой цепочки преобразований окажется хотя бы одно преобразование, изменяющее корректность решаемой задачи? Тогда эту задачу нельзя уже отнести ни к корректным, ни к некорректным. Подобные задачи в 1998 году, в публикации [10] было предложено выделить в отдельный, третий класс – класс задач, меняющих свою корректность при эквивалентных преобразованиях, используемых в ходе их решения.

Открытие нового класса задач имеет как теоретическое, так и прикладное значение. С теоретической точки зрения оно интересно тем, что ранее, начиная с открытия Адамаром в 1902 году класса некорректных задач, считалось, что все задачи математики, физики техники делятся на два класса – на класс корректных и класс некорректных задач. В 1998 году выяснилось, что существует еще один своеобразный класс – класс «задач-перебежчиков», перебегающих из класса корректных в класс некорректных (или, наоборот) при эквивалентных преобразованиях, используемых при их решении.

С практической точки зрения открытие третьего класса задач интересно тем, что оно раскрывает еще один источник возможных ошибок при вычислениях. Действительно, пусть в ходе цепочки эквивалентных преобразований исходной математической модели изменилась корректность задачи, и она стала не корректной. Тогда даже сколь угодно малая погрешность в вычислениях (например, погрешность округления) может коренным образом изменить результат решения, привести к серьезной ошибке.

Покажем все это на примере известной задачи о вычислении собственных значений системы однородных линейных уравнений с параметром.

Рассмотрим, систему из трех уравнений с параметром λ :

$$\left. \begin{aligned} (1-\lambda)x_1 + x_2 + 2x_3 &= 0, \\ x_1 + (1-\lambda)x_2 + 3x_3 &= 0, \\ x_1 + x_2 &= 0 \end{aligned} \right\} \quad (17)$$

(в последнее уравнение параметр λ не входит).

Поскольку все уравнения однородны, то система, безусловно, имеет тривиальное нулевое решение $x_1=x_2=x_3=0$, но при некоторых значениях параметра λ возможны не нулевые решения. Такие значения параметра λ называют собственными значениями, а вычисление собственных значений является, как известно, одной из наиболее важных задач вычислительной математики и имеет многочисленные приложения в физике и технике.

Собственные значения системы (17) являются корнями определителя

$$\begin{vmatrix} (1-\lambda) & 1 & 2 \\ 1 & (1-\lambda) & 3 \\ 1 & 1 & 0 \end{vmatrix} = 5\lambda, \quad (18)$$

откуда сразу видно, что единственным собственным значением является $\lambda=0$. Для простой системы (17) вычислить определитель несложно, и задача его вычисления вполне корректна, однако, для систем, состоящих из большого числа уравнений, прямое вычисление затруднено. Одним из возможных методов вычисления является последовательное сокращение числа уравнений и числа переменных в порядке их индексов.

Применим этот метод к системе (17). Помножив второе из уравнений (17) на $-(1-\lambda)$ и сложив с первым, мы исключим x_1 и придем к уравнению:

$$(\lambda^2 - 2\lambda)x_2 + (1 - 3\lambda)x^3 = 0. \quad (19)$$

Вычитая из второго уравнения (17) третье, получим:

$$\lambda x_2 - 3x_3 = 0. \quad (20)$$

Мы пришли к системе двух уравнений (19) и (20) с двумя переменными x_2 и x_3 . Исключив x_3 , придем к одному уравнению

$$5\lambda x_2 = 0, \quad (21)$$

откуда сразу находим единственное собственное значение $\lambda=0$.

Однако внимательное исследование обнаруживает, что задача вычисления собственных значений для системы уравнений (19)–(20) некорректна.

Действительно, пусть мы сделали малую ошибку в вычислениях и получили в уравнении (20) коэффициент перед x_3 не 3, а $3(1+\epsilon)$. Тогда после исключения x_3 из уравнений (19) и (20) мы получим вместо уравнения (21) уравнение:

$$(3\epsilon\lambda^2 - 6\epsilon\lambda - 5\lambda)x_2 = 0, \quad (22)$$

из которого находим два собственных значения:

$$\lambda_1 = 0, \quad \lambda_2 = 2 + 5/3\epsilon. \quad (23)$$

Второе (ложное) собственное значение будет присутствовать при любых, сколь угодно малых ε и при $\varepsilon \rightarrow 0$ совсем не приближается к первому, что говорит о некорректности задачи вычисления собственных значений для системы (19)–(20).

Сколь угодно малая погрешность в вычислении всего лишь одного коэффициента приводит к принципиальной ошибке в решении – вместо одного собственного значения получаем два различных, не близких друг к другу.

Таким образом, мы привели пример преобразования, эквивалентного в классическом смысле, но не в расширенном, причем на этот раз без привлечения физических соображений, пользуясь лишь математикой.

Теперь рассмотрим, какие практические следствия вытекают из недавно обнаруженного существования подобных преобразований. Из этого следует, что в любом вычислительном алгоритме, состоящем из цепочки эквивалентных преобразований (а таких вычислительных алгоритмов очень много) необходимо каждое преобразование проверять на эквивалентность в расширенном смысле. Если такой проверки не делать, то возможны любые ошибки в решениях. Действительно, если хотя бы одно преобразование оказалось эквивалентным в классическом смысле, но не в расширенном, то уже сколь угодно малая, неизбежная на практике погрешность в вычислениях (например, ошибка округления) может привести к самой грубой ошибке в окончательном результате вычисления и стать причиной тяжелой аварии, и даже катастрофы. Примеры приведены в [2]. Отсюда следует важность знания теории преобразований, эквивалентных в расширенном смысле.

Заметим, что когда мы говорим, что отсутствие проверки на эквивалентность в расширенном смысле *может* привести к грубой ошибке, то это не означает, что погрешность *обязательно* появится. Некоторые вычислительные алгоритмы безопасны. Так, исследование, проведенное в [2] показало, что в классической задаче о вычислении собственных значений, когда каждое из уравнений системы содержит параметр λ в первой степени, потери корректности при понижении порядка не происходит. В то же время в не менее часто встречающейся обобщенной задаче вычисления собственных значений, когда часть уравнений не содержит λ , потеря корректности возможна и происходит достаточно часто.

Проверка показала, что потеря корректности и связанные с ней ошибки (появление лишних собственных значений, зависящих от сколь угодно малых погрешностей округления) действительно систематически возникают у систем уравнений выше четвертого порядка и обычные методы борьбы с погрешностями округления, такие как пере-

программирование вычислительной машины на удвоенную точность, в данном случае не помогают.

Теперь становится понятным, почему существование преобразований, эквивалентных в классическом смысле, но не в расширенном, было обнаружено так поздно. При ручном счете в обобщенной задаче о собственных значениях было, безусловно, удобнее сначала исключить часть переменных, пользуясь уравнениями, не содержащими параметра λ . После этого обобщенная задача переходит в классическую, алгоритм решения которой безопасен. И только после широкого распространения программируемых вычислительных машин стало более удобным решение по единой программе, с исключением переменных в порядке их индексов. Только тогда ранее на замечавшиеся преобразования, эквивалентные в классическом смысле, но не в расширенном, показали все свое коварство и были обнаружены.

Общий вывод: распространение вычислительной техники требует дополнительной проверки и ревизии используемых алгоритмов. Алгоритмы, столетиями надежно применявшиеся при ручном счете, могут проявить новые свойства при машинных вычислениях.

12. Советы и рекомендации

Завершая цикл лекций, хочу посоветовать относиться более внимательно к «грамматике» науки – правилам ведения научных дискуссий, методам доказательства и опровержения теорем и т.д. Неоценимую пользу может принести знакомство с историей науки, которая, в настоящее время, преподается, к сожалению, далеко не во всех высших учебных заведениях. Однако тот, кто не прослушал курса истории своей науки в ВУЗе, может прочесть самостоятельно любые из многочисленных хороших книг по этому вопросу. Из недавно изданных и имеющихся в продаже отметим книги [11], [12], в библиографии которых вы найдете обширный перечень ранее изданных книг. В [11] проанализирована, в частности, история развития теории автоматического управления и регулирования. Она рассмотрена как один из разделов прикладной математики; там же, в [11] изложена история исследования класса некорректных задач (начиная с Адамара) и рассказано о предпосылках открытия задач промежуточных между корректными и некорректными.

Для тех, кто ищет интересные темы для научной работы, я бы очень рекомендовал исследовать круг вопросов, связанных с преобразованиями, эквивалентными в классическом, но не в расширенном смысле, и с задачами, относящимися к третьему классу, промежуточному между ранее известными классами корректных и классом некорректных задач математики, физики и техники. Это – новая область.

Область недавнего научного прорыва, где еще очень многое не исследовано, остается в «тумане», где исследования, по существу, еще только начались. В такой сфере, где все еще только складывается, работать приятно и интересно. Она ждет своих молодых и энергичных исследователей. Необходимо, прежде всего, пополнить список задач третьего класса: На сегодняшний день известны и исследованы лишь очень не многие, на деле их гораздо больше, а между тем каждая непредвиденная, неожиданная встреча с подобной задачей может обернуться для инженера серьезным просчетом.

Недостаточно исследованы свойства преобразований, эквивалентных в расширенном смысле. Уже первые итоги в их исследовании показали, что эти свойства очень интересны, во многом – парадоксальны (см. [2], стр.120–125). Безусловно, при дальнейшем изучении можно ожидать новые сюрпризы.

Для тех, кто не будет заниматься исследованиями, для тех, кто интересуется, прежде всего, практическими приложениями, вытекающими из недавно сделанных научных открытий, следует отметить, что эти открытия, к сожалению, не облегчают, а усложняют жизнь инженера. Раньше он был уверен, что если у проектируемой им системы все корни характеристического полинома лежат в левой полуплоскости комплексного переменного далеко от мнимой оси, или же, если для системы построена функция Ляпунова, то устойчивость и параметрическая устойчивость обеспечены. Теперь мы видим, что это не так, и чтобы гарантировать параметрическую устойчивость, необходимы дополнительные расчеты, дополнительная работа. Точно так же инженеры и физики были уверены, что приведение системы уравнений к нормальной форме, необходимое для использования стандартных программ, вполне справедливо и ничего в решениях не меняет. Теперь мы видим, что и это не так, что необходимы дополнительные вычисления, дополнительная работа, для того, чтобы проверить правильный ли ответ дает выбранная математическая модель, не изменились ли при преобразовании к нормальной форме некоторые тонкие, но важные свойства решений. Понятно, что дополнительную работу делать никому не хочется, и поэтому результаты, опубликованные в [2, 6], еще не получили широкого прикладного применения. Однако, несложные дополнительные проверки, описанные в [2] и в [6] страхуют от ошибок, уменьшают вероятность аварий. Постепенно это начинает осознаваться (см. публикации [7, 8, 9] и совсем недавнюю статью [13]). Поэтому рано или поздно (и скорее – рано) использовать дополнительные проверки придется всем. Причем фирма, первая внедрившая у себя дополнительные проверки, может с полным основанием утверждать, что вероятность аварий в выпускаемой ею продукции меньше, чем у фирм-

конкурентов; это даст такой фирме преимущество в конкурентной борьбе за заказы, а этим пренебрегать нельзя.

ЛИТЕРАТУРА к «ПРИЛОЖЕНИЮ»

1. *Петров Ю.П.* Устойчивость линейных систем при вариациях параметров // Автоматика и телемеханика, 1994, № 11, с. 186–189.
2. *Петров Ю.П.* Неожиданное в математике и его связь с авариями и катастрофами последних лет. – СПб: Изд-во СПбГУ, 1-е изд. – 1999, 108 с.; 3-е изд. – 2002, 141 с.
3. *Летов А.М.* Аналитическое конструирование регуляторов // Автоматика и телемеханика, 1960, № 4, 5, 6.
4. *Петров Ю.П.* Синтез оптимальных систем управления при не полностью известных возмущающих силах. Уч. пособие. – СПб: Изд-во ЛГУ, 1987. 289 с.
5. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. 3-е изд. – М.: Наука, 1986. 288 с.
6. *Петров Ю.П.* Новые главы теории управления. Уч. пособие. – СПб: Изд-во СПбГУ, 2000. 156 с.
7. *Данилевич Я.Б., Петров Ю.П.* О необходимости расширения понятия эквивалентности математических моделей // Докл. РАН, 2000, т. 371, № 4, с. 473–475.
8. *Гайдук А.Р.* К исследованию устойчивости линейных систем // Автоматика и телемеханика, 1997, № 3, с. 153–160.
9. *Подчукаев В.А.* К проблеме грубости // Сборник «Аналитические методы синтеза регуляторов». – Саратов, 1997, с. 206–235.
10. *Петров Ю.П.* Третий класс задач физики и техники – промежуточных между корректными и некорректными. – СПб: Изд-во СПбГУ, 1998. 30 с.
11. *Петров Ю.П.* Лекции по истории прикладной математики. – СПб: Изд-во СПбГУ, 2001. 337 с.
12. *Рыбников К.А.* История математики. Учебник. – М.: Изд-во МГУ, 1994. 496 с.
13. *Тарарыкин С.В., Тютиков В.В.* Робастное модальное управление динамическими системами // Автоматика и телемеханика, 2002, № 5, с. 41–55.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Аварии 22, 68, 232
Аналитическое конструирование оптимальных регуляторов 44
Априорная информация о решении 18, 178
«Белый шум» 88, 90, 152
Вариации (погрешности) номинальных коэффициентов 8, 12, 68
«Вариация нуля» 12
Вектор-столбец 123, 172
Вектор-строка 123
Вектор-функция 141, 174
Воздействие возмущающее 19
– управляющее 19
Восстановление дефокусированных изображений 198, 203
Восстановление смазанных изображений 188, 194–197
Вырожденные и плохо обусловленные СЛАУ 10, 129, 172
Гармоника Фурье 76, 152, 166
Гладкость решения 161, 172
Датчик случайных чисел RNDAN 164, 183
 δ -функция 81, 104, 106, 221
Дефокусированное изображение 198
Динамическая система 90
Дискретизация 149, 164, 181
Дифференциальное уравнение с постоянными коэффициентами 19
Дифференцирование функций с помехами 94, 144
Единственность решения 138
Задача корректная (well-posed) по Адамару 138
– – по Тихонову (условно корректная) 160
– – по Фикера 138
– Коши 132, 142, 178
– некорректная (ill-posed) 138
– неустойчивая 133, 143
– обратная 164, 188, 190, 213
– плохо обусловленная 60
– прямая 164, 189
– рентгеновской томографии 204
– синтеза магнитного поля в ЯМР-томографе 212
– существенно некорректная 160
– условно корректная (корректная по Тихонову) 160
– устойчивая 161
Задачи, промежуточные между корректными и некорректными 41, 49
– синтеза оптимальных систем управления 19
Изображение 188
– по Лапласу 66, 136
– по Фурье 135
Интегральное (непрерывное) преобразование Лапласа 136

- Фурье (НПФ) 135
- Хартли 136
- Искомая функция 148
- Исходный процесс 135
- Качество управления 27
- Квадратурная формула 149, 163
- Классические методы решения интегральных уравнений 148
- Компакт 178
- Контраст изображения 162, 192
- Корректность и некорректность по Адамару (классическая) 138
 - по Тихонову (условная) 160
 - по Фикера 138
 - интегральных преобразований 83
- Корреляционная функция 73
- Косинус-преобразование Фурье 70, 135
- Коэффициенты математической модели (номинальные и реальные) 6, 8, 12, 31, 57
 - уравнений 8, 19, 29
- Линейная алгебра 122
- Локальный носитель supp 168
- Матрица 123
 - полиномиальная 30
- Метод квадратур 148, 163
 - локальной регуляризации 208
 - наименьших квадратов (МНК) Гаусса 128, 152
 - неопределенных множителей Лагранжа 160
 - преобразования Фурье 151, 191, 201
 - псевдообратной матрицы (МПОМ) Мура–Пенроуза 128, 152
 - регуляризации Тихонова 159, 166, 191, 201, 208
- с ограничениями на решение 218
 - решения на компакте 178
 - синтеза систем управления 44
- Минимизация функционала 160, 166, 170, 173, 179
- Множество 116
 - корректности 160
 - функций монотонных ограниченных 179
 - выпуклых 181
 - ограниченной вариации 182
 - ограниченных выпуклых 181
- Напряженность магнитного поля 214
- Невязка 128, 153
- Неоднородность магнитного поля 213
- Неопределенный множитель Лагранжа 161
- Непрерывная зависимость решения от начальных данных, от параметров, от правой части и т.д. 31
- Неустойчивость решения 150
- Норма (вектора, матрицы, оператора, решения) 118, 125
- Нормальная СЛАУ 128, 153
 - форма Коши системы ОДУ 32, 130
- Нормальное псевдорешение 157
 - распределение 164, 168
 - решение 156, 173
- Обратная связь 55
- Ограничения на решение 180–182, 218
 - на управляющее воздействие 20
- Оператор 121

- дифференцирования $D = \frac{d}{dt}$ 95
- обратный 122
- Определитель 123
- Пакет программ IMAGE 192
- Параметр m 32
- регуляризации α 161, 173, 209
- Параметрическая устойчивость 22, 48
- Параметры математической модели 8
- «Пила» 151
- Погрешности коэффициентов и параметров математической модели и уравнений 6
- Погрешность исходных данных (оператора, правой части, граничных условий) 10, 159, 164, 168
- решения 10
- Полином 12
- Помеха 94
- Порядок регуляризации q 167, 201, 209
- Правая часть 148, 168
- Предсказание будущего 100
- Преобразование Лапласа 136
- Радона 206
- Фурье (ПФ) 134, 151, 167
- быстрое (БПФ) 167, 193, 212
- двумерное 135
- дискретное (ДПФ) 152, 167, 193, 212
- непрерывное (НПФ) 134
- обратное (ОПФ) 135, 151
- Хартли (ПХ) 136
- Преобразования, эквивалентные в классическом смысле 42
- в расширенном смысле 46
- Прикладные задачи 188, 198, 204, 212
- Примеры реконструкции изображений 194–197, 203
- Программы 164, 167, 212
- Пространства C, L_1, L_2, W_2^1 и др. 119, 144
- Пространство банахово, гильбертово, евклидово, линейное, метрическое, нормированное и т.д. 116–121
- Псевдообратная матрица Мура–Пенроуза 129, 156
- Псевдорешение 128, 153, 173
- Ранг матрицы r 125, 172
- расширенной матрицы ρ 127, 172
- Распределение тока вдоль обмотки катушки томографа $J(a)$ 214
- Регулируемые переменные 45
- Регуляризация 14, 18, 25, 159
- «интуитивная» 208
- локальная 208
- Регуляризирующий алгоритм (РА) 160
- оператор (РО) 160
- Регуляризованное решение 167, 171, 174, 191
- уравнение 177
- Регулярные методы 116
- Регулятор 20, 24, 44
- оптимальный 21, 23, 25, 27, 37
- Реконструкция дефокусированных изображений 188, 198
- рентгеновских изображений 207, 211
- смазанных изображений 146, 188
- Решение неустойчивое 131, 141, 145, 150, 152, 166
- нормальное 156, 173
- регуляризованное 167, 171,

174, 191
 – точное 151, 164
 – устойчивое 46, 131, 138, 147, 160, 166
 Свертка 134, 147, 151, 166, 170
 Сглаживающий функционал 160, 166, 170, 173
 Сетки узлов по s, t, x, ω 149, 163, 180
 Сингулярное число μ 125, 140
 Синтез магнитного поля в ЯМР-томографе 212
 – систем управления 19
 Синус-преобразование Фурье 135
 Система линейных алгебраических уравнений (СЛАУ) 129, 149
 – – – – вырожденная 129, 172
 – – – – недоопределенная 127, 139, 155, 172
 – – – – нормальная 128, 153
 – – – – однородная 28
 – – – – определенная 127, 139, 172
 – – – – переопределенная 127, 139, 153, 172
 – – – – плохо обусловленная 10, 128, 129, 172
 – обыкновенных дифференциальных уравнений (ОДУ) 32, 130, 141
 – – – – с управляющей функцией 131, 174
 Системы управления 54
 Скалярное произведение (x, y) 120
 След матрицы Sp или Tr 124
 Смаз (сдвиг, смещение) 190
 Смазанное изображение 146, 188

Собственное значение (число) λ 29, 122, 125, 140
 Собственный вектор (собственная функция) оператора 122
 Совокупность интегральных уравнений 191
 Спектр возмущающих воздействий 19, 88
 – – – – дробно-рациональный 19
 – матрицы 125
 – мощности 135
 – оператора 122
 – Рахманина-Фирсова 23, 29
 – Фурье 135
 Спектральная плотность мощности 80
 Способы выбора параметра регуляризации α : невязки, обобщенной невязки, подбора 162, 192, 201
 Среднеквадратическая погрешность 147, 164, 168
 Суда и системы судовой автоматики 70
 Существенно некорректная задача 160
 Существование решения 138
 «Телеграфный сигнал» 77
 Теорема о непрерывной зависимости решений ОДУ от параметров 32
 Теоремы о существовании, единственности и устойчивости решения системы ОДУ 130
 Теория оптимального управления 19–22
 Типы уравнений 129
 Томограф рентгеновский 206
 – ядерно-магнитно-резонансный 212

- Томография 204
 - рентгеновская 204
 - ядерно-магнитно-резонансная 212
- Третий класс задач 41
- Трехпараметрический спектр 93
- Узлы дискретизации по s , t , x , ω 168
- Умножение матриц и векторов 127
- Уравнение дифференциальное 19
 - в частных производных 176, 192
 - Лапласа 132, 142, 176
 - теплопроводности 132, 143, 176
 - интегральное 133
 - Вольтерра 134, 145
 - двумерное 134, 200
 - одномерное 133
 - Радона 206
 - регуляризованное 177
 - типа свертки 134, 200
 - Фредгольма 133, 145
 - операторное 134, 138, 159
 - характеристическое 124
 - Эйлера–Тихонова 161
- Условие Липшица 130
- Условия граничные 132, 177
 - начальные 132, 141, 177
- Устойчивость решения 20, 130
- Устойчивые методы 114, 162
- Формула Барлинга 91
 - квадратурная трапеций 149, 163
 - Эйлера 135
- Функционал 117
 - сглаживающий 160, 166, 170
 - стабилизирующий 175
- Функция дробно-рациональная 19
 - искомая 148
 - корреляционная 73
- Фурье-образ 134
- Характеристика направленности антенны 165
- Характеристический полином 32
- Характеристическое уравнение 124
 - число 125
- Частота Фурье 135
- Частотная характеристика (ЧХ) спектрометра 166
- Численные примеры 9, 15, 164, 168, 182, 186, 202, 211, 219
- Число обусловленности cond 128, 140
 - сингулярное 124
 - узлов 149, 164, 168
 - характеристическое 125
- Шаг дискретизации 152
- Шум 197
- Эквивалентные преобразования 42, 147
 - (равносильные) системы уравнений 42
- Экстремальные задачи 13
- Электропривод постоянного тока 49
- Эффект Гиббса 169
- Ядро интегрального уравнения 133, 148
 - разностное 168
 - фредгольмовское 133
- ЯМР-изображение 213
- ЯМР-томограф 212
- ЯМР-томография 212

УЧЕБНОЕ ИЗДАНИЕ

Петров Юрий Петрович
Сизиков Валерий Сергеевич

**КОРРЕКТНЫЕ, НЕКОРРЕКТНЫЕ
И ПРОМЕЖУТОЧНЫЕ ЗАДАЧИ
С ПРИЛОЖЕНИЯМИ**

Оригинал-макет подготовлен авторами

Подписано в печать 16.12.12.
Электронных текстовых данных 65,5 Мб
Электронный текст подготовлен ОАО «Издательство «Политехника»».
191023, Санкт-Петербург, Инженерная ул., д. 6.
www.polytechnics.ru